Teaching and subjects on bio-medical engineering

Approaches and experiences from the BIOART-project

This book originated from the BIOART-project "Innovative Multidisciplinary Curriculum in Artificial Implants for Bio-Engineering BSc/MSc Degrees (586114-EPP-1-2017-1-ES-EPP-PKA2-CBHE-JP)".

The work groups a number of selected chapters and summaries of papers on topics from the new developed curricula. The book is intended to be an educational book, outlining the methodology in the curricula BIOART, its introducing in the consortium to build new robust and modern curricula to prepare engineers for the bio-medical fields.

Key features:

- Introduction on the quality and accreditation systems in the consortium countries.

- Section "Educational approaches" on the educational background and suggested educational methods for the bio-medical engineering fields.

- Section "System design of artificial implants" on the principal elements and methods of how bio-artificial systems are designed and of which general building blocks they are composed of.

- Section "Materials for bio-medical engineering applications" on the use of different materials for bio-medical engineering including biomaterials/smart materials and the according manufacturing and production technologies.

- Section "Applications and case studies" describes different applications in the bio-medical engineering fields and are as such exemplary for the different topics of the other chapters and serve as reference for implementation of bio-medical systems.

Printed by Acco cv, Leuven (Belgium).

ISBN 978-94-641-4233-4

TEACHING AND SUBJECTS ON BIO-MEDICAL ENGINEERING

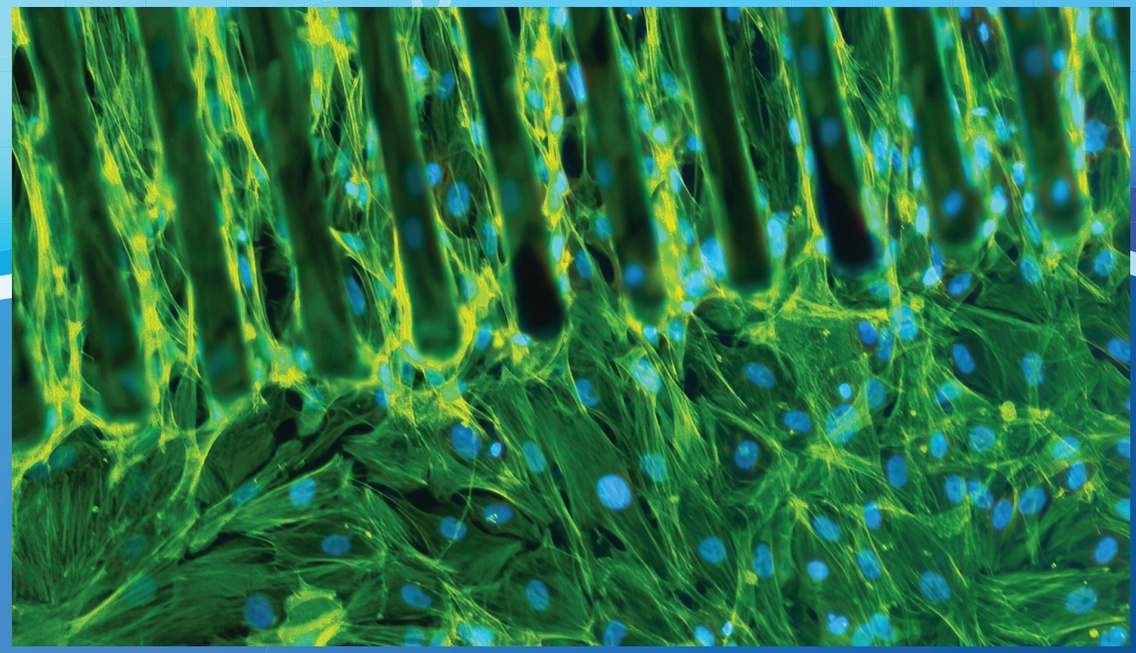
Co-funded by the Erasmus+ Programme of the European Union

BIOART Engineering education for human welfare

PETER ARRAS AND DAVID LUENGO (EDS.)

TEACHING AND SUBJECTS ON BIO-MEDICAL ENGINEERING

APPROACHES AND EXPERIENCES FROM THE BIOART-PROJECT



2021

Teaching and subjects on bio-medical engineering

Approaches and experiences from the BIOART-project

Teaching and subjects on bio-medical engineering

Approaches and experiences from the BIOART-project

Peter Arras and David Luengo (Eds.)

Printed version ISBN: 978-94-641-4233-4

Epub ISBN: 978-94-641-4245-7

Keywords: Bio-medical Engineering, BIOART, implants

Cover Design: Andreiev M.O.

Cover photo: Picture by Rosalba M. Ferraro & Miriam Seiti, AML group KU Leuven in partnership with UniBs: “Cell Culturing on Patterned Substrates for Morphological Stimulation in Neural Tissue Engineering”.

First edition: 2021

The BIOART-project (Innovative Multidisciplinary Curriculum in Artificial Implants for Bio-Engineering BSc/MSc Degrees), 586114-EPP-1-2017-1-ES-EPPKA2-CBHE-JP, has been funded with the support of the Education, Audiovisual and Culture Executive Agency of the European Commission.

The European Commission’s support for the production of this publication does not constitute an endorsement of the contents, which reflect the views only of the authors, and the Commission cannot be held responsible for any use which may be made of the information contained therein.

Preface

This book originated from the BIOART-project “Innovative Multidisciplinary Curriculum in Artificial Implants for Bio-Engineering BSc/MSc Degrees (586114-EPP-1-2017-1-ES-EPPKA2-CBHE-JP)”, a European subsidized educational project aimed at adapting, modernising and restructuring existing curricula in artificial implants for bio-engineering degrees in Ukraine and Israel, by developing new courses, testing the innovative curricula and disseminating the results.

The work groups a number of selected chapters and summaries of papers on topics from the new developed curricula. The contributions are made by the different partners of the BIOART-consortium and often in a collaborative coproduction between partners to combine their expertise. The book is intended to be an educational book, outlining the methodology in the curricula BIOART, its introducing in the consortium to build new robust and modern curricula to prepare engineers for the bio-medical fields.

Seeing the evolutions in medical and bio-medical sciences and techniques over the last decade on the technological fields as electronics and data-analysis, material sciences and mechanical structures it is clear that engineers prepared to work in the high-tech world of bio-medical engineering need a very broad set of skills and competences.

From the vast range of bio-medical applications BIOART focusses on the design and use of artificial medical implants. An artificial implant is a man-made medical device manufactured to replace a missing biological structure, support a damaged biological structure/process , or enhance an existing biological structure/process.

The use of medical implants has been rising for many years thanks to improved knowledge on materials, electronics and new manufacturing methods. As such it is also an absolute necessity that study cycles aimed at preparing engineers for the design, manufacturing and maintenance of these implants follow the same pace of the state-of-the-art and are in need of new structured contents and methods.

The book gives an overview of the methods and topics the consortium implements in the updated curricula and introduces the different aspects to undergraduate and graduate level engineering students. Yet, the intended use of this book is not confined to a specific course in the bio-medical engineering curriculum but can also be used as a general introduction and overview to artificial implants, touching the specific areas on which the BIOART-project focusses. Furthermore, the different chapters can also be extracted and used in specialized courses of the curricula both as broadening knowledge as well as deepening knowledge on specific topics.

The book starts with an introduction on the quality and accreditation systems in the consortium countries for new and updated curricula.

The first section ”Educational approaches” is on the educational background and suggested educational methods for the bio-medical engineering fields.

A second section “System design of artificial implants” focuses on the principal elements and methods of how bio-artificial systems are designed and of which general building blocks they are composed of.

A third section “Materials for bio-medical engineering applications” is on the use of different materials for bio-medical engineering including biomaterials/smart materials and the according manufacturing and production technologies.

Finally, the last section ”Applications and case studies” describes different applications in the bio-medical engineering fields and are as such exemplary for the different topics of the other chapters and serve as reference for implementation of bio-medical systems

Acknowledgments

A book of this kind can only be made with the help, advice, criticism and ideas from many people.

Many colleagues of the BIOART consortium have been willing to share their expertise and thoughts and devote their time to write a chapter for this book as a help to students in bio-medical engineering.
We gratefully acknowledge the team of international reviewers whose valuable work and comments helped to complete this book.

Editors:

Editorial board:

List of authors

Aikin Nikita , National University “Zaporizhzhia Polytechnic”, Ukraine
Altukhov Alexander , Donbass State Engineering Academy, Ukraine
Arras Peter , KU Leuven, Belgium
Avrunin Oleg , Kharkiv National University of Radio Electronics, Ukraine
Azarkhov Oleksandr , Pryazovskyi State Technical University, Israel
Bardea Amos , Holon Institute of Technology, Israel
Benavente C., Universidad Politécnica de Madrid, Spain
Benis Arriel , Holon Institute of Technology, Israel
Bezsmertnyi Yurii , Vinnytsia National Technical University, Ukraine
Burova Dariya , Priazovskyi State Technical University, Ukraine
Chabak Yuliia , Pryazovskyi State Technical University, Ukraine
Cheiliakh Oleksandr, Pryazovskyi State Technical University, Ukraine
Chekhmestruk Roman , Scientific Research Institute of Invalid Rehabilitation on the base of Vinnytsia Pirogov , Ukraine
Cherniavskyi Anton , Donbass State Engineering Academy, Ukraine
Cheylyakh Yan , Priazovskyi State Technical University, Ukraine
Chorniy Vadim , Zaporizhzhia State Medical University, Ukraine
Dobriak Serhii, Donbass State Engineering Academy, Ukraine
Efremenko Vasily , Pryazovskyi State Technical University, Ukraine
Efremenko Bohdan , Pryazovskyi State Technical University, Ukraine
Ferraris Eleonora , KU Leuven, Belgium
Gibney Rory , KU Leuven, Belgium
Gladkova Olga , National University “Zaporizhzhia Polytechnic”, Ukraine
Goldfrad Keren , Bar Ilan University, Israel
Gribkov Eduard , Donbass State Engineering Academy, Ukraine
Harpak Katarzyna , Cracow University of Technology, Poland
Hrushko Oleksandr , Vinnytsia National Technical University, Ukraine
Kapliienko Tetiana , National University “Zaporizhzhia Polytechnic”, Ukraine
Korniejenko Kinga , Cracow University of Technology, Poland
Kovalenko Andrii , Donbass State Engineering Academy, Ukraine
Łagan Sylwia, Cracow University of Technology, Poland
Liber-Kneć Aneta , Cracow University of Technology, Poland
Luengo David , Universidad Politécnica de Madrid, Spain
Mak-Mak Natalia , Priazovskyi State Technical University, Ukraine
Malysheva Inna , Priazovskyi State Technical University, Ukraine
Meltzer David , Universidad Politécnica de Madrid, Spain
Mikula Janusz , Cracow University of Technology, Poland
Mykhaylov Pavlo , Vinnytsia National Technical University, Ukraine
Nestawal Stephanie , Danube University Krems, Austria
Nosova Yana , Kharkiv National Medical University, Ukraine
Osés D., Universidad Politécnica de Madrid, Spain
Parkhomenko Anzhelika , National University “Zaporizhzhia Polytechnic”, Ukraine
Parkhomenko Andriy , National University “Zaporizhzhia Polytechnic”, Ukraine
Pavlov Sergii , Vinnytsia National Technical University, Ukraine
Petrova Olga , National University “Zaporizhzhia Polytechnic”, Ukraine

Table of contents

Introduction

Quality Assurance in Multinational Curriculum Development: the role of the Bologna Process. Introduction.

Keren Goldfrad[1][0000-0003-3411-8018], Stephanie Nestawal [2][0000-0001-7317-1482], Galyna Tabunshchyk [3][0000-0003-1429-5180]

[1] Bar Ilan University, Ramat Gan
[2] Danube University Krems, Dr.-Karl-Dorrek-Str. 30, 3500 Krems, Austria
[3]National University "Zaporizhzhia Polytechnic", Zaporizhzhia, Ukraine

Abstract. The aim of this chapter is to describe and discuss the aspect of quality assurance of the Bologna policies in higher education in context of implementation of the Erasmus+ CBHE projects. The authors consider applications of the Bologna policies from the EU, UA and IL perspectives as it was experienced during BIOART project lifetime.

Keywords: Bologna Process, quality assurance, curriculum development.

1 Quality Assurance and the Bologna Reform

The European Union has been developing policies in order to increase the quality of higher education, with the Bologna Process being one of the most important developments in this regard. Since its implementation, the Bologna Process has influenced developments in quality assurance and development at European higher education institutions with the "Standards and Guidelines for Quality Assurance in the European Higher Education Area" (ESG) eventually being adopted in 2005. The ESG are deemed as the defining principles and over the years have become the main basis for the design of internal and external quality assurance procedures and -systems that have developed within higher education.

The principles of quality assurance and quality development, however, have already emerged and developed as central themes in the political and scientific debate in most European countries in the 1980s. Although the reasons for this vary from country to country, the Agency for Quality Assurance and Accreditation in Austria identified three transnational trends in their report from 2015:

1) A strong expansion of higher education becoming evident during the 1970s and 1980s with the proportion of people with a university degree in a particular age cohort multiplying.

2) The manifestation of quality problems arising from structural changes triggered by the expansion of higher education and followed by the disproportionate provision of financial resources.

3) The deregulation of higher education systems and the emerging of "New Public Management" combined with a growing accountability and obligation on the part of

the universities and the emergence of quality assurance instruments and introduction of actors in the governance of higher education institutions and higher education systems. [1]

These developments formed an important framework for the statutory obligation to carry out internal university quality assurance measures on the one hand and to establish external quality assurance systems and corresponding quality assurance agencies.

Since the signing of the Bologna Declaration in 1999, a great deal of progress has been observed in the development of the EHEA. The essential and exceptional element of the Bologna Process is the voluntary rapprochement rather than a binding agreement. This leaves it up to individual countries to tailor the realisation of the visionary EHEA to their national circumstances, rather than implementing it at the lowest common denominator.

While some of the objectives of the Bologna Process are specific measures, others are more like declarations of intent. Some of these concrete actions and instruments already existed before the Bologna Reform: the ECT System and the Diploma Supplement, came into being under the Lisbon Recognition Convention (1997) of the European Council and UNESCO. The unique feature of the Bologna Process therefore is its consolidation function: it bundles policies and develops them further, draws up guidelines on how policies are to be implemented and enables the monitoring of the pursuit of agreed objectives through institutionalized communication, applying benchmarking measures and information exchange in transnational and national policy networks. At the same time, the "Bologna model" leaves sufficient scope for the participating countries to shape higher education policies according to national circumstances and policy preferences. [2] [3]

It is suggested that the acceptance and popularity of the Bologna Reform is based on the very fact that it primarily focuses on "product control" and that central coordination mechanisms are voluntary. There is no legal framework which could enforce the implementation of the jointly agreed reforms in the participating countries. This also explains the uneven implementation across different policy areas and across participating countries (Paris Communiqué 2018). [4] After 20 years, the EHEA resembles a patchwork of 48 different higher education systems that have implemented similar structural reforms at the macro level but with limited compatibility at the degree and program level (BFUG Working Group 2 2018). [5] Implementation deficits are also related to the fact that political leaders in most countries who have negotiated on the nature and scope of policy changes for higher education are not explicitly responsible for how universities teach, how scientists work, how students learn, or how employers recruit.

[6] Deficits may also be due to a lack of interest among stakeholders at the national level in implementing policies in whose formulation they have not or have not sufficiently been involved. Perhaps the sole idea of a unified European Higher Education Area, in which higher education policies and programs are fully integrated, comparable and compatible, are a utopia. It is also utopian to assume that policies that are defined "top-down" will be implemented one-to-one by subordinate but more or less autonomous institutions. [7] [8]

2 The Bologna Process in Israeli institutions of higher education

Israeli institutions of higher education have a complex "relationship" with the Bologna Process. There is a strong drive for Israeli universities to attract students from around the world and in that respect become more international. This notion of internationalization has generated a very positive attitude towards the Bologna Process, which in turn led Israeli policy makers to request to join this initiative and become full members. The two official requests that were filed in 2007 and 2008 were denied, but Israel was granted observer status [15].

Despite this denial, Israel is striving to adhere to the three main objectives of the Bologna Process, namely: the three-cycle degree system, a mutual recognition of qualifications and learning periods, and a quality assurance system [11]. The first objective, the three-cycle degree system, has been present in the Israeli higher education system since the 1950s. The second and third objectives posed a challenge for the Israeli Council of Higher Education (CHE). As of now, the national qualification framework has not been implemented, but quality assessment and quality assurance of higher education are regarded as a national priority.

The CHE is situated under the authority of the Minister of Education and serves as the only regulator of quality assurance in HE. Accreditation of new degrees and programs have to be approved by the Quality Assessment and Assurance Unit, which was established in 2004. Although this unit oversees the degrees and programs of all HEIs, its goal is to promote the development of internal mechanisms for quality assurance within each institution. The intention is to build a culture of self-assessment, and indeed, there were many local initiatives within individual universities and colleges to create an internal system of quality assurance. These initiatives were sporadic and not uniform since there were no authoritative guidelines from above. To that end, the CHE nominated a quality assurance committee in 2013. Its purpose was to report on the quality of teaching in the framework of higher education and set the stage for national quality assurance regulations and procedures.

The survey that the quality assurance committee conducted among all the institutions of higher education in Israel revealed that Israeli academics associate quality assurance directly with teaching and learning related aspects. Based on the committee's report, the CHE formulated a list of criteria that all universities and colleges must implement to receive government funding. In order to fulfil these criteria, each institution should establish a mechanism which oversees issues relating to quality assurance, teaching and pedagogy. Such a mechanism should include professional pedagogical training of the teaching faculty, ongoing teaching surveys, internal system advancing lecturers with low survey scores, institutional department receiving student complaints, and publication of detailed syllabi. These regulations set into motion a chain of policies and procedures within each institution.

As mentioned earlier, many universities and colleges commenced a few local initiatives prior to these CHE regulations. One such initiative was students' surveys which focus on their evaluation of quality of teaching, or their satisfaction from the level of

teaching. Another initiative that was instituted in many Israeli institutions is the outstanding teaching award for lecturers who have excellent quantitative and qualitative survey responses, and who contribute to the teaching effort within their departments. The exact selecting process and the number of lecturers chosen varies from institution to institution. The national student union began a similar process on a national level and is now giving awards to inspiring lecturers from HEIs around the country.

One of the major initiatives that began in only a few institutions before the CHE regulations and gained considerable weight over time are the Teaching Enhancement Centers (TEC). These centers are foremost in charge of providing faculty members with pedagogical training and support in order to promote their teaching skills. This support consists of a number of services. The centers provide individual counselling to any academic instructor on campus. Instructors are observed teaching, and then receive individual feedback from a pedagogical expert. Every year more and more researchers request this service to improve their teaching strategies and pedagogical skills. The teaching enhancement centers train new lecturers during their first year in an intensive two/three-day workshops. The centers also provide professional workshops on pedagogical issues such as increasing motivation among learners, creating effective PowerPoint presentations, using digital tools to increase student engagement, writing exams, conducting project-based learning activities, and building synchronous and asynchronous active learning units. In the context of this paper, it is important to note that quite a few institutions provide faculty with a workshop that focuses on how to write a syllabus according to the Bologna Process guidelines.

In addition to these two basic services, personal support and professional workshops, each center has other local activities. The Teaching Enhancement Centre at Bar-Ilan University (BIU) established in 2011, for example, has been conducting in-depth interviews to 3rd year students in each individual department in order to receive more information regarding students' satisfaction level from their learning experience at BIU. These interviews, which began a few years ago, usually consist of six to ten students who are interviewed as a group by a member of the TEC team. The students' identities are concealed, so that they feel comfortable relating their feelings and learning experiences. Students are also asked whether they feel they have acquired the skills necessary for the labor market. A written report is produced, and two TEC members go over the report with the head of the department. Each department chair has the authority to decide which points should be addressed and what needs to be done within the department. In many cases, certain lecturers are asked to participate in certain pedagogical workshops.

Teaching Enhancement Centers in HEIs have been mandated by the Council of Higher Education, but on a voluntary basis, there is an unspoken solidarity where more established centers are helping and guiding newer centers. These collegial collaborations are facilitated by the Israeli National Forum for Teaching Enhancement Centers which was created in 2002. All the heads of the TECs in Higher Education Institutions are members of this forum. The forum holds meetings, workshops, and even organizes international group trips so that its members have transnational connections and learn from one another. These trips generated not only worldwide friendships, but also strengthened collaboration among the forum members which proved very beneficial

during the COVID19 crisis. Since the beginning of the crisis, forum members started functioning as a support group and facilitated the sharing of strategies and ideas between institutions to further enhance the teaching level in HEI. As Adelman (2009) points out, "the Bologna Process is an analogue to the macroeconomic theory of convergence, the ways in which nations move from different stages of development to a more-or-less common platform of performance".

In general, the Bologna Process had a positive impact on HEI in Israel. First, it heightened the awareness that there is a need for common grounds on which to base mutual recognition of academic courses on the national and international levels. This awareness led to the understanding that the structure of syllabi must be standardized, and that syllabi should become more "student-centered" and less "teacher-centered". As mentioned earlier, workshops are conducted in many institutions to train lecturers to write syllabi based on the recommendations of the BP and particularly the learning outcomes approach. Moreover, international mobility and mutual recognition demands that all syllabi are translated to English, and "speak" a common language.

The Bologna Process also emphasized the necessity of a quality assurance mechanism on a national level through the Council of Higher Education and within each institution. At BIU each department has a teaching committee that goes over and approves the syllabi within the department. These syllabi are then sent to the faculty teaching committee which in turn are forwarded to the Vice Rector's office for approval by the institution's central teaching committee. On a national scale, the CHE conducts teaching quality assessments across the disciplines. Each year several disciplines are chosen by the quality assurance committee within the CHE to examine the quality of teaching in all the HEIs according to a specific discipline. The CHE sets up an international professional team of experts which evaluates the discipline-specific departments in the country. This process led to revisions of the curricula, to more standardized qualification frameworks and to a common credit system.

Despite the apparent advantages and benefits brought about by the Bologna Process, it is important to mention the challenges that it poses to creativity and innovation. Some researchers, such as Teelken and Wihlborg [14], point out that the BP negate the notions of academic freedom, autonomy, and independence. Indeed, one can argue that governmental policies and regulations striving for comparability place more emphasis on assessment and control and neutralize the competitive aspect between different programs and institutions. In this way, the inherent goal of the BP of quality assurance, which leads to standardization and uniformity, may decrease diversity and uniqueness among academic programs.

Whereas European institutions of higher education are mostly struggling with the notion of the three-cycle degree system, their Israeli counterparts are grappling with the external influences of marketization and the consumerism brought about partly by the BP's student-centered attitude. Accountability and quality assurance procedures also force universities to comply with a consumerist approach which take into account the demands of multiple stakeholders including students and employers [12]. This external pressure is perceived by many faculty members as an unwelcome intrusion on their academic freedom.

Taken as a whole, the Bologna Process enabled a process of international collaboration and harmonization between different countries within and outside the European Union. It forced us all to "speak the same language" and use comparable quality assurance frameworks. Hopefully, this movement of convergence will create a "zone of mutual trust" [10], while preserving the program or institution's independence, autonomy and academic freedom.

3 The Ukrainian perspective on the Bologna process

The Bologna process started in Ukraine on the 19th of May 2005, with the introduction of an action plan for quality assurance in Ukrainian Higher Education and integration into the European Educational Network by 2010. In 2014 the new national standard in Higher Education was accepted. This new standard became the new driving force for the implementation of the Bologna process.
In 2017 a new Law of Education was accepted which in general is compatible with the Bologna Process.

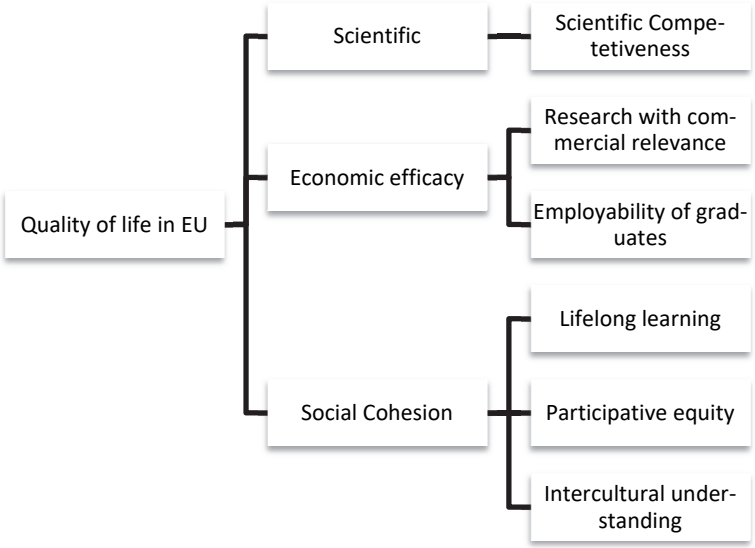


Fig. 1. Primary goals and levels of the Bologna System Model [16]

Let us consider the results of implementation of the primary levels of the Bologna system model (**Fig. 1**) in Ukrainian HEIs.
To comply with the Bologna goals an improved national qualifications framework (NQF) was put into place.
The NQF includes:

- 11 levels of qualification;
- 4 cycles in higher education: Junior Bachelor (short cycle), Bachelor-Master-PhD and 4th cycle – Doctor of Science)

- Implementation of the European approach to assure equal access to education and inclusive learning;
- Implementation of national regulations to support mobility;
- Implementation of a life-long learning approach;
- Widen university autonomy.

3.1 Quality assurance process and accreditation of curricula.

One of the important steps was done in establishing quality assurance. The European experience was studied, and the National Agency for Quality Assurance in Higher Education was created. «The mission of the National Agency for Higher Education Quality Assurance (NQAA) is to catalyse positive changes in higher education and to shape its quality culture»[16,17].

In the new Educational Law, announced more autonomy for HEIs. One of the important steps which all Ukrainian universities should comply with passing accreditation for its degree studies, is that each HEI should implement internal quality assurance rules and regulations [18].

In Ukraine the HEI departments are responsible for the implementation of the new curricula and for the organization of the accreditation. The list of possible curricula is set on the governmental level with standards, which contains required competences, skills and program results. According to the Educational law there should be a strong correlation between educational standards and new curricula and NQF.

The implementation of a completely new curriculum such as the one developed within the BIOART project is possible only as a specialty within an existing one, as there are national standards with the list of possible specialties. This structure of defined specialties makes it more difficult to start with new interdisciplinary degree studies, as from its nature, multidisciplinary studies cannot be put under one specific specialty.

University autonomy allows universities not to pass national accreditation and produce local diplomas of the university. We conducted research among local businesses and despite the fact that the main factor is the authority of the university all suppose that it is more convenient to employ graduates from curricula which passed national accreditation.

The process of national accreditation for new studies approximately takes three months and consists of the following steps []:

Step 1: The electronic application is signed by the rector of the HEI: this opens the request for an accreditation and confirms that the university itself supports the idea of the new curriculum.

Step 2. The guarantor of the curricula prepares a self-assessment and apply it in electronic system. The guarantor is the professor who is in charge of the department or is the main expert in the proposed study-field and as such guarantees the academic quality of the contents of the curricula. The self-assessment report describes how the HEI will organize the new curricula/or existing curricula and proofs/states to the accreditation agency that curricula-changes/renewals is necessary (from national and international

benchmarks). It also explains how academic and pedagogic quality will be achieved and how the curricula comply with the laws (and the Bologna goals).

Step 3. The accreditation agency prepares the process of the accreditation and selects experts (accreditation committee) who will conduct an assessment.

Step 4. The accreditation committee which consists of three experts with one student among them analyses the self-assessment report as a basis for the on-site audit.

Step 5. The onsite audit visits the HEI to check the facts stated in the self-assessment report and to interrogate staff and students to assess how the curricula is implemented and to see what the strong and weak points are. (Currently, this process is conducted on site and online due to the pandemic situation.)

Step 6. The experts prepare their conclusions. And the university analyses the conclusions and prepares its feedback.

Step 7. Expertise by committee: Conclusions.

Step 8. Final decision.

The result of the accreditation process could be:

1. Passed. The result is then published on the website of the NQAA and the HEI receives the accreditation certificate.
2. Rejected. University could apply for appeal
3. Conditional accreditation for one year - students who complete their studies and successfully defend their BA or MA thesis receive a governmental diploma. But such curricula requires new accreditation in one year.

3.2 Key performance indicators for HEI's in the quality process

According to the NQAA the most important key performance indicators for a high-quality educational system are based on international standards. The ranking in the international rating systems, the citation criteria (h-index) of the educational staff, number of graduates employed according to the diploma (labor market, world tendencies, scientific research), which highly correlated with the Bologna System model (fig.1).

To reach these standards in NU "Zaporizhzhia Polytechnic" there is conducted annually a self-assessment of pedagogical and scientific staff. The assessment includes formal categories: general staff characteristics, educational work within the reporting period, scientific activities, publishing activities, organizational activities with students.

The progress rates of the departments also include the number of publications, citation index, research grants and income in the field of teaching and learning, national and global awards and prices, projects.

The results of this self-assessment are used by university management for motivation.

A self-assessment methodology and software for it was developed by the educational department in cooperation with the software tools department. As all of this is new to the university, the tools are being improved and extended to make it easier in use, and more performant in outcomes.

For curricula measures the number of graduates, the success rate in finished degrees, and the employability success rate of graduates are the main measuring tools.

University standard for the new master's curricula which was implemented in NU "Zaporizhzhia Polytechnic" in the frame of BIOART is organized in the following way: general cycle (9ECTS), professional cycle 51 ECTS (disciplines and internship) and scientific part as master thesis 30 ECTS.

The compliance to the goals of intercultural understanding is mainly a matter of internationalizing the degree studies. For this a major measuring tool is the mobility numbers for students and staff to and from the university.

Erasmus+ CBHE projects in cooperation with Erasmus+ KA1 help Ukrainian universities to support international mobility of staff and students. Within the BIOART project 77% of educators of NU "Zaporizhzhia Polytechnic" involved in the new curricula had a possibility to participate in internships in European universities and in training abroad. Students participate in KA107 with Polytechnic in Madrid for one semester. Universities are still struggling with the implementation of individual educational courses and academic recognition of the results, which has a negative influence on the number of students who are participating in international mobility as its increase load on them. The rigid national accreditation system with its standards and defined contents make it difficult for students to make a selection of courses in a host-institution. According to the system only copies of the Ukrainian subjects are accepted, but this is of course far from reality. The consequence is that mostly acquired credits abroad are not automatically converted to acquired credits at the home university.

The BIOART project also is a support for the sustainability goals by funding equipment to help universities to reach the program results. In NUZP this funding led to the development of two new laboratories at two departments involved in the new curricula.

For supporting inclusive education, the university system for distance learning is widely used. The system contains all teaching materials. The Software Tools department has great experience in the implementation of the remote laboratories, which give students the possibility to have access to the equipment 24/7 [20-22]. These possibilities are also "exported" to the community stakeholders and as such is used for in-company training and for kids-university (course and summer school activities for pupils of secondary schools) and for student competitions like BEST (Board of European Students of Technology).

For the Ukrainian partners, the implementation of new curricula within the international CBHE projects has great advantages: – all syllabi are developed in cooperation with and within and international network of project partners, HEI's can invest in equipment in modern laboratories, and there is an international forum and outreach for staff and students.

However, a big challenge remains: the implementation of the multidisciplinary master curricula - as described in BIOART. When establishing new curricula, you should first of all reach the programs results, which are described in the national standard. This means that it causes more self-study for the students when we are implementing interdisciplinary.

Since students with different subject backgrounds can be enrolled in the new master program to cover the gap, we are limiting the list of such specialties and set elective

blocks in the bachelor curricula. Unfortunately, this sometimes causes an overload for the students.

As the general term for master studies in Ukraine is 1.5 year (90 ECTS equivalent), the best solution for multidisciplinary curricula would be to increase it to 2 years (120 ECTS equivalent).

Another open question is how to check if the program results were reached. Nowadays to answer this question, we are using students' questionnaires, but in general student's activity/motivation to study is very low. Generally conducted research [23] on students understanding their carrier strategy only 55% have a clear idea about it (15% its best graduates and 40% those who are oriented towards success in their professional activities), 30% are so called "disadaptants" and 15% are unmotivated student youth with vague plans for professional growth. Involvement students into the research national and international projects, internship at the national and international companies aimed to increase the number of motivated students.

4 Implementation of the main targets of the Bologna Reform in Austria

Austrian universities have grown rapidly between 1999 - the year in which the Bologna Declaration was signed - and 2018. During these two decades, universities have implemented study law reforms within the framework of the Bologna Process - in addition to a large-scale organizational reform. Those reforms - amongst others - included the implementation of the three-cycle degree architecture and the introduction of comprehensible and comparable academic degrees. Between 1999-the year in which the Bologna Declaration was signed-and 2018, Austrian universities have expanded rapidly. In addition to the large-scale organizational change, universities have adopted study law changes in the sense of the Bologna Process over the last two decades. Those changes included the adoption of the architecture of the three-cycle degree and the introduction of comprehensible and equivalent academic degrees, among others.

As of 2020, the multi-tier study architecture is already well implemented in Austria, but processes such as the emphasis on learning outcome, manageable study duration and mobility windows still need further development and continuous improvement.

The implementation of the Bologna process in Austria initially was not easy and a lot of it was based on trial and error. Nevertheless, the country demonstrated a strong commitment to the Bologna Process reforms from the earliest years, taking an organized, methodical and pragmatic approach to implementation. A national "Bologna Follow-Up Group" was established involving the higher education institutional representatives and students' union, and all institutions appointed a Bologna Coordinator. Degree structures were radically changed – moving away from the old diploma system to the Bologna cycles. The speed of implementation varied considerably according to the area of studies, with natural science programs moving first to the Bologna cycles and social sciences and humanities taking longer.

It is argued that one of the main drivers that sped up the implementation process was the internationalization within the sector of Universities of Applied Sciences. These institutions promoted the intensive examination and analysis of the Bologna Process and its associated challenges: leaner content and improved competence development, implementation of qualification profiles, integration of mobility windows, transparency, ECTS and the focus on learning outcomes.

Nevertheless, the lack of experience with the multi-cycle system and the consideration of what constitutes a bachelor's or master's degree program, led to uncertainties and misunderstandings. Hence resulting in misinterpretations and undesirable developments such as crowded curricula, purely consecutively designed study programs, little room for mobility or a structurally driven allocation of ECTS without acknowledging the content, the level or learning outcomes.

Learning outcomes and the associated paradigm shift were and still are the biggest challenges, as they question the traditional approach to teaching and learning and the role of the teacher. The idea of purely consecutive study courses and the indirectly associated preservation of the old diploma study programs soon proved to be a misjudgment, as did the belief that one's own bachelor's graduates could only participate in their own consecutive master's programs. Students very soon recognized the advantage of the Bologna system and the associated individualization of their educational path and sought variety and focus for their second cycle. This also brought up the question of permeability, combined with the question of recognition and equivalence.

The complex interplay of different reforms and framework conditions becomes evident when the effects of the Bologna Process are viewed from the students' perspective: they now have to complete a three-year, tightly structured course of study, in which the pressure to perform has increased while opportunities for combining work and study have been reduced. Furthermore, at the end of their studies, students find themselves without a vocational qualification because their degree is not recognized on the labor market. The consecutive master's program seems indispensable and has in fact replaced the traditional diploma degree by becoming the unofficial "first degree".

It could be argued that the acceptance of Bologna conform first degrees would grow if the development of study programs, as a consequence of the Bologna process, would finally aim at a rigid replacement of content centricity in favor of outcome – orientation. This approach indeed involves many challenges and requires a theory-based and systemic approach. The Constructive Alignment Concept according to Biggs and Tang (2011) provides a suitable basis for such a competence-oriented study program development. [10]

The lack of understanding of core concepts and practices of the Bologna reform and their consequences for curriculum development and design, shows as one of the key constraints for the successful quality assurance of any curriculum development projects.

Throughout BIOART it became clear, that the development of a high-quality study course is highly complex, as it is not only important to focus on didactic elements during the entire development process, but also to consider the structural features of a program, to apply instruments such as taxonomies of learning objectives and to design the

courses and assessments in a way that is appropriate for the learners, which should be based on a careful selection of theoretical models.

5 Conclusions

One of the main project objectives of Erasmus+ KA2 project 586114-EPP-1-2017-1-ES-EPPKA2-CBHE-JP «Innovative Multidisciplinary Curriculum in Artificial Implants for Bioengineering BSc/MSc Degrees» [BIOART] was to develop an improved, best practice-curriculum that ensures the appropriateness of academic content for students and in accordance with industry demands. The sustainability of the project is ensured by two aspects: The first aspect was the development and introduction of adapted (bio-)engineering syllabi, comparable to similar international study programs or courses delivered at top universities in partner and program countries and established in a manner to ensure sustainability but also allow flexibility and responsiveness in a dynamic discipline over the next few years. This was achieved through integrative content development which was based on a preliminary stakeholder and labor market analysis and supported by academic governance and quality assurance, and the setting of national standards and the adherence to Bologna system standards in curriculum design. The second aspect focused on the establishing of BIOART labs using state-of-the-art equipment, hence building an infrastructure that allows students to form a learning environment that exposes them to modern tools and equipment that best prepares them for an industrial setting.

As a consequence of the BIOART project partners' collaborative efforts, new and updated curricula in the field of bioengineering have become more flexible and adaptive by means that curricula bridge the gap between abstract theories learning and generic engineering knowledge on the one hand and the more contextual, continuously changing and demanding realities of the labor market and biotech and bioengineering industry. The adapted or newly developed courses also became more open to adjustments based on the deficiencies identified during preparatory analysis. Furthermore, established practices and processes of curriculum development were reviewed and validated, and new contents were tested and assessed during pilot teaching sessions.

The process of curricula development and the updating of the courses was challenging for most partner universities involved in the project, due to the rigid time frame and the compliance with national and/or institutional accreditation regulations as well as meeting the formal Bologna requirements, such as the allocation of ECTS or the definition of learning outcomes. BIOART certainly contributed to the understanding of formal quality assurance mechanisms and acknowledging the quality assurance instruments as defined by the Bologna reform. Furthermore, project partners familiarized themselves with the concept of learner-centered curriculum design throughout the process. Unfortunately, a thorough qualitative assessment and evaluation of the developed curricula and/or adapted courses according to the different design levels that constitute a competence-oriented curriculum development could not be carried out due to the scope of BIOART.

One of the major obstacles that could be identified during the project's implementation was the lack of a common understanding and interpretation of the basic concepts of quality assurance in curriculum design according to Bologna standards. BIOART project partners are at different stages or levels of participation in the Bologna process and the implementation of its principles. There are also reservations related to the core concepts of the Bologna reform and their realistic application that cannot be ignored. Nonetheless, an effect is achieved through an indirect contribution to the discussion, adaptation and implementation of the Bologna reform and, thus, the alignment of the partner and program universities involved.

6 Acknowledgement

This work is carried out partly with the support of Erasmus+ KA2 project 586114-EPP-1-2017-1-ES-EPPKA2-CBHE-JP «Innovative Multidisciplinary Curriculum in Artificial Implants for Bio-Engineering BSc/MSc Degrees» [BIOART].

References

1. Birke, B., Hopbach, A.: Qualitätssicherung an österreichischen Hochschulen. Eine Bestandsaufnahme. Bericht gemäß §28 HS-QSG zum Stand 2015. Facultas, Wien (2016).
2. The European Higher Education Area in 2015: Bologna Process Implementation Report, https://eacea.ec.europa.eu/sites/eacea-site/files/european_higher_education_area_bologna_process_implementation_report.pdf, last accessed 28.09.2020
3. The European Higher Education Area in 2018: Bologna Process Implementation Report, https://eacea.ec.europa.eu/national-policies/eurydice/content/european-higher-education-area-2018-bologna-process-implementation-report_en, last accessed
4. Paris Communiqué 2018, http://www.ehea.info/media.ehea.info/file/2018_Paris/77/1/EHEAParis2018_Communique_final_952771.pdf, last accessed 2020/09/03.
5. Bologna Follow Up Group (BFUG), http://www.ehea.info/page-ministerial-conference-paris-2018, last accessed 2020/09/03.
6. Bergan S., Deca, L., Twenty Years of Bologna and a Decade of EHEA: What Is Next? In: European Higher Education Area: The Impact of Past and Future Policies, pp. 295-319, Springer, Cham (2018).
7. Vögtle, E.: Higher Education Policy Convergence and the Bologna Process. A Cross-National Study. 1st ed. Palgrave Macmillan, Basingstoke (2014).
8. Vögtle, E.: 20 years of Bologna - a story of success, a story of failure: Policy convergence and (non-) implementation in the realm of the Bologna Process. In: Innovation: The European Journal of Social Science Research 32 (3) 1-23, 2019.
9. Biggs, J., Tang, C. Teaching for Quality Learning at University. 1st ed. McGraw-Hill Education, Maidenhead (2011).
10. Adelman, C. (2009) The Bologna Process for US Eyes: re-learning higher education in the age of convergence. Washington, DC: Institute for Higher Education Policy. www.ihep.org/Research/GlobalPerformance.cfm.
11. European Commission (2013) The Bologna Process: towards the European Higher Education Area. http://ec.europa.eu/education/higher-education/bologna_en.htm
12. Hodson, P. & Thomas, H. (2003). Quality Assurance in Higher Education: fit for the new millennium or simply year 2000 compliant? Higher Education, 45(3), 375-387.

13. Teelken, C. & Wihlborg, M. (2010). Reflecting on the Bologna Outcome Space: Some pitfalls to avoid? Exploring universities in Sweden and the Netherlands, European Educational Research Journal, 9(1), 105–115.
14. Wihlborg, M. & Teelken, C. (2014). Striving for Uniformity, Hoping for Innovation and Diversification: A Critical Review concerning the Bologna Process – Providing an Overview and Reflecting on the Criticism. Policy Futures in Education, 12(8), 1084-1100.
15. Zahavi, H. (2019). The Bologna Process in Israel as a reflection of EU-Israel relations, European Journal of Higher Education, 9(1), 102-117.
16. Haack A., Braun L., Bach U., Vossen R., Jeschke S. (2014) Mapping the Goals of the Bologna Process. In: Jeschke S., Isenhardt I., Hees F., Henning K. (eds) Automation, Communication and Cybernetics in Science and Engineering 2013/2014. Springer, Cham. https://doi.org/10.1007/978-3-319-08816-7_22
17. National Agency for Higher Education https://en.naqa.gov.ua/
18. Strategy of the national agency for higher education quality assurance to 2022 https://en.naqa.gov.ua/wp-content/uploads/2020/04/Strategy-to-2020.pdf
19. Regulations on the system of providing the National University "Zaporizhzhya Polytechnic" with the quality of educational activities and the quality of higher education (internal quality assurance systems) of Zaporizhia: Educational Department, Educational and Methodical Department of NU "Zaporizhzhya Polytechnic"», 2019. – 21 p. (In Ukrainian)
20. The Remote Labs as an Effective Tool of Inclusive Engineering Education //Anzhelika Parkhomenko, Andriy Parkhomenko, Galyna Tabunshchyk, Karsten Henke, Heinz- Dietrich Wuttke / Conf. proc. of the XIVth International Conference on Perspective Technologies and Methods in MEMS Design (MEMSTECH), Lviv, Polyana, 18-21 April, -PP. 209-214. 10.1109/MEMSTECH.2018.8365735
21. Engineering Education for HealthCare Purposes: A Ukrainian Perspective // Galyna Tabunshchyk, Anzhelika Parkhomenko, Serhij Morshchavka, David Luengo / Conf. proc. of the XIVth International Conference on Perspective Technologies and Methods in MEMS Design (MEMSTECH), Lviv, Polyana, 18-21 April, -PP. 245 - 249 DOI: 10.1109/MEMSTECH.2018.8365743
22. Tabunshchyk G., Kapliienko T., Arras P. (2019) Sustainability of the Remote Laboratories Based on Systems with Limited Resources. In: Auer M., Langmann R. (eds) Smart Industry & Smart Education. REV 2018. Lecture Notes in Networks and Systems, vol 47. Springer, Cham. https://doi.org/10.1007/978-3-319-95678-7_22
23. M. Kuzmina, A. Karpenko, G. Tabunshchyk, V. Kuzmin, N. Karpenko, V. Popovych Career Strategies Approach for the Digitalised World Requirements. ICL2020 – 23rd International Conference on Interactive Collaborative Learning, 23–25 September, Virtual Conference (TalTech, Tallinn, Estonia), pp.985-992 (2020)

Section 1: Educational Approaches

Problem and Project-Based Learning in the fields of Health Informatics and Digital Health

Arriel Benis [0000-0002-9125-8300]

Holon Institute of Technology, Golomb St. 52, Building 1, Floor 6, Holon, 5810201, Israel
arrielb@hit.ac.il

Abstract. Problem-Based and Project-Based Learning approaches are used for teaching students in an active and inquiry-based learning way. Therefore, they are gaining knowledge and developing skills in the field. Informatics and digitalization are at the cross-road of Science, Technology, Engineering, Management, Innovation studies, and at least one application field. Health Informatics or more globally Digital Health are dynamic and stimulating fields of practice for developing a large range of capabilities dealing with creativity and leadership. Nonetheless, the specificities of both Health Informatics (standards, methodologies, human- or animal-focused information) and Digital Health (tools and systems) require an additional involvement from the students to deliver projects fitting with real-world needs and constraints. In this chapter, we will explain how to run a Problem-Based and Project-Based Learning in these challenging areas. This is required to have some awareness of health and medicine concepts and their specific constraints. This active and inquiry-based learning journey in the fields of Health Informatics and Digital Heath must give some day-to-day tools for being able to cultivate new scientific, technologic, engineering, and managerial know-how. It is essential to point out that Health Informatics and Digital Health are used herein as application fields. The methods and examples presented in this chapter can be used and mirrored in any other field and more particularly in bio-engineering and biomedical engineering.

Keywords: Healthcare Informatics, Digital Health, eHealth, Problem-Based Learning, Project-Based Learning.

1 Introduction

Problem-Based (PrBL) [1] and Project-Based Learning (PBL) [2] approaches are used for teaching students in an active and inquiry-based learning way. Therefore, they are gaining knowledge and developing skills in the field. Informatics and digitalization are at the cross-road of Science, Technology, Engineering, Management, Innovation studies, and at least one application field.

Health Informatics or more globally today Digital Health are dynamic and stimulating fields of practice for developing a large range of capabilities dealing with creativity and leadership. However, the specificities of both Health Informatics (standards, meth-

odologies, human- or animal-focused information) and Digital Health (tools and systems) call for an additional involvement from the students to deliver projects fitting with real-world needs and constraints.

In this chapter, we will explain how to run a PrBL or a PBL in these challenging areas. This requires having some awareness of health and medicine concepts and their specific constraints. This active and inquiry-based learning journey in the fields of Health Informatics and Digital Health must give some day-to-day tools for being able to cultivate new scientific, technologic, engineering, and managerial know-how.

It is essential to point out that Health Informatics and Digital Health are used herein as application fields. The methods and examples presented in this chapter can be used and mirrored in any other field and more particularly in bioengineering and biomedical engineering.

In the first section, we will describe what the PrBL and PBL approaches are respectively. We will explain synthetically their "What, Why, Where, When, Who, How, How much".

The second section will deal with what are Health Informatics and Digital Health. We will point-out their links to bioengineering and biomedical engineering.

In the section and last part, we will give some examples of implementing PBL focusing on Health-focused computerized systems. Also, we will give some ideas on possible projects to implement for educational purposes.

2 Problem and Project-Based Learning

Problem-Based Learning [1] and Project-Based Learning [2] approaches are unlike the traditional learning one. They are student-centered and by essence an experimentation-based educational method.

By exploring real-world situations, issues, needs, or comprehensively "challenges", PrBL and PBL, are allowing starting the learning process by disclosing a case or a situation that must be handled. The learner is not getting, as a premise, the "tool", the knowledge for solving it.

PrBL and PBL philosophy involve the learners in such a way that they need to discover and to identify what they need to know, and then to apply it. This is in contrast with the traditional learning methods which start by teaching what the learners need to know for solving a “problem” and then illustrating it with an example.

2.1 Learning to be a part of a whole

Globally, learning by investigating a problem or working on developing a product as a project gives the students and so the learners to develop real-life applicable and improvable capabilities.

First of all, these learning approaches being "student-centered", the learner is the project owner. This means that the students' voices and choices are one fundamental component of the learning and self-development process.

The 21st century is a connected world. In other terms, all the things around us (and by extension each one of us) are linked to others. More than talking out the Internet of Things (IoT) [3] we need to talk about the Internet of People (IoP) [4] which leads to Human-to-Human Interaction (HHI) (in a digital context) [5]. These HHIs are fundamentals in the communication process and in the opinion debate which are crucial bricks of a problem-solving or a project development [6]. From a teaching perspective, this means that both PrBL and PBL support the improvement and the reinforcement of interpersonal and teamwork skills. The students learn to be a part of a whole.

Moreover, these learning approaches drive to high-quality solutions. In this framework, quality means "going deeper" and searching to handle the question by thinking "out-of-the-box" for delivering a relevant, rigorously built, and real-world related answer.

Thus, working on a subject-matter chosen by students themselves and interesting them tends to augments self-learning intent.

2.2 Being efficient in the real-world, having real-life skills

The 21st-century skills, which can be also called "soft skills" [7, 8] which are developed and stimulated by the Problem and Project-Based Learning frameworks are well-summarized, for example, by the 6C's integrative concept [9, 10].

One important ability and skill required anywhere and at any time is "Communication". Individuals must communicate with their peers and their environment at the right time and in the right way for getting the right answer or support. Learners have the opportunity to actively exchange between them, and with other actors such as their peers (involved in handling another problem or project), teachers, mentors, or advisors. This is allowing them to create, to share, and to update the content of communicational processes such as "talking and listening", writing and reading", and "drawing and visualizing" information, and share knowledge.

Being able to do these tasks with the right individuals, with the right tools, and on the right channels are other bricks allowing acquiring and building collaboration proficiency.

Cultivating communication abilities also involves developing "critical thinking" skills. Indeed the learners must understand the information and the knowledge that they got, and judge if it is the right one for them and reliable. Additionally, the students must be able to judge themselves. They must be able to define if they are handling correctly the project (or the problem) that they have to deal with.

A good understanding of what critical thinking is leading to learn, to develop, and to enhance the individual and the collective creativity by proposing individually or collectively new or alternative solutions.

Finally, communicating efficiently and effectively, by delivering the right content in the right collaboration framework, in a creative and critical (but constructivist) manner is the result of a self-confidence development process happening over time and by handling numerous problems and projects.

Nowadays, these abilities and skills (Communication, Content development, Collaboration, Critical thinking, Creativity, Confidence) are critical ones for working efficiently and being a part of real-world projects (i.e. not training one).

2.3 Similarities and differences

Problem-Based Learning and Project-Based Learning have common aims and look similar. We previously pointed-out that both are:

- Teaching students in an active and inquiry-based learning way;
- Involving the students in such a way that they are building 21st-century skills (6C's);
- Taking more time than "classical learning" because they are actively engaging the students in the learning process by allowing them to define the question(s) and/or issue(s) that they will investigate.

Nevertheless, some important differences exist between them.

First, the starting point of a PrBL is commonly a use-case (real or fictive) disclosed in a structured form. On the second hand, PBL deals mainly with real-life situations and needs.

Accordingly, PrBL is more generally focusing on a single-subject when PBL is more often multi-subject. Indeed, a project by essence requires looking to different facets for managing it in a transversal manner.

PrBL being more abstract it may need small groups of students like PBL which can also be done alone.

Then, PrBL can be completed shortly (i.e. between a few hours to a few weeks) when PBL is lengthy (i.e. between a few weeks to a few months).

Additionally, the result of a PrBL consists generally in a written document or a presentation, and in the case of a PBL, the students create a new product by taking into account its environment (e.g. financial or regulatory constraints, end-user specific needs, etc.).

2.4 Step-by-step frameworks

Considering PrBL and PBL as a unified framework, articulated around 7 main steps helping the students working on a case study to investigate (for PrBL) or a real-life situation (for PBL).

Table 1 shows compactly and simply the different steps and the specific tasks to complete for each one.

It is important to keep in mind that for both frameworks, the students are

- Developing and enhancing their 6C's skills -Communication, Content development, Collaboration, Critical thinking, Creativity, Confidence-;
- Learning to work with S.M.A.R.T. goals [11, 12];
- Going deeper in their understanding by asking and answering a standard set of questions, the well-known 5W2H –c– [13].

Working S.M.A.R.T. means for each step having Specific (and Simple) and Measurable (and Manageable) oriented questions, and Achievable (and Action-oriented), Relevant (Realistic and Reasonable), and Time-bounded (and trackable) answers.

Table 1. Problem-Based Learning vs. Project-Based Learning

Steps	Problem-Based Learning	Project-Based Learning
0	• Get a **case study**	• Get a **real-life situation**
1	• **Examine** the issue (case study). • **Clarify** unclear terms.	• **Look** at the Real World. • **Search** for new needs.
2	• **Identify** the problem and its outcomes. • **State** what is **known**.	• **Define** the "Challenge" (the final product). • **State** what is **known**.
3	• **Brainstorm** and define hypotheses • Determine the **information to get** and the **tools to use** to solve the problem.	• **Brainstorm** about the "Challenge" (the final product). • **Develop** a relevant "Subject Matter **Expertise**" (Inquiry, Self-Directed Learning)
4	• **Suggest** potential solutions. • **Explain** each solution. • **Build bridges** between solutions.	• **Innovate** and look forward • **Model** the product. • **Implement** a product.
5	• **Select** relevant suggestions. • **Develop** deeply each solution	• **Experiment** (Run) the product. • Get **feedback**.
6	• **Assess** each solution(s). • **Solve** the problem.	• Self-**assess** and get assessments by peers and mentor • **Revise** and **improve** the product
7	• **Synthesis** and performance review. • **Share** findings.	• **Report** the overall product development process • **Disseminate** both report and product.

For example, working on "Step 1" of a Problem-Based Learning must be based on dealing, for example, with the following questions: "What is the issue?", "Why is it an issue?", "Where is it an issue? ", "When is it an issue? ", "Who is impacted by this

issue? ", "How did it tend to be an issue? ", " How much did it cost to solve this issue (financially, timely, humanly)?".

2.5 Take away message

- Problem-Based Learning [1] and Project-Based Learning [2] are engaging the student as an active learner.
- Discovery by investigating and experimenting are the fundamental elements of these teaching approaches.
- Both case-study and real-life situations are defined in such a way that they allow acquiring a "Subject Matter Expertise" by Self-Directed Learning.
- At each one of the learning steps defining S.M.A.R.T. goals with the 5W2H method helps efficiently brainstorm and prepare a deliverable based on earned information, knowledge, and experience.

3 Health Informatics and Digital Health

Health Informatics and Digital Health could sound for the non-specialist as similar fields but they have different aims and objectives even if they are complementary.

3.1 Health Informatics

The term "Health informatics" is used herein as a generic term referring to a large spectrum of strongly related and interdependent specialties, such as healthcare informatics, medical informatics, nursing informatics, clinical informatics, biomedical informatics, bioinformatics.

Health data, information, and knowledge are concepts that relate to an individual's health and care history [14].

The main initial aim of Health Informatics is to handle patient' health and care-related data, information, and knowledge, generated through the use of information systems and other technological complementary tools and devices [15].

This core definition has been expanded over time and today it takes into account the needs

- To stimulate and to evaluate the technology adoption and its performance;
- To continuously evaluate and improve the safety, quality, effectiveness, and efficiency of care [16].

Nowadays, Health Informatics is defined as:

(i) Pursuing "[…] the effective uses of biomedical data, information, and knowledge for scientific inquiry, problem-solving, and decision making, motivated by efforts to improve human health", and;

(ii) Investigating and supporting "[…] reasoning, modeling, simulation, experimentation, and translation across the spectrum from molecules to populations, dealing with a variety of biological systems" [17].

Health informatics is at the cross-road of healthcare, "computer and information sciences, engineering and technologies", the other sciences, and engineering domains as a

whole, and management. Researchers and Practitioners must have dual expertise, one from the health and medicine arena and one from the computer and information sciences and engineering.

3.2 Digital Health

The Internet is a full part of the day-to-day life of each one since the beginning of the 2000s, eHealth (electronic health) as a new and innovative concept, and the field appears. eHealth is extending the focal point of "health informatics" towards clinical information systems [18]. eHealth is taking the health services customers at the center and so emphasizing the "consumer health informatics" [19]. Disseminated over the two last decades the eHealth conceptualization is strongly related to the current economic digital revolution of health services [20]. The term "Digital Health" (dHealth) is nowadays used as a rebranding of eHealth. This term also sounds related to the use of "advanced technologies" used in the daily practice of health and medicine for non-fully technological communities [21]. eHealth and dHealth are synonyms even the second one must be considered as an evolution of the first one - a kind of popularization of the field [22]. We are using in this chapter the terms of Digital Health and dHealth for being consistent with the present-day uses of it both in academia and industry.

Digital Health is built around 4 main axes:

1. Data, information and knowledge collection, storage, mining, and dissemination for improving healthcare practice;
2. Personalization of precise healthcare;
3. Health promotion, wellbeing, and efficient self-management;
4. Optimization of the economic dimension and understanding of sociological one in this new context.

These 4 axes are strongly interrelated and interdependent.

The first one corresponding to the formal acquisition and processing of dHealth data is fundamental for the discovery of new knowledge [23–25]. This novel knowledge boosts the understanding of the market needs in terms of software and hardware frameworks and infrastructures. The last decade's move to “big data” [26, 27] has also induced a change in the manner to deal with the data’s generators and users. More specifically, the data must be “smart” for both and induce innovation for taking continuously the overall health sector to the next step [28].

Globally, dHealth has for aim to enhance and to improve the efficiency of historical surveillance systems used for preventing and controlling disease spreading. The effects of such changes will contribute to raising the quality of the methods, processes, and systems involved in, for example, (1) delivering wellbeing and healthcare services, (2) dealing with outbreaks prevention, detection, and mitigation outbreaks [29–32].

3.3 Take away message

Health Informatics and Digital Health are complementary fields with a large number of overlaps.

On the first hand, Health informatics is, at the intersection of healthcare, sciences, engineering, and management. Dual expertise, in a health-related field and another one

related to computer and information sciences and engineering. The scientific field mainly focuses on (i) the efficient and effective uses of data improving human health", and (ii) investigating and supporting health-related knowledge engineering and management for supporting translational exchanges between research to practice [17].

One the second hand, Digital Health has an integrative purpose focusing on combining various sources of cyber-physical technologies [33, 34] for improving the quality of the methods, processes, and systems involved in health services delivery in regular and disruptive times [29–32].

4 Problem and Project-Based Learning in the fields of Health Informatics and Digital Health

The use of PrBL and PBL (and other related approaches) established Higher Education and more particularly in medical and health-related fields.

4.1 eHealth and eLearning: PrBL and PBL as a whole

One key point allowing dealing efficiently with PrBL and PBL pedagogical methods which are actively involving the students was the increasing digitalization of the educational supports during the 2000's.

Similarly to eHealth at the beginning of this century [18], eLearning [35, 36] allowed to expand the approaches used for actively involving students in their educational experience. Virtual biomedical universities are today one of the effective and efficient ways [37–39].

A challenge for Health Informaticians is to develop tools supporting the training and the continuous education of clinicians. As some examples, the following kinds of tools and systems can be developed as a Problem or Project-Based Learning in this field [37]:

- Simulation and virtualization of ***biological and physiological process***es;
- Handling ***small and targeted use cases***, such as understanding a clinical context, interpreting clinical and/or biological data, and proposing a diagnosis [40];
- Using ***realistic simulation and interactive patients' data*** (such as Electronic Patient Records or Virtual Reality [41]) for training students to deal with at least one of the following: history taking, physical examination, ordering and investigating laboratory and imaging examinations, treatment prescription, and follow-up [42].

It is important to highlight that Health Informatics and Digital Health are studied by students outside of health-related curriculums.

For example, during their B.Sc. in Industrial Engineering and Technology Management curriculum at the faculty of Technology Management at the Holon Institute of Technology in Israel, the students are taking courses in Computer Integrated Manufacturing (CIM) and Introduction to the Internet of Things (IoT). The students are learning the fundamentals, advanced, and up-to-date concepts needed in their future practice. A large part of these courses are using the PrBL and PBL approaches. Moreover, the same

students have to complete a final graduation project during their 4th year [43]. A large and increasing number of these students are interested in discovering the Health Informatics and Digital Health world. These two frameworks are opportunities for these students without health-related background to have a first or extended contact with these fields.

Allowing students to familiarize themselves with Digital Health and Health Informatics with Project-Based Learning can consist of developing a "smart household pet feeding monitoring system". This kind of project and challenge relates to application in Veterinary Medicine and so Veterinary Medical Informatics. Accordingly, the students (ideally a team of 3) must design and implement a prototype of a pet feeding station. This system provides health and behavior monitoring services, collecting data and information for allowing health pet follow-up by looking at feeding habits and weight changes over time. Accordingly, the students had to deal with the overall PBL process and more specifically by:

1. Looking at the Real World and searching for new needs.
   e.g. Such a kind of pet feeding station exists in the market. So answering briefly to the 5W2H for understanding the potential customers' needs (the pets and their owners).
2. Defining what will be the prototype of the final product
   e.g. What will be included? Why and Who will it be used? Where? At home then how will it look? When will it work? How much will it cost to produce and then merchandise it?
3. Developing the relevant “Subject Matter Expertise”
   e.g. Even the students have a background in electronics, computing, automation, information technology, IoT, and management, they had to deliver a realistic product. They developed their knowledge and understanding of some fundamentals in animal physiology, behavior science, health informatics (focusing on veterinary aspects), and digital health (as a ubiquitous environment).
4. Innovating and looking forward to designing and implementing a prototype.
   e.g. the feeding station system was developed around 4 main components:
   (1) a Cyber-Physical System: an IoT device collecting and sending data related to feeding (RFID for identifying a pet, filling and emptying of the food and water tanks), “health indicators” such as pet health-related data (number of visits for eating or hydrating, quantity consumed at each time; weighing before feeding and/or hydrating; other physiological parameters -e.g. temperature, pulse…-);
   (2) Storing and managing the data related to a specific pet (associated with clinical data from the Electronic Medical Records managed by a Veterinary clinic);
   (3) Running an analytical system detecting deviations of the “health indicators” from norms and sending “alarms” by comparing current data to historical data of a pet or of “similar” pets.
   (4) Reporting in an interactive and user-friendly way reports to the pet-owner and optionally to a Doctor of Veterinary Medicine.

5. Experimenting the prototype with prospects (e.g. the students' families) and getting feedback.
6. Revising and improving the prototype according to the previously got feedback.
7. Disseminating both report and product.
   e.g. Submitting a final graduation report, presenting to peers and faculty staff the overall project.

A PBL like this one allows learning some elementary knowledge, methods, and tools used in the Health Informatics and Digital Health domains. The students in the current case discovered the specificity of managing a project involving telemonitoring, physiological signal processing, health information management (e.g Electronic Health Records), analysis and reporting of biological and medical (veterinary) data, information, and knowledge for clinical purposes and population surveillance [44].

4.2 Suggestions for PrBL and PBL subjects

The list below consists of a set of examples of PrBL and PBL subjects. The teacher will define it as a problem or as a project, depending on the needs of the course where they will be used.

1. Simulating a physiological process [45, 46]. e.g. automatic detection of QRS components in ECG signals.
2. Create an interactive family tree with known genetic/phenotypic characteristics for each family member.
3. Designing and developing a model looking at a patient with diabetes treatment adherence from a Systems Medicine perspective [47].
4. Simulating an epidemic / a pandemic (AIDS, COVID, Influenza…) [48].
5. Designing a smart home for supporting the elderly.
6. Designing a smart household pet feeding monitoring system [44].
7. Developing an application managing personal health records.
8. Developing an alert system to publicize the spread of viruses and recommendations for individual and collective protection.
9. Writing a review of applications used by patients with diabetes for managing their treatment.
10. Writing a position paper about the potential future ethical issues induced by new Health Informatics and Digital Health developments.

4.3 A positive experience for students

Given the limited-time that health professionals and the increasing needs to consolidate the teaching resources for students, Problem-Based Learning and Project-Based can be both run in a traditional way with face-to-face meetings (in class) or in a digital context (fully online or blended). The effectiveness of moving a part or all the PrBL or the PBL online has a positive impact on satisfaction. Moreover, it enhances self-improvement abilities to acquire new knowledge and develop 21st-century skills [41].

Looking at a more specific assessment of the use of PrBL (and by extension PBL) as a strategy for teaching Medical Informatics chapters such as Information and Communication Technology in the medical schools we can notice also a students' high acceptance and satisfaction with this method. Additionally, it looks like previous PrBL experience and so related skills influence positively this assumption [49].

4.4 Take away message

PrBL and PBL are pedagogical philosophies actively involving the students. Today they are powerfully based on digital technologies. Since the beginning of the 21st century, eHealth use and impact in constantly increasing [18]. eLearning [35, 36] has the same trend. Therefore, eHealth comprises an eLearning component focusing on instruction and training of professionals involved in health-related practice and services [37–39].

PrBL and PBL are used for teaching Health Informatics and Digital Health realistically.

5 Conclusions

In this chapter, we introduced Problem-Based Learning and Project-Based Learning, 2 student-centered, investigated-focused, and continuous-engagement educational methods.

In a second step, we explained what are Health Informatics and Digital Health.

Then, we described how PrBL and PBL are and can be used for efficiently and effectively teaching Health Informatics and Digital Health. This is a “game” having for goal to integrate prior knowledge and cultivate new scientific, technologic, engineering, and managerial know-how when they are not mastered [50].

The success of both, PrBL and PBL is accordingly dependent on the responsibility of students in being engaged in the “game” (the learning task) and so being involved in a self-regulated learning (SRL) process and environment [51]. Indeed, today, PrBL and PBL are well-fitting with online teaching for both biomedical sciences and engineering education and public health literacy [52].

Nota Bene

- All the references provided in this chapter are additional resources for teachers and students interested in going deeper into their mastering of Problem-Based Learning and Project-Based Learning.
- Problem-Based Learning and Project-Based Learning are two of the numerous methods used in the student-centered, active, and inquiry-based learning sphere.

References

1. Wood, D.F.: ABC of learning and teaching in medicine: Problem based learning, (2003). https://doi.org/10.1136/bmj.326.7384.328.
2. Bell, S.: Project-Based Learning for the 21st Century: Skills for the Future. The Clearing House: A Journal of Educational Strategies, Issues and Ideas. 83, 39–43 (2010). https://doi.org/10.1080/00098650903505415.
3. Sunyaev, A.: The Internet of Things. In: Sunyaev, A. (ed.) Internet Computing. pp. 301–337. Springer International Publishing, Cham (2020). https://doi.org/10.1007/978-3-030-34957-8_10.
4. Conti, M., Passarella, A.: The Internet of People: A human and data-centric paradigm for the Next Generation Internet. Comput. Commun. 131, 51–65 (2018). https://doi.org/10.1016/j.comcom.2018.07.034.
5. Sharma, R., Pavlovic, V.I., Huang, T.S.: Toward multimodal human-computer interface. Proc. IEEE. 86, 853–869 (1998). https://doi.org/10.1109/5.664275.
6. Mainela, T., Ulkuniemi, P.: Personal interaction and customer relationship management in project business. Jnl of Bus & Indus Marketing. 28, 103–110 (2013). https://doi.org/10.1108/08858621311295245.
7. Hadim, H.A., Esche, S.K.: Enhancing the engineering curriculum through project-based learning. In: 32nd Annual Frontiers in Education. pp. F3F–1–F3F–6. IEEE. https://doi.org/10.1109/FIE.2002.1158200.
8. Robles, M.M.: Executive Perceptions of the Top 10 Soft Skills Needed in Today's Workplace. Business Communication Quarterly. 75, 453–465 (2012). https://doi.org/10.1177/1080569912460400.
9. Golinkoff, R.M., Hirsh-Pasek, K.: Becoming brilliant: What science tells us about raising successful children. American Psychological Association, Washington (2016). https://doi.org/10.1037/14917-000.
10. Taufiq, M., Wijayanti, A., Yanitama, A.: Implementation of blended project-based learning model on astronomy learning to increase critical thinking skills. J. Phys. Conf. Ser. 1567, 042049 (2020). https://doi.org/10.1088/1742-6596/1567/4/042049.
11. Ogbeiwi, O.: Why written objectives need to be really SMART. British Journal of Healthcare Management. 23, 324–336 (2017). https://doi.org/10.12968/bjhc.2017.23.7.324.
12. Aghera, A., Emery, M., Bounds, R., Bush, C., Stansfield, R.B., Gillett, B., Santen, S.A.: A Randomized Trial of SMART Goal Enhanced Debriefing after Simulation to Promote Educational Actions. West. J. Emerg. Med. 19, 112–120 (2018). https://doi.org/10.5811/westjem.2017.11.36524.
13. van Gaasbeek, J.R.: Before Requirements: What, Who, Where, When, Why, and How. Insight. 2, 11–14 (2000). https://doi.org/10.1002/inst.20002411.
14. Vest, J.R., Gamm, L.D.: Health information exchange: persistent challenges and new strategies. J. Am. Med. Inform. Assoc. 17, 288–294 (2010). https://doi.org/10.1136/jamia.2010.003673.
15. Haux, R.: Medical informatics: Past, present, future. Int. J. Med. Inform. 79, 599–610 (2010). https://doi.org/10.1016/j.ijmedinf.2010.06.003.
16. Agarwal, R., Gao, G., DesRoches, C., Jha, A.K.: Research Commentary —The Digital Transformation of Healthcare: Current Status and the Road Ahead. Information Systems Research. 21, 796–809 (2010). https://doi.org/10.1287/isre.1100.0327.
17. What is Informatics?, https://www.amia.org/fact-sheets/what-informatics, last accessed 2020/09/24.

18. Khillar, S.: Difference Between Digital Health and Health Informatics | Difference Between, http://www.differencebetween.net/technology/difference-between-digital-health-and-health-informatics/, last accessed 2020/06/26.
19. Eysenbach, G.: What is e-health? J. Med. Internet Res. 3, E20 (2001). https://doi.org/10.2196/jmir.3.2.e20.
20. Eysenbach, G.: Celebrating 20 Years of Open Access and Innovation at JMIR Publications. J. Med. Internet Res. 21, e17578 (2019). https://doi.org/10.2196/17578.
21. Ahmadvand, A., Kavanagh, D., Clark, M., Drennan, J., Nissen, L.: Trends and Visibility of "Digital Health" as a Keyword in Articles by JMIR Publications in the New Millennium: Bibliographic-Bibliometric Analysis. J. Med. Internet Res. 21, e10477 (2019). https://doi.org/10.2196/10477.
22. Khillar, S.: Difference Between Digital Health and Health Informatics | Difference Between, http://www.differencebetween.net/technology/difference-between-digital-health-and-health-informatics/, last accessed 2020/06/26.
23. Fayyad, U., Piatetsky-Shapiro, G., Smyth, P.: The KDD process for extracting useful knowledge from volumes of data, http://dx.doi.org/10.1145/240455.240464, (1996). https://doi.org/10.1145/240455.240464.
24. Chandola, V., Sukumar, S.R., Schryver, J.C.: Knowledge discovery from massive healthcare claims data. In: Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining - KDD '13. p. 1312. ACM Press, New York, New York, USA (2013). https://doi.org/10.1145/2487575.2488205.
25. Levner, E., Kriheli, B., Benis, A., Ptuskin, A., Elalouf, A., Hovav, S., Ashkenazi, S.: Entropy-Based Approach to Efficient Cleaning of Big Data in Hierarchical Databases: 9th International Conference, Held as Part of the Services Conference Federation, SCF 2020, Honolulu, HI, USA, September 18-20, 2020, Proceedings. In: Nepal, S., Cao, W., Nasridinov, A., Bhuiyan, M.D.Z.A., Guo, X., and Zhang, L.-J. (eds.) Big Data – BigData 2020. pp. 3–12. Springer International Publishing, Cham (2020). https://doi.org/10.1007/978-3-030-59612-5_1.
26. Gandomi, A., Haider, M.: Beyond the hype: Big data concepts, methods, and analytics. Int. J. Inf. Manage. 35, 137–144 (2015). https://doi.org/10.1016/j.ijinfomgt.2014.10.007.
27. Mooney, S.J., Westreich, D.J., El-Sayed, A.M.: Commentary: Epidemiology in the era of big data. Epidemiology. 26, 390–394 (2015). https://doi.org/10.1097/EDE.0000000000000274.
28. VanderWaal, K., Morrison, R.B., Neuhauser, C., Vilalta, C., Perez, A.M.: Translating Big Data into Smart Data for Veterinary Epidemiology. Front Vet Sci. 4, 110 (2017). https://doi.org/10.3389/fvets.2017.00110.
29. Who Regional Office for Europe: From Innovation to Implementation: EHealth in the WHO European Region. World Health Organization (2016).
30. Inkster, B., O'Brien, R., Selby, E., Joshi, S., Subramanian, V., Kadaba, M., Schroeder, K., Godson, S., Comley, K., Vollmer, S.J., Mateen, B.A.: Digital Health Management During and Beyond the COVID-19 Pandemic: Opportunities, Barriers, and Recommendations. JMIR Ment Health. 7, e19246 (2020). https://doi.org/10.2196/19246.
31. Whitelaw, S., Mamas, M.A., Topol, E., Van Spall, H.G.C.: Applications of digital technology in COVID-19 pandemic planning and response. The Lancet Digital Health. 2, e435–e440 (2020). https://doi.org/10.1016/S2589-7500(20)30142-4.
32. Budd, J., Miller, B.S., Manning, E.M., Lampos, V., Zhuang, M., Edelstein, M., Rees, G., Emery, V.C., Stevens, M.M., Keegan, N., Short, M.J., Pillay, D., Manley, E., Cox, I.J., Hey-

mann, D., Johnson, A.M., McKendry, R.A.: Digital technologies in the public-health response to COVID-19. Nat. Med. 26, 1183–1192 (2020). https://doi.org/10.1038/s41591-020-1011-4.
33. Yang, G., Pang, Z., Jamal Deen, M., Dong, M., Zhang, Y.-T., Lovell, N., Rahmani, A.M.: Homecare Robotic Systems for Healthcare 4.0: Visions and Enabling Technologies. IEEE J Biomed Health Inform. 24, 2535–2549 (2020). https://doi.org/10.1109/JBHI.2020.2990529.
34. Marques, G., Miranda, N., Kumar Bhoi, A., Garcia-Zapirain, B., Hamrioui, S., de la Torre Díez, I.: Internet of Things and Enhanced Living Environments: Measuring and Mapping Air Quality Using Cyber-physical Systems and Mobile Computing Technologies. Sensors . 20, (2020). https://doi.org/10.3390/s20030720.
35. Gunasekaran, A., McNeil, R.D., Shaul, D.: E-learning: research and applications. Ind and Commercial Training. 34, 44–53 (2002). https://doi.org/10.1108/00197850210417528.
36. Gachago, D., Morkel, J., Hitge, L., van Zyl, I., Ivala, E.: Developing eLearning champions: a design thinking approach. Int J Educ Technol High Educ. 14, 229 (2017). https://doi.org/10.1186/s41239-017-0068-8.
37. Le Beux, P, Fieschi, M.: Virtual biomedical universities and e-learning. Int. J. Med. Inform. 76, 331–335 (2007). https://doi.org/10.1016/S1386-5056(07)00060-3.
38. Buendía, F., Gayoso-Cabada, J., Sierra, J.-L.: Generation of Standardized E-Learning Content from Digital Medical Collections. J. Med. Syst. 43, 188 (2019). https://doi.org/10.1007/s10916-019-1330-5.
39. Galvin, J.R., D'Alessandro, M.P., Erkonen, W.E., Knutson, T.A., Lacey, D.L.: The virtual hospital: a new paradigm for lifelong learning in radiology. Radiographics. 14, 875–879 (1994). https://doi.org/10.1148/radiographics.14.4.7938774.
40. Ricci, F.L., Consorti, F., Pecoraro, F., Luzi, D., Mingarelli, V., Tamburis, O.: HIN - Health Issue Network as Means to Improve Case-Based Learning in Health Sciences Education. Stud. Health Technol. Inform. 255, 262–266 (2018).
41. Tudor Car, L., Kyaw, B.M., Dunleavy, G., Smart, N.A., Semwal, M., Rotgans, J.I., Low-Beer, N., Campbell, J.: Digital Problem-Based Learning in Health Professions: Systematic Review and Meta-Analysis by the Digital Health Education Collaboration. J. Med. Internet Res. 21, e12945 (2019). https://doi.org/10.2196/12945.
42. Ricci, F.L., Consorti, F., Pecoraro, F., Luzi, D., Mingarelli, V., Miotti, S., Tamburis, O.: Understanding Petri Nets in Health Sciences Education: The Health Issue Network Perspective. Stud. Health Technol. Inform. 270, 484–488 (2020). https://doi.org/10.3233/SHTI200207.
43. Benis, A., Nelke, S.A., Winokur, M.: Upgrading Industrial Engineering and Management curriculum to Industry 4.0, http://dx.doi.org/10.1109/icit45562.2020.9067243, (2020). https://doi.org/10.1109/icit45562.2020.9067243.
44. Benis, A.: Healthcare Informatics Project-Based Learning: An Example of a Technology Management Graduation Project Focusing on Veterinary Medicine. Stud. Health Technol. Inform. 255, 267–271 (2018).
45. Physiological Process, https://www.sciencedirect.com/topics/immunology-and-microbiology/physiological-process, last accessed 2020/10/03.
46. Hmelo, C.E., Holton, D.L., Kolodner, J.L.: Designing to Learn About Complex Systems. Journal of the Learning Sciences. 9, 247–298 (2000). https://doi.org/10.1207/S15327809JLS0903_2.
47. Comte, B., Baumbach, J., Benis, A., Basílio, J., Debeljak, N., Flobak, Å., Franken, C., Harel, N., He, F., Kuiper, M., Méndez Pérez, J.A., Pujos-Guillot, E., Režen, T., Rozman, D., Schmid, J.A., Scerri, J., Tieri, P., Van Steen, K., Vasudevan, S., Watterson, S., Schmidt, H.H.H.W.: Network and Systems Medicine: Position Paper of the European Collaboration

on Science and Technology Action on Open Multiscale Systems Medicine. Netw Syst Med. 3, 67–90 (2020). https://doi.org/10.1089/nsm.2020.0004.
48. Huang, C.-Y., Tsai, Y.-S., Wen, T.-H.: Simulations for Epidemiology and Public Health Education. In: Mustafee, N. (ed.) Operational Research for Emergency Planning in Healthcare: Volume 2. pp. 176–202. Palgrave Macmillan UK, London (2016). https://doi.org/10.1007/978-1-137-57328-5_9.
49. Burgun, A., Darmoni, S., Duff, F.L., Wéber, J.: Problem-based learning in medical informatics for undergraduate medical students: an experiment in two medical schools. Int. J. Med. Inform. 75, 396–402 (2006). https://doi.org/10.1016/j.ijmedinf.2005.07.014.
50. Jin, J., Bridges, S.M.: Educational technologies in problem-based learning in health sciences education: a systematic review. J. Med. Internet Res. 16, e251 (2014). https://doi.org/10.2196/jmir.3240.
51. English, M.C., Kitsantas, A.: Supporting Student Self-Regulated Learning in Problem- and Project-Based Learning. Interdisciplinary Journal of Problem-Based Learning. 7, (2013). https://doi.org/10.7771/1541-5015.1339.
52. Hansen, M.M.: Versatile, immersive, creative and dynamic virtual 3-D healthcare learning environments: a review of the literature. J. Med. Internet Res. 10, e26 (2008). https://doi.org/10.2196/jmir.1051.

Content and competency structure of IT specialties students in the field of bioengineering

Oleksandr Tarasov[1[0000-0002-0493-1529]], Pavlo Sahaida[2[0000-0002-4700-8160]],

Volodymyr Tankut[3[0000-0002-9901-4036]], Liudmyla Vasylieva[4[0000-0002-9277-1560]],

Kinga Korniejenko[5[0000-0002-8265-3982]]

[1, 2, 4] Donbass State Engineering Academy, Akademichna str.,72, Kramatorsk, 84313, Ukraine
[3] Sytenko Institute of Spine and Joint Pathology of National Ukrainian Academy of Medical Sciences, Pushkinskaya str., 80, Kharkiv, 61024, Ukraine
[5] Cracow University of Technology, Warsaw str. 24, 31-155, Kracow, Poland
alexandrtar50@gmail.com

Abstract. This lecture is devoted to the introduction of multidisciplinary training for the specialists in the field of design and implementation of high-tech implants with elements of intellectual behavior. The end-to-end process of designing such implants is considered: from the study and analysis of functional requirements, as well as the study of new materials, 3D-modeling of geometric parameters, and modeling of mechanical properties – to the development of electronic and software components of sensory and control subsystems.

Identification of the main stages of implant development made it possible to identify and formulate the necessary competencies that need to be formed by specialists in this field. Students will be acquainted with the general direction of training, the skills, and abilities that they are going to acquire during the training. In the process of system analysis of the development of implants, an ontological approach was applied. It allowed formalizing the analysis results in a form convenient for the preparation of regulatory documents. Information support for the organizational and methodological work of teachers was also improved. This lecture also increases the students' motivation to study disciplines of the bioengineering field.

A competency-based approach to the organization of the educational process requires the development of an integrated model of educational processes and the objects that participate in them, based on high-level abstractions.

Keywords: bioengineering, competence approach, implant, ontological modeling.

1 Introduction

The rate of renewal of engineering knowledge and competencies is constantly increasing. In most industries, there is a reduction in the time of the innovation cycle – the time between scientific development and the introduction of the technologies into production. This is especially true concerning information technology (IT) education. Technical skills are also evolving rapidly. Many students currently studying at universities will have new professions in the future that do not yet exist. The skills they will need to possess are not yet determined. For many students, retraining will become a common practice, as we enter an era of continuing evolution and education. At the same time, the engineering problems and tasks are changing in connection with the introduction of technologies into all spheres of life and the economy. Technical systems are becoming more complex and interconnected. Solving these problems and managing such systems requires new approaches to education. It is necessary to take into account not only the technical components of complex systems but also their impact on social, environmental, economic, and other fields [1].

Currently, educational strategies for the development of learning systems are developing in technologically advanced countries and include various specialized programs for different levels of education. Curricula are designed as a set of interdisciplinary competencies provided by the disciplines to be studied.

For the students to learn to make innovative creative decisions, it is necessary to include components related to the research technological solutions and application of mathematics as a general method of modeling in education.

This is especially important when executing multidisciplinary training programs combining natural sciences and other disciplines. For example, during the implementation of the international educational project ERASMUS + "Innovative Multidisciplinary Curriculum in Artificial Implants for Bio-Engineering BSc/MSc Degrees". The main objective of the project is to improve the process of students' training in the field of medical device design.

To increase the students' motivation effectively, it is necessary to constantly improve the structure and methods of the educational process. This is possible through the development of e-learning methods using modern information, communication technologies and global networks [2]. It is also necessary to improve the content of disciplines, which provide knowledge, skills and competence of graduates. The set of knowledge and skills that a student receives in the learning process, taking into account the technologies used, forms a model of the educational process. This model is more or less consistent with the model of future professional activity. Therefore, the development and improvement of the content of education should be based on an analysis of the directions of technological development in this area of knowledge. The complexity of these tasks requires constant analysis of technology development trends. The result of this analysis are recommendations for the improvement of the content of the educational process of IT specialists training [3, 4], as well as new technologies for teaching and knowledge assessment [5, 6]. It is also necessary to consider the requirements for the training of specialists of enterprises and organizations where they carry out their professional activities.

The competency-based approach to the organization of the educational process requires the development of an integrated model of the processes occurring during the learning process and the objects taking part in them.

During the training process, students acquire knowledge of their future professional activities. An important basis for the information support of engineering activities is the effective presentation of knowledge about the work of the subject area (SA), as well as standardization of tasks and standardization of the educational process. Some of them are fundamental, while some are aimed at the development of modern technologies and tools for professional work. Ensuring the necessary balance between these parts determines the effectiveness of training. Both parts change over time, and the sphere of activity of the corresponding specialty expands [3, 4].

Under these conditions, it is necessary to constantly clarify the boundaries of the specialization area, since new branches of knowledge are identified constantly as well. All this requires steady replenishment of the knowledge and skills of the graduates in the subject area and adjustments of the educational process. Solving this problem will increase the competitiveness of graduates in the labor market. Taking these tasks into account, the future of higher education is linked to the following processes: integration of education, science and innovation; the use of technologies based on the participation of students in innovation, research, etc., which provides the competency-based approach to learning [5, 6], the growth of knowledge and skills due to the active participation of students in the learning process.

The purpose of the work is to formalize knowledge obtained during the educational process when forming an interdisciplinary educational program for training BSc/MSc in IT and bioengineering. Application of the ontological approach [7] to the formalization of knowledge about the educational process makes it possible to demonstrate the connection between the elements of students' training.

2 Content of competencies for an IT program in bioengineering

The taxonomy of educational activity with highlighted attributes of concepts and their meanings is shown in Fig. 1. The department is a part of the university and trains students in a specialty that corresponds to its profile, scientific, personnel, material and technical resources. The specialty of BSc/MSc training allows future employees, who obtained the necessary level of qualifications, to participate in a certain type of economic activity, performing professional duties, following the job description of their workplace. Using their professional skills, employees perform various work functions, generally distributed by type (research, control, design, technological, forecasting, organizational) and levels (stereotyped, operational, technological, research).

During the process of implementation of the work functions, students (BSc/MSc) perform various tasks of activity, which also differ in types (professional, social, production, etc.). Qualitative performance of such tasks should be ensured with appropriate knowledge and understanding of production and scientific tasks and algorithms for their solution. From the point of view of educational activities, the functions and tasks of

future employees should be ensured by the formation of different competencies. Accordingly, competencies are supported by the skills formed during the study of disciplines of different cycles.

The formed skills determine the volume and content of academic disciplines and vary by type (subject-practical, subject-intellectual, sign-practical, sign-mental).

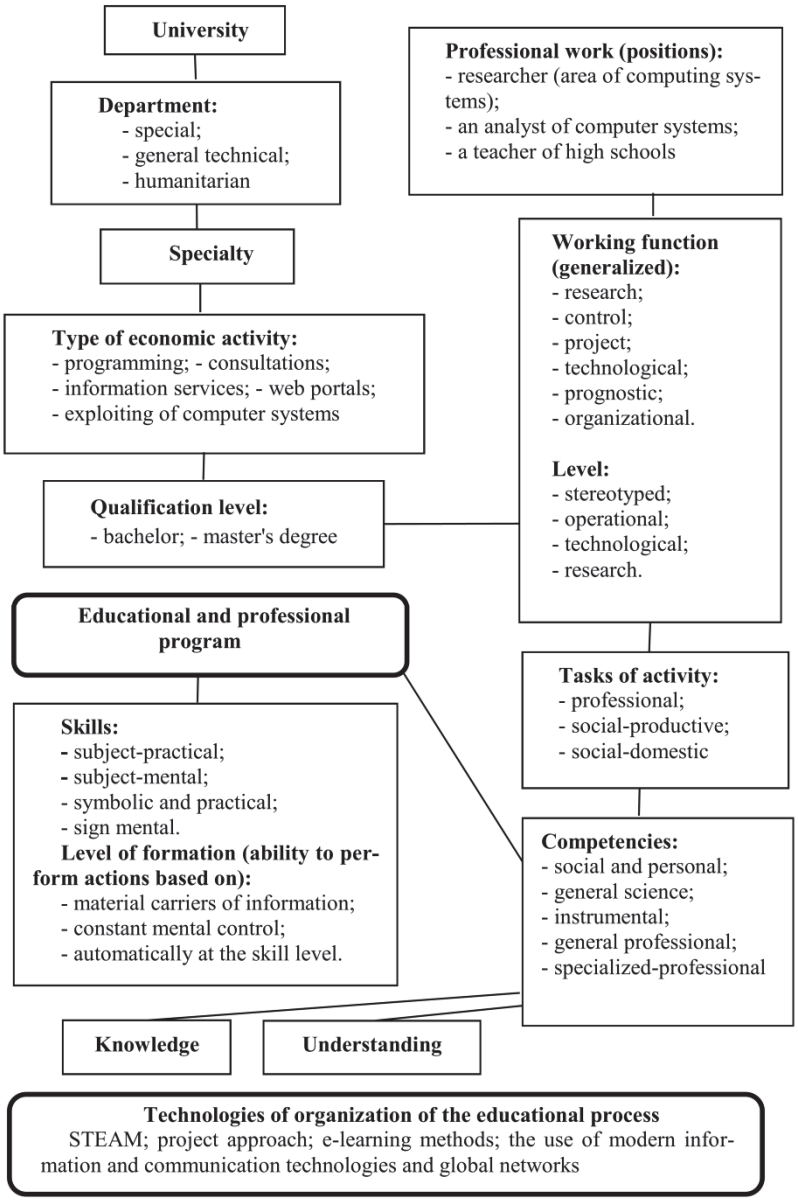


Fig. 1. Taxonomy of educational activity with selected attributes of concepts and their meanings

The skills are also divided according to the level of formation (the ability to perform actions on material carriers, due to constant mental control, automatically according to skill levels, etc.) [3, 4]. Also, when developing curricula, it is necessary to adhere to general and national standards [1, 3].

Thus, the development of the content of training IT students in medicine requires consistent implementation of the following tasks [6, 8, 9, 10]:

- formation of typical production functions according to the classifier of occupations and types of economic activities;
- determination of the training level of bachelors or masters to perform the assigned functions and key competencies for a given educational and qualification level;
- development of these competencies in the field of professional activity of a graduate in the specialty "IT in medicine";
- formation of MSc skills that differ from the requirements for the "level of formation" of the BSc abilities;
- development of the structure and content of the disciplines, within which the necessary knowledge and skills are formed;
- determination of responsibilities of the corresponding qualification level;
- determination of the list of competencies formed by students within the framework of the educational process.

Subsequently, for these competencies, the skills should be determined to be formed when studying the sections of the relevant training modules of the professional training of the IT BSc/MSc in medicine.

Development of the content of the disciplines for the specialty "IT in Medicine" was performed using the elaborated taxonomy of educational activities (see Figure 1) and the identified tasks for students' training. Based on the ontological approach to modeling of the processes in the subject area "Training of IT MSc in Medicine", basic competencies were identified. A fragment of the applied ontology of the subject area " Training of IT MSc in Medicine" is shown in Fig. 2. The ontology presents the skills that ensure the MSc professional competence. The production skills of graduates as specialists in the field of information and communication technologies in medicine are also presented [8, 9].

Thus, the students should know the basic concepts of the development of medical technologies related to the receipt, transmission and processing of information for various purposes. Fields of IT application should include nanomedicine, orthopedics, stimulation, diagnosis and use of implants, etc. Students should be able to solve standard tasks of professional activity based on informational and bibliographic culture and using information and communication technologies. In addition, it is necessary to take into account the requirements of information security, ethical and legal aspects of the use of medical information. This requires the competence of students in the field of data protection in medicine [9].

All the applied methods and studied disciplines can be divided into two parts. The main part is necessary for general training of students. This knowledge and skills are necessary at all stages of a specialist's activity. The second part is a special one and it

is designed to solve problems in a chosen subject area. In particular, mathematics is a general method for studying systems. Artificial intelligence methods can also be applied at all stages of the automation of a specialist's activities.

For general students' training, basic mathematical disciplines are of particular importance, which develop [11, 12]:

- the ability for abstract thinking, critical analysis, evaluation and synthesis of new ideas, for searching, processing and analyzing information from various sources, for building logical conclusions, using formal mathematical models;
- the ability for mathematical and logical thinking, formulation and research of mathematical models, discrete mathematical models in particular;
- the ability to substantiate the choice of methods for solution of theoretical and applied problems in the field of computer science;
- the ability to use, develop and research mathematical methods and algorithms for data processing in various subject areas (technical and medical systems).
- the ability to process and interpret the results obtained, to analyze, comprehend and present them, to justify the proposed solutions at the modern scientific and technical level.
- the ability to master new methods and technologies for studying systems for various purposes [13].

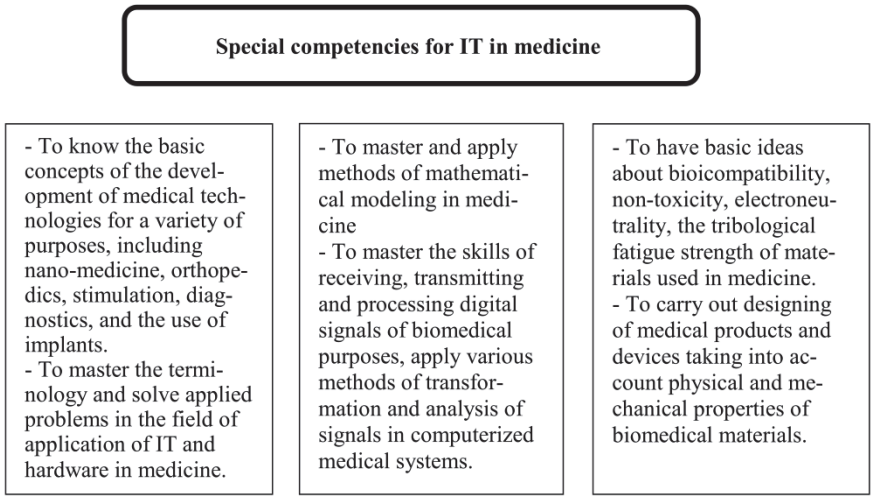


Fig. 2. A fragment of the ontology of the subject area "Master's Degree for IT in Medicine Specialty"

The mastery of the methods of machine learning and artificial intelligence by the students expands the mathematical apparatus necessary for the development and study of models of various phenomena and processes. As a result, there appears a need to

create intelligent devices and software systems [14–16]. To do this, students should be able to:

- carry out a formal description of the tasks of studying operations in organizational, technical and socio-economic systems of different purposes, to determine their optimal solutions using methods of machine learning and artificial intelligence;
- realize the mathematical representation of models of systems of different types, to perform the iterative development of a model; use models and modeling process to test hypotheses, evaluate the adequacy of models;
- create mathematical models and algorithms to research knowledge in databases and data repositories (knowledge discovery in a database – KDD), including knowledge acquisition (Data Mining);
- develop information models of the medical diagnostic process at health care institutions, evaluate the efficiency of the system for obtaining, collection, processing, transmission and protection of medical information;
- master the documentation on the development and implementation of a model and present it in oral and written form.

Given the focus of the BioArt project on training IT students in medicine, special knowledge in the field of medicine as a subject area [8, 9, 13] should be limited within the framework of the corresponding curriculum. Subsequent specialization at medical institutions will allow students to supplement their knowledge in the field of bioengineering.

3 Model of learning activities in the field of design and use of implants

Let's consider the formalization of knowledge about business processes in organizational and technical systems involved in the development and production of biomaterials, measuring and control devices. Let's also present the process of designing, manufacturing and using implants. Consideration of the areas of application of IT in medicine will justify the motivation for students who study in programs related to bioengineering.

It is difficult to single out all the areas of students' research and training related to the use of IT in medicine since this is a very large field of activity. Therefore, we consider the content of students' training for IT specialties, which is carried out as a part of the BioArt project. To determine the set of necessary knowledge and skills, the sequence of operations of the process of an implant creation are considered. Also let us present the information base necessary for implementation of this process. This will help to intelligently approach the formation of the set of competencies required for the work of an IT specialist in the field of bioengineering. Consideration of the fields of application of IT in medicine will also allow substantiating the motivation of students enrolled in programs related to bioengineering.

The model of educational activity with the highlighted activities and their attributes is shown in Fig. 3. After assembly and testing of implants, they are implanted to patients (Fig. 4). Then, periodic monitoring of a patient and an implant is performed. The results

obtained are processed by statistical methods to improve the quality of the entire process: design, manufacture, installation and operation.

The concepts presented in these figures and the relationship between them reflect the structure of educational activities in the field of creation and operation of implants [10, 17–19]. Bioengineers perform various job functions applying their professional skills. These skills are usually categorized by type (research, control, design, technological, predictive, organizational) and levels (stereotyped, operational, technological, research).

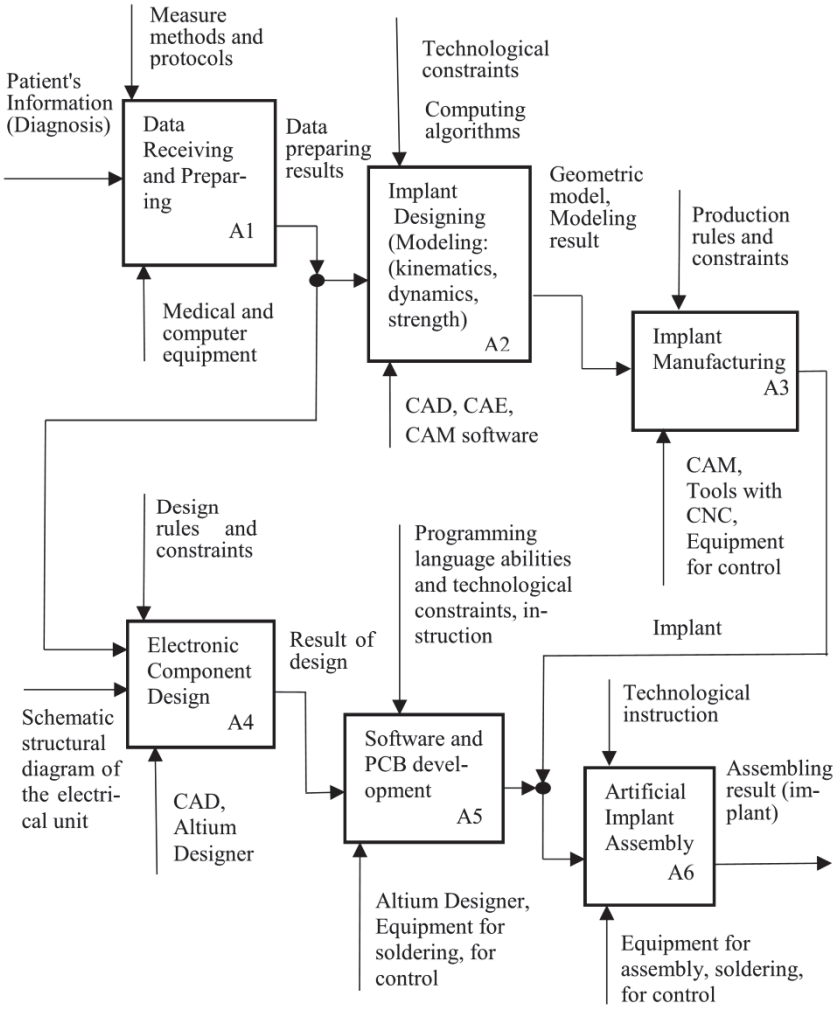


Fig. 3. Model of educational activity for an implant design with highlighted activities and their attributes.

Modern implants are complex products that include mechanical parts and electronics [20–23]. The task of the mechanical part of the implant is to ensure the strength of the structure, its kinematics and durability of work after implantation to a patient. Electronics allows to control the parameters of a person and an implant, as well as transmit the necessary information for storage in the control system or for managing the implant. The third component of the overall system is the software subsystem that ensures joint operation of its elements, the logic of processing the received data [12, 13, 16, 24, 25].

The competencies that students should receive based on the developed curriculum should be formulated and grouped according to these components of an implant.

Patients' information is the starting point for selecting or designing an implant (Fig. 3, A1). Much of this information is obtained from instrumental research. This information is obtained using various sensors and devices recording a person's parameters. The resulting parameters are usually converted into electrical quantities and presented in graphical form. Thus, students should learn the physical fundamentals of sensors, know the different types of sensors and the peculiarities of their application. Moreover, to be able to combine sensors into sensor networks for simultaneous monitoring of a number of human parameters [8, 13, 14, 20], the students should:

- know the types and understand the features of the application of medical sensors, topology and characteristics of the wireless sensor network;
- master the principles of operation and scope of application of biosensors, converters, their features and limitations, various components of biosensors and biosensor networks;
- master the principles of designing and modeling of biosensor networks, know the features of applied bioelectronics components, potential benefits and achievements of creation of wireless sensor networks.

Part of the information is obtained as images with subsequent processing of these images by software subsystems [14, 15]. In particular, for implants of joints and teeth, their volumetric geometric model is required. Thus, an important competence for bioengineers is signal processing, in particular, processing of graphic images of various types [8, 13]. Therefore, students need to master the technologies for obtaining and processing of signals:

- to know and understand the fundamental concepts of digital signal processing, to master the use of digital filters for sound and image transformation;
- to possess skills of testing, data collection, and processing of digital signals for biomedical purposes, to apply various methods of transformation and analysis of signals in computerized systems of medical purpose;
- to know the methods of simulation and statistical processing of signals.

One of the most important issues in effective diagnostics is the quality of design, manufacture and operation of medical equipment [11, 13, 21]. The equipment provides coordination of measurements, processing of results, and transformation of the received information for visual presentation. Data are presented in the form of digital values,

graphs, images, sound signals. For the correct design and operation of medical equipment, students should:

- know definitions of the terms for medical systems equipment, trends and problems of their design, software for automated development of medical equipment;
- master and apply methods of mathematical modeling of medical equipment, create mathematical models that describe the operation of medical equipment, create software for modeling;
- carry out diagnostics of malfunctions, master methods of control and testing of medical equipment;
- understand the basic concepts in the field of nano- and microelectromechanical systems, the possibilities of their application in medicine, master their structure and device, as well as understand the principles of their computerized design and use.

A decision on the design or choice of an implant is made after processing the received medical data on the physical characteristics of the patient and his diagnosis [8, 9, 22, 23]. Depending on the type of implant, its mechanical elements are designed (A2) [10, 11, 13]. The electronic device is also designed according to the tasks of the prosthetics (A4), if necessary. At present, computer-aided design (CAD) and CAE systems (A2, A4) [10, 11, 21] are widely used to design implants, medical instruments and equipment. 3D scanners are often employed to build 3D models [12, 21]. Most often, implants are of a complex geometric shape, therefore, automated processing systems on CNC machines (A3, A5, A6) are used for production [11]. In addition to CAM systems, experiments are actively carried out on application of 3D printing to obtain implants of complex shapes. To use these technologies, students must be able to apply universal and specialized design systems:

- possess computer simulation skills in the design of medical equipment and implants, taking into account the individual anatomical characteristics of a person;
- develop and implement software for the creation and production of equipment and implants in MCAD / MCAM packages, integrate with these systems and work with 3D printers.

An important issue for the development of implants is the choice of materials for medical purposes [18, 19]. A number of requirements are imposed on these materials, which are associated with implantation inside the human body. Reasonable choice of materials in some cases is the basis for successful prosthetics. To use the materials (A2, A5, A6) correctly, students should:

- understand the requirements for the medical materials and products made of them, know physical and mechanical properties of the materials, as well as know the methods of their identification and statistical processing [24, 25];
- have a basic understanding of bioinertness (biocompatibility), electroneutrality, nontoxicity, tribological characteristics of the fatigue strength of materials used for implants [17, 18, 19].

Knowledge of biomechanics allows students to provide the necessary movements of an implant or prosthesis elements to perform the required functions, as well as to calculate loads of the structural elements. Knowledge of the loads and displacements of structural elements relative to each other allows to choose the right materials that ensure the strength and durability of implants, prostheses, and elements of medical equipment (A2). To properly design implants, students need to:

- to master the topographical anatomy of a person, master the means of modeling kinematics and dynamics of the human motor apparatus;
- set tasks and perform strength calculations using CAE - systems taking into account mathematical models of biomedical materials behavior [12, 21].

To master the processes of modeling of all components of implants and medical equipment (A2, A4), the competencies formed during the study of the disciplines of the mathematical group are of great importance [12, 16, 23, 25]. These include disciplines: Mathematical modeling, Mathematical methods of operations research, etc. As a result of mastering the disciplines of a specialized cycle of mathematical training, students should be able to:

- apply modern technologies for mathematical modeling of objects, processes and phenomena, be able to develop computational models and algorithms for the numerical solution of mathematical modeling problems, taking into account the errors of the approximate numerical solution of professional problems;
- carry out a formal description of research tasks in organizational, technical, socio-economic systems for various purposes; determine optimal solutions for the tasks.

Medical electronic devices perform the functions of processing the received information, controlling the operation of implants. The development of medical electronic devices requires knowledge of the physical foundations, purpose and principles of operation of electronic components. Computer-aided design systems for electronic devices ensure their development, in particular, the design of printed circuit boards (PCBs). The use of ECAD systems for these purposes (A4, A5) [8, 20] requires students to:

- develop medical systems using the results of measurements of various parameters of human organism subsystems;
- know the basic concepts, stages and technologies of design, manufacture and control of printed circuit boards;
- choose the necessary electronic components for the project, master practical techniques of designing printed circuit boards;
- apply automated means of electronic circuits and PCBs design.

Programming is of great importance in designing medical systems (A5). The programs implement the logic of receiving, processing and storing information, ensure the coordinated operation of intelligent implants and other medical devices. The programs implement mathematical methods of information processing in the form of appropriate

algorithms. In addition to the design of the mechanical and electronic parts of the implant, information processing, testing and assembly of all components of the product using robotic systems (A3, A5, A6) are automated.

The most difficult part of the technology of using implants is the installation of an implant into the human body (Fig. 4, A7). Robotic systems are also being developed and used to perform medical operations in an experimental mode. Of great importance in the design of medical systems is programming (A5). Programs implement the logic of obtaining, processing, and storing information, ensure the coordinated operation of intelligent implants, medical devices. Programs implement mathematical methods of information processing in the form of appropriate algorithms. Thus, today not only the design of the mechanical and electronic parts of the implant is automated, but the processing, testing, assembly of all components of the product using robotic systems (A3, A5, A6).

The most complex part of the technology of using implants is the installation of an implant in the human body (Fig.4, A7). But also for the performance of medical operations, robotic systems are developed and applied in an experimental mode.

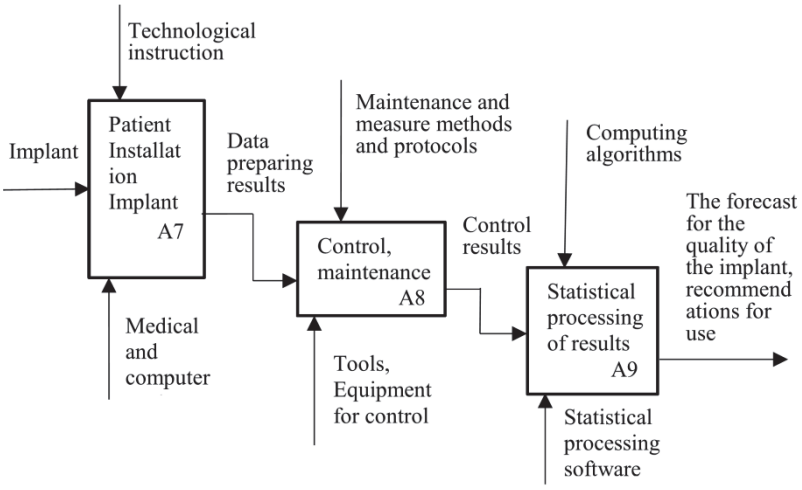


Fig. 4. The model of educational activity in the statistical processing of implant installation results

In medicine, information systems are widely used that provide information transfer at various levels, centralized storage (A8) [8, 9]. Automated information systems allow you to remotely monitor a patient's condition, perform diagnostics based on the use of network technologies, telemedicine, conduct statistical studies, and a number of other tasks (A9) [9, 14]. To do this, you must:

– to apply signal processing technologies, images and video to diagnose and monitor patient status, including the use of microcontrollers and wireless networking technologies for remote diagnostics and monitoring;

- to understand the principles of data transmission, know the types and routing protocols, master the design and development of new network architectures and MAC protocols;
- to master the application of various types of network infrastructure in telemedicine for the transmission of images, video, and remote interactions of equipment in biomedical diagnostic systems.

The collection of biomedical information and its mathematical processing make it possible to analyze and model processes both in the human body and in society (A9) [12, 14, 24, 25]. In this regard, the aim of the course «Mathematical processing of biomedical data» that students should be able to:

- understanding of the basic principles and methodology of processing statistical data / experimental and the ability to use them in practice: process data and interpret them;
- the ability to develop measures for the improvement, modernization of existing computer systems, and the selection of more effective technical and instrumental means with the necessary substantiation and calculations in the processing of biomedical data.

A developed set of competencies is not mandatory for full use. This set of competencies can be reduced or expanded depending on the basic specialty and profile of the university.

4 Conclusions

1. It has been determined that an important basis for information support of engineering activities is the effective presentation of knowledge and standardization of training tasks in investigated subject areas. The competency-based approach to the organization of the educational process and the requirements for training specialists from the enterprises and organizations in which they will carry out their professional activities require the development of a comprehensive model of the educational process based on high-level abstractions.
2. . The need for knowledge in the field of medicine for IT students has been substantiated, which motivates studying specific issues for the high-quality implementation of tasks in the field of bioengineering. The formalized implant development process has substantiated the need for medical knowledge for IT students.
3. Within the framework of the development of STEM - and other technologies of education, students study engineering disciplines that are necessary for the formation of a modern engineer and specialist in the field of information technology. In the process of learning, students master a number of technologies that are the basis for the implementation of modern software systems for various purposes: for engineering, medicine, and other fields of activity.
4. The model of the educational process has helped to identify and formulate requirements for the content and scope of the acquired knowledge and the results of the

acquisition of skills in the process of acquiring the students the competencies necessary for the implementation of the types and tasks of activity in the field of bioengineering, as envisaged by the project BioArt.

References

1. Task Group on Information Technology Curricula: Information Technology Curricula 2017. Curriculum Guidelines for Baccalaureate Degree Programs in Information Technology. New York, NY, USA: Association for Computing Machinery (ACM), 2017. ISBN: 978-1-4503-6416-4, DOI: 10.1145/3173161, Web link: https://dl.acm.org/citation.cfm?id=3173161
2. Curriculum Development: Foundations and Modern Advances in Graduate Medical Education. November 2018/ in book: Contemporary Topics in Graduate Medical Education [Working Title] DOI:10.5772/intechopen.81532
3. Computer Science Curricula 2013: Curriculum Guidelines for Undergraduate Degree Programs in Computer Science. ACM, IEEE Computer Society, 2013. https://www.acm.org/binaries/content/assets/education/cs2013_web_final.pdf.
4. Key Competences for Lifelong learning: A European Reference Framework. https://ec.europa.eu/education/education-in-the-eu/council-recommendation-on-key-competences-for-lifelong-learning_en.
5. Tarasov, O. Sahaida, P., Vasylieva, L.: Improvement of educational process based on software development for virtual and remote labs. Proceedings of the International Symposium on Embedded Systems and Trends in Teaching Engineering, Nitra: Constantine the Philosopher University, 220-224 (2016).
6. Smirnova, E.V. & Clark, R.P. (2018). Handbook of Research on Engineering Education in a Global Context (Advances in Higher Education and Professional Development). IGI Global. ISBN 978-1522533955.
7. Guizzardi, G. (2005). Ontological Foundations for Structural Conceptual Models. Veenendaal, The Netherlands, Universal Press.
8. Kyung, C.-M., Yasuura, H., Liu, Y. & Lin, Y.-L. (2017). Smart Sensors and Systems: Innovations for Medical, Environmental, and IoT Applications. Springer International Publishing.
9. Eren, H. & Webster, J. G. (2016). The e-medicine, e-health, m-health, telemedicine, and telehealth handbook. Volume II, Telehealth and mobile health. CRC Press.
10. Rusiński, E. & Pietrusiak, D. (Eds.). (2019). Proceedings of the 14th International Scientific Conference: Computer Aided Engineering (Lecture Notes in Mechanical Engineering). Springer. ISBN 9783030049744. DOI: doi.org/10.1007/978-3-030-04975-1.
11. Advances on Mechanics, Design Engineering and Manufacturing. (2017). Proceedings of the International Joint Conference on Mechanics, Design Engineering & Advanced Manufacturing (Lecture Notes in Mechanical Engineering) Springer International Publishing AG. ISBN 978-3319457802. DOI: https://doi.org/10.1007/978-3-319-45781-9.
12. Bibb, R., Eggbeer, D. & Paterson, A. (2015). Medical Modelling: The Application of Advanced Design and Rapid Prototyping Techniques in Medicine. Second edition. Elsevier Ltd. ISBN 978-1-78242-300-3 (print), ISBN 978-1-78242-313-3 (online).
13. Bronzino, J. D. & Peterson, D. R. (2015). The Biomedical Engineering Handbook, Fourth Edition - 4 Volume Set: Biomedical engineering fundamentals; Medical Devices and Human Engineering; Biomedical Signals, Imaging, and Informatics; Molecular, Cellular, and Tissue Engineering. CRC Press.

14. Artificial Intelligence and Machine Learning in 2D/3D Medical Image Processing. (2020). Rohit Raja, Sandeep Kumar, Shilpa Rani, K. Ramya Laxmi (Eds.). CRC Press. ISBN 978-0367374358.
15. Smith, N. B., & Webb, A. Introduction to Medical Imaging: Physics. Engineering and Clinical Applications 2011. ISBN: 978-0521190657.
16. Nettleton, D. (2014). Commercial Data Mining: Processing, Analysis and Modeling for Predictive Analytics Projects. Elsevier. ISBN 978-0-124-16602-8.
17. AlMangour, B. (Ed.). (2019). Additive Manufacturing of Emerging Materials. Springer. ISBN : 978-3-319-91712-2. DOI: https://doi.org/10.1007/978-3-319-91713-9.
18. Agrawal, C.M., Ong, J.L., Appleford, M.R. & Mani, G. (2013). Introduction to Biomaterials: Basic Theory with Engineering Applications. Cambridge University Press. ISBN: 978-1139035545. DOI: 10.1017/CBO9781139035545.
19. Zhang, L.-C. and Chen, L.-Y. (2019). A Review on Biomedical Titanium Alloys: Recent Progress and Prospect. Adv. Eng. Mater., 21: 1801215. https://doi.org/10.1002/adem.201801215.
20. Lee, E.A. & Seshia, S.A. (2017). Introduction to Embedded Systems, A Cyber-Physical Systems Approach, Second Edition. MIT Press. ISBN 978-0-262-53381-2, 2017.
21. Webster, J. G. (2020). Medical Instrumentation Application and Design, 5th Edition. CRC Press. ISBN: 978-1119457336.
22. Peate, I. & Evans, S. (2020). Fundamentals of Anatomy and Physiology For Nursing and Healthcare Students. Third Edition. John Wiley & Sons. ISBN: 978-1119576518.
23. Arus, E. (2017). Biomechanics of Human Motion. Applications in the Martial Arts, Second Edition. CRC Press. ISBN 978-1138555532.
24. Montgomery, D. C. (2017). Design and analysis of experiments. Eighth Edition. John Wiley & Sons. ISBN: 978-1118-14692-7.
25. Wu, C. J., & Hamada, M. S. (2011). Experiments: planning, analysis, and optimization (Vol. 552). John Wiley & Sons. ISBN: 978-0-471-69946-0.

Section 2: System design of artificial implants

Electrical Signals in Biosensors

Amos Bardea [0000-0002-0512-4120]

Holon Institute of Technology- HIT, Holon, Israel
amos.bardea@hit.ac.il

Abstract. The field of diagnostics is a cornerstone of modern health care especially in personal medicine. As the need for simple, rapid and cost effective medical care expands in the next few decades, it will become increasingly important to develop highly sensitive, specific, selective, portable and high throughput methods for detection of relevant biological entities. Such biosensor schemes are critical not only to garner vital patient information in a timely fashion but also to accelerate progress toward a better understanding of the complex biological pathways that govern disease. Electrical biosensors offer a simple and inexpensive device with computability to IT (Information Technology) world and IoT (Internet of Things). The major electrical biosensing strategies, including impedimetric, potentiometric, and amperometric detection. Implanted biochip are one of the gentle solutions for continuous monitoring diabetic diseases, metabolic changes following drug and chemotherapy treatment. This chapter will describe the principles of various types of electrical signals in Biosensors.

Keywords: Biosensor, Bioelectronics, Electrochemical detection, Quartz Crystal Microbalance.

1 Prologue

Integration of biology and electronics resulted in the scientific area of bioelectronics. Bioelectronics is actually a multi-facet scientific and technological area that includes electronic (or optoelectronic) coupling of biomolecules, or their natural or artificial assemblies, with electronic or optoelectronic devices [1]. Interfacing of biomaterials and electronic devices can be used to transduce chemical signals generated by biological components into electronically (or photonically) readable signals, or to activate the biomaterials by applying electronic (or optical) signals, thus resulting in the switchable/tunable performance of the biological components, Figure 1, [2]. The bioelectronic (optobioelectronic) systems can be used to develop biofuel cells, implantable biofuel cells for biomedical applications, self-powered biosensors [3]. Also, the bioelectronic systems are developed for Biocomputing applications. The general vision of biological systems as very complex systems transducing matter, energy and information, allows us to consider the operation of biochemical cycles as computing processes.
One of the main applications of the bioelectronic systems is sensing devices, such enzyme-based biosensors, DNA-sensors, immunosensors, etc [4, 5]. This chapter will describe the principles of various types of electrical signals in biosensors and biochips.

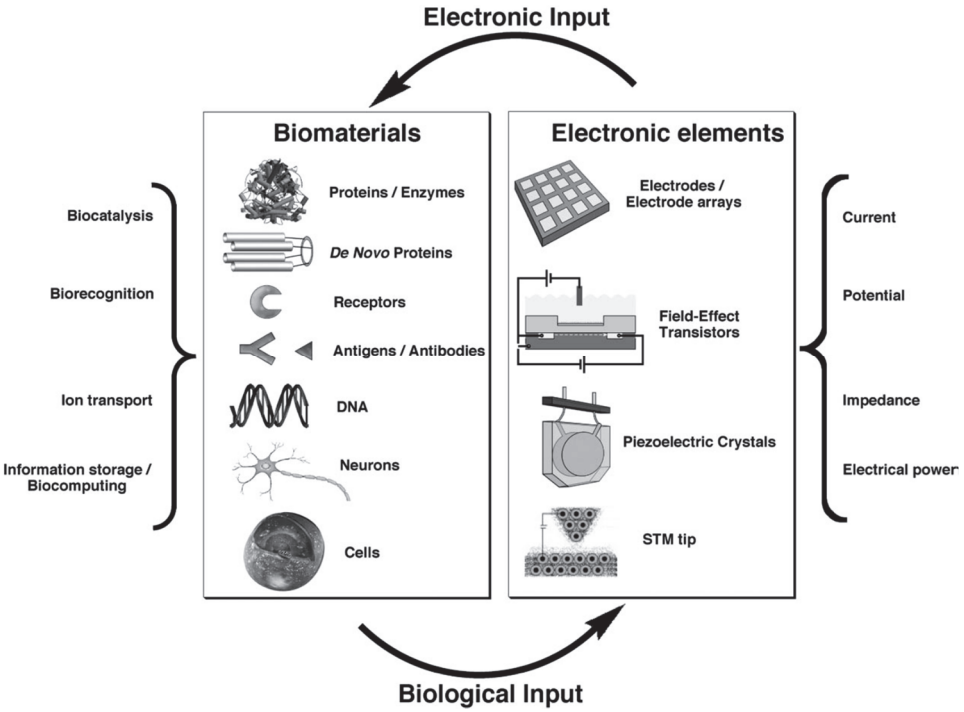


Fig. 1. The bioelectronics concept (adopted from [1]).

We require biosensors because we need to know what we eat (food quality control), what is our physical condition (health care), and who carries a bomb (security). In the 19th century people were producing the food by themselves or bought it directly from the farmers and they trusted the supplier. In the 19th century and even at the beginning of the 20th century knowing your health condition was not important since there was not much that could be done about it. The main mode of transportation was the horse and carriage. It wasn't until the latter part of the century that railways changed people's lives and habits. But even after the advent of the railway, remote areas still relied on the horse for local transport. People trusted their travelling companions. Today we have to regulate food quality, we need to conduct medical tests to fit medicines to patients and we need to be sure that our mean of transportation will be secure. Furthermore, the last decade medicine shifts the paradigm forward the personal medicine, which selects appropriate and optimal therapies based on the context of a patient's genetic content or other molecular or cellular analysis. This new model is requiring high throughput of the diagnostic process and testing. Hence, highly sensitive, specific, selective, portable, cost effective, real-time and multi analytes biosensor are require to address the need of the modern medicine and healthcare policy. In addition, sensitive biosensing is crucial in the forensic field when very low amount of bio molecules, such as DNA, need to be detected. High sensitivity of biosensor is related to the capability of producing a signal when the concentration of the analyte (sensing agent)

is very low and is $<10^{-10}$ Molar. The sensitivity of the sensor required for early detection of diseases, mainly viral, cancer and cardiovascular diseases. Specificity and selectivity are criteria of efficiency of the biosensor. The of signal-noise ratio (SNR) is important criteria for any validation of sensor. The high frequency signals of false positive or false negative cause the sensor to be ineffective. The criteria of portability, user friendly and cost effective are related to the needs of remote medicine and point-of-care tests that can be performed at the bedside or near the patient place. This contrasts with the historical pattern in which testing was wholly or mostly confined to the medical laboratory, which entailed sending off specimens away from the point of care and then waiting hours or days to learn the results, during which time care must continue without the desired information [9].

The high throughput diagnostic systems currently not meet the criteria above, since facilities required are big and expensive. Spectrophotometry is the quantitative measurement of the reflection, absorbance or transmission of light, and require photodetector for electromagnetic signal readout. The same is a fluorescence spectroscopy system that analyzes fluorescence signal from a sample. Enzyme Linked Immunosorbent Assay-(ELISA) are based as well on photometric system and require an optical signal readout system. Same the chemo luminescence that measuring the emission of light as the result of a chemical reaction require an optical signal readout system. In contrast a biosensor device applying electrical signal is suitable and computable for producing portable, rapid, real time detection and cost effective devices with interface to online IT world.

2 Biosensor – Principles of operation

The sensing process of biological matter applying electrical signals has a need to couple between two systems with different languages. Unlike the solid phase and electrical conductive material of electronic devices, biological material is mostly in liquid solution and has a very low electronic conductance. The challenge of designing and manufacturing biosensors is the ability to translate biological activity into electrical signal or express the concentration of chemical or biological agents by electrical signal.

Biosensors consist the bio-mediator and transducer, both transfer the test agent, named “Analyte” to an electrical signal (Fig.2). The bio-mediator consists matrices of probes that interact with the analyte. The specificity and selectivity of the bio-sensing depends on the specificity and selectivity of the interaction between the matrices and the analyte based on bio-recognition of complementation of the biomaterial, such as Enzyme-Substrate, Antigen-Antibody, Receptor-Ligand and complementary of two strands of DNA (Deoxyribonucleic Acid), RNA (Ribonucleic Acid) and PNA (Peptide nucleic Acid). The matrices are part of the bio-mediators and usually consist conductive or semi-conductive support, and the electronic transduction of the biological functions associated with the biological matrices. The basic feature of a bioelectronic device is the immobilization of a biomaterial onto a conductive or semi-conductive support.

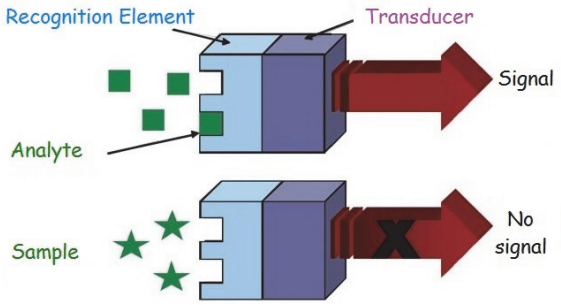


Fig. 2. A scheme of electrical biosensor configuration. (https://loosa.paginas.ufsc.br/en/biossensores/)

The electrical biosensing signals are including impedimetric that measure the electone transfer rate, potentiometric, amperometric, capacitance, conductivity, semiconductor field effect transistor (FET), microgravimetric Quartz Crystal Microbalance (QCM) and electrical power.

Table 1. Classification of biosensor.

Probe	Signal
❖ Redox Enzyme Sensor ❖ Immunosensors ❖ DNA sensor ❖ Receptor / Ligand ❖ Cells sensors	❖ Potential /Current (potentiometric / amprometric) ❖ Capacitance ❖ Conductivity ❖ Impedance ❖ FET-Field Effect Transistor ❖ Electrical Power / AFM, STM ❖ Crystal Microbalance Microgravimetric Analyses (QCM)

3 Biosensor – Biomediator

The immobilization of a biomaterial onto a surface of substrate of bio mediator is obtained by surface chemistry. Three main strategies are applied: a. Physical adsorption b. Chemical covalent binding c. biorecognition, (Fig. 3), [6]. The interacting force of physical adsorption (physisorption) is weak bonds such hydrogen bond, electrostatic bond, dipole-dipole interaction or Van der Waals (VDW) forces. Even though the interaction energy is very weak (~10–100 meV), the multi sites interaction is able to immobilize the molecule on the surface. Another method for immobilizing biomaterial onto surfaces using physical adsorption is obtained by anchoring the biomaterial into a polymer (Fig. 4).

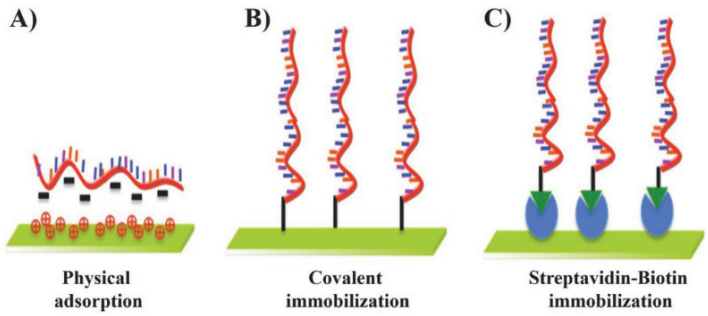


Fig. 3. Immobilization methods of DNA. (adopted from [6])

Chemical covalent binding is obtained by binding the molecule covalently to the surface by chemical bonds. Covalent linkage of proteins to conductive or semiconductive supports often utilizes the availability of functional groups on the surface of the solid support. For example, metal-oxide materials such as TiO_2, SnO_2, contain surface hydroxyl groups that are synthetically useful for the coupling of organic materials. Noble metals (Au, Pt) were chemically or electrochemically pre-treated to generate such surface hydroxyl functions. Carbon electrodes, after appropriate chemical treatment, contain different surface-associated functional groups (carboxylic, carbonyl, lactone, hydroxyl, etc.) capable of attachment to proteins [7].

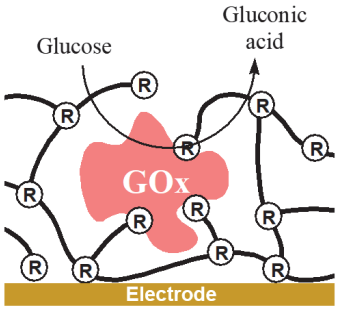


Fig. 4. Physical adsorption of redox enzyme (Glucose Oxidase-GOx) onto surfaces by anchoring the enzyme into redox polymer.

Biomolecule monolayer is assembling and bind on the surfaces using direct or non-direct process. For example: thiolated monolayers is self-assemble (SAM) on top of gold surface since the thiol group bind to gold (Au) by chemical interaction. Functional

group (X) such amino or carboxylic acid-functionalized thiolate monolayers associated with Au electrodes are covalently coupled to complementary functional group (carboxylic or amino groups which create amide bond) of bio molecule such protein or DNA (Fig.5). The SAM is dictate the binding of the chain molecules in same direction.

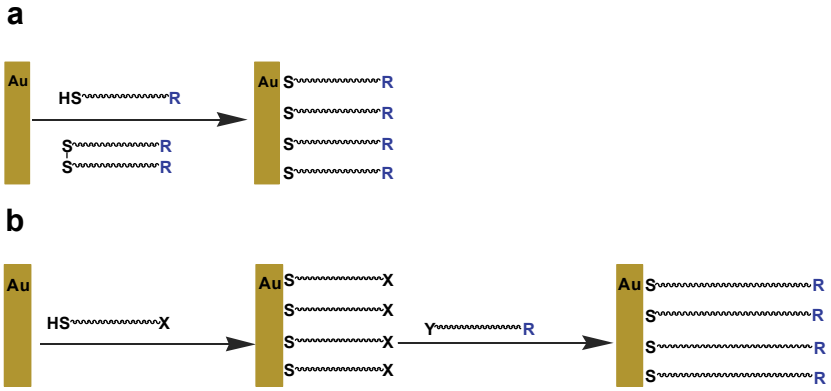


Fig. 5. Direct binding of biomolecules (R) (a) or non-direct (b) process Self-assembly of functionalized, thiolated monolayers by thiol group (SH) or disulfide (S-S) on Au electrodes. X is a functional group for coupling with a biomolecule. Y is a complementary functional group to X.

The immobilization biomolecules onto the surface is appling by harness the biorecognition process, such as Biotin-Avidin recognition. The biorecognition is taking place when the two complementary biomaterial interact like “the lock and key”.

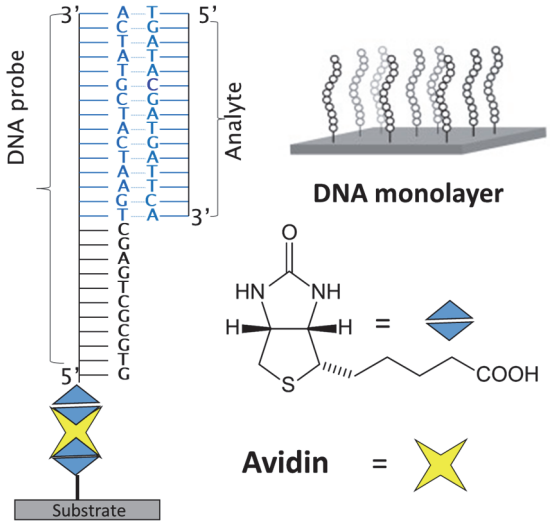


Fig. 6. Self-assemble DNA probe monolayer onto surface of the substrate by Biotin-Avidin interaction.

The hen egg-white Avidin protein has very high affinity for Biotin (vitamin B7) with high specificity and selectivity. Each Avidin protein consists of four binding sites for four Biotin molecules. The two polar sites allow anchoring the biomolecules onto the surface as described in figure 6. The advantage of this method is the ability to reproduce the surface probe by rinsing the surface with a solution that denature the Avidin protein. The disadvantage of this method is the stability over the time [8].

Classification of the types of biosensor can be accomplished by the type of probe or by type of signal according the table 1 and figure 1. The probes are part of the matrices and bind to the transducer that execute the electric signal. It can base on the reduction and oxidation (Redox) reaction and mostly catalyzes by redox enzyme. The Immunosensors are based on Atigen-Antibody (Ag-Ab) complex. In case the analyte is Ab the probe is Ag and vice versa. The DNA/RNA sensor presents the same strategy when the single strand DNA (ssDNA), a mostly oligo nucleotide is bound to the materials and used as the probe to detect the analyte of a complementary strand (Fig. 6). In the same method biosensor can be design for any bioreceptor and ligands such Biotin-Avidin interaction. Furthermore, biosensing can be obtained to the entire cell, such nerve cell for measuring their signal processing or other cell for measure the metabolic process according the cell environment.

4 Transducer and Electrical Signals

The transducer of the biosensor is required to express the concentration of analytes, biorecognition activity and biochemical process to electrical signal. In spite that biological materials are not conductive they can involve on the electrochemical process. Many biosensors signal utilizes this feature to generate an electrical signal. The electrochemistry is the reciprocal reaction between chemical energy to electrical energy. The main instrument for applying the electrochemical measurement is the electrochemical cell. In this section we will describe the biosensing based on electrochemical detection and biosensing based on Microgravimetric quartz-crystal-microbalance (QCM) analyses.

Electrochemical detection

Blood glucose monitoring reveals individual patterns of blood glucose changes, and helps in the planning of meals, activities, and at what time of day to take medications. Also, testing allows for quick response to high blood sugar (hyperglycemia) or low blood sugar (hypoglycemia). This might include diet adjustments, exercise, and insulin (as instructed by the health care provider) [12]. The most typical example of first-generation biosensor is the Clark electrode. Leland C. Clark Jr. (1918–2005) was an American biochemist and the inventor of the Clark electrode: The electrode is covered with a GOx-soaked membrane. The measured compound here is oxygen, the concentration of which changes with changed enzyme activity depending, in turn, on glucose (analyte) concentration. The Clark cell measure the glucose in solution by the amprometric method. Amperometric-based electrodes measure the current in the electrochemical cell

instrument. The current is correlated to the concentration of the substrate. Glucose biosensor measure the analyte concentration in non-direct way since the current is correlate to the concentration of oxygen based on equations 1 and 2.

The glucose is the substrate (S_{red}) of GOx (E). GOx is catalyzed by oxidation of glucose to producing gluconic acid (S_{ox}) and reduction the oxygen to producing hydrogen peroxide by electron transfer from glucose to hydrogen peroxide according the equation 2.

$$S_{red} + O_2 + 2H^+ \xrightarrow{E} S_{ox} + H_2O_2 \quad (1)$$

$$\beta - D - glucose + O_2 + 2H_2O \xrightarrow{GOx} D - gluconicacid + 2H_2O_2 \quad (2)$$

Hence, concentration of glucose is proportional to the O_2 consumption, which is quantified by the resulting electrical current.

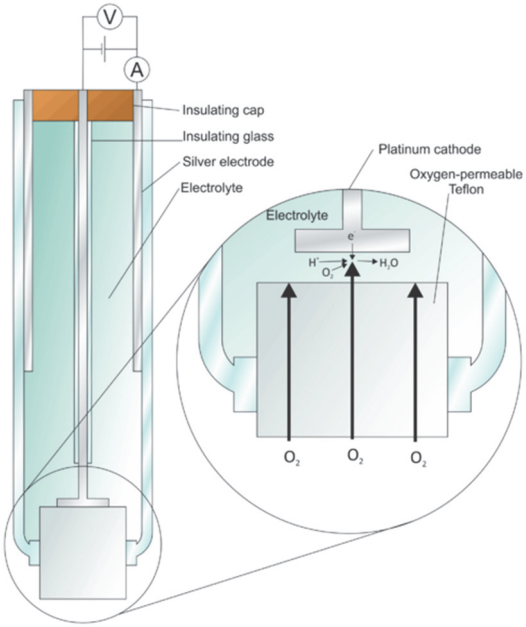


Fig. 7. Schematic representation of the Clark-based metabolite electrode. A permselective membrane covers the Pt electrode, the membrane is impregnated with immobilized glucose oxidase (GOx)

Basic amperometric electrode is the oxygen electrode (Clark electrode), which consists of an oxygen-permeable membrane covering a platinum cathode. The metabolite electrode has an additional membrane (gel layer) with an immobilized enzyme, which is in contact to the O_2 permeable plastic membrane; this arrangement allows the set to become permeable to both, oxygen and the specific metabolite, and the electrode develops a current proportional to the analyte. Figure 7 shows the schematic representation of the construction of a metabolite electrode. In one half-cell oxygen reduces by platinum (Pt) anode and in the second half-cell silver cathode (Ag) oxidize as described

in equations 3 and 4 respectively and the net reaction in equation 5. Each redox reaction transfers 4 electrons into the circuit. The current is proportion to the concentration of the reactants and to the reaction rate restricted by diffusion and penetration of the analytes and oxygen.

$$O_2 + 4H^+ + 4e^- \rightarrow 2H_2O \quad (3)$$

$$4Ag + 4Cl^- \rightarrow 4AgCl + 4e^- \quad (4)$$

$$O_2 + 4H^+ + 4Ag + 4Cl^- \rightarrow 4AgCl + 2H_2O \quad (5)$$

The voltage applyed on the Clark cell is -0.6 V, it calculates based on standard electrode potential (E°) required for a reduction or oxidation occur in each half-cell refer to hydrogen half-cell, H_2/H^+, under standard condition: 1 molar in aqua solution, in standard temperature (25°C) and pressure (1 atm) – STP. The potential is fit to the real concentration according Nernst equation 6. For a complete electrochemical reaction, the equation can be written as where E is the cell potential (electromotive force) at the temperature of interest, E° is the standard cell potential, R is the universal gas constant (8.3144 J K^{-1} mol^{-1}), T is the temperature in Kelvins, n is the number of electrons transferred in the cell reaction or half-reaction, F is the Faraday constant, the number of coulombs per mole of electrons (F = 96485.3321C mol^{-1}). Q is the reaction quotient of the cell reaction describe in equation 7. Equation 8 expresses the Nernst equation by logarithmic form.

$$E = E° - \frac{RT}{nF} \ln Q \quad (6)$$

$$\alpha A + \beta B \rightarrow \rho R + \sigma S \implies Q = \frac{[R]^\rho [S]^\sigma}{[A]^\alpha [B]^\beta} \quad (7)$$

$$E = E° - \frac{2{,}303RT}{nF} log_{10} Q \quad (8)$$

Three-electrode cells (see Fig. 8) are commonly used in controlled potential experiments. The cell is containing the three electrodes working electrode (WE), reference electrode (RE), and counter electrode (CE).

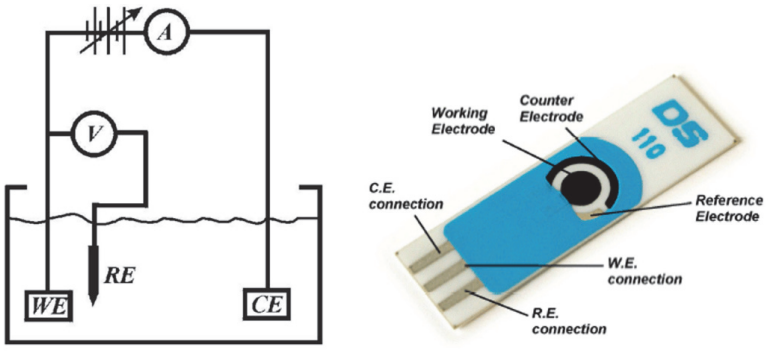


Fig. 8. Schematic representation of the three electrodes electrochemical cell and the implementation cell in a chip. (Metrohm DropSens company and http://www.porous-35.com)

Electrodes are immersed in the analyte solution. WE is reacting with analytes. RE provides a stable and reproducible potential and it insulated from the sample through an intermediate bridge to minimize contamination of the sample solution. The potential of WE is compared against RE. The RE is used as electrochemical half-cell and is "buffering" against potential changes. The buffering is achieved by a constant composition of both forms of its redox couple, such as silver–silver chloride (Ag/AgCl) or saturated calomel reference electrodes (Hg/Hg_2Cl_2). An inert conducting material, such as platinum wire or graphite rod, is usually used as the current-carrying counter auxiliary electrode (CE). Analyzer (or potentiostat) instrument can apply various techniques of measurement such applying voltage versus RE to the WE and measuring current between WE and CE or vice versa.

The most widely used technique for acquiring qualitative information about electrochemical reactions are linear sweep voltammetry (LSV) and cyclic voltammetry (CV). Linear sweep voltammetry is a voltammetry method where the current at a WE is measured while the potential between the WE and a RE is swept linearly in time. Oxidation or reduction of species is registered as a peak or trough shape in the current signal at the potential at which the species begins to be oxidized or reduced, while the CV results rapidly provide considerable information on the thermodynamics of redox processes and the kinetics of heterogeneous electron transfer reactions and on coupled chemical reactions or adsorption processes. CV consists of scanning linearly the potential of a stationary, using a triangular potential waveform (Fig. 9). The resulting current versus potential plot is termed a cyclic voltammogram [11].

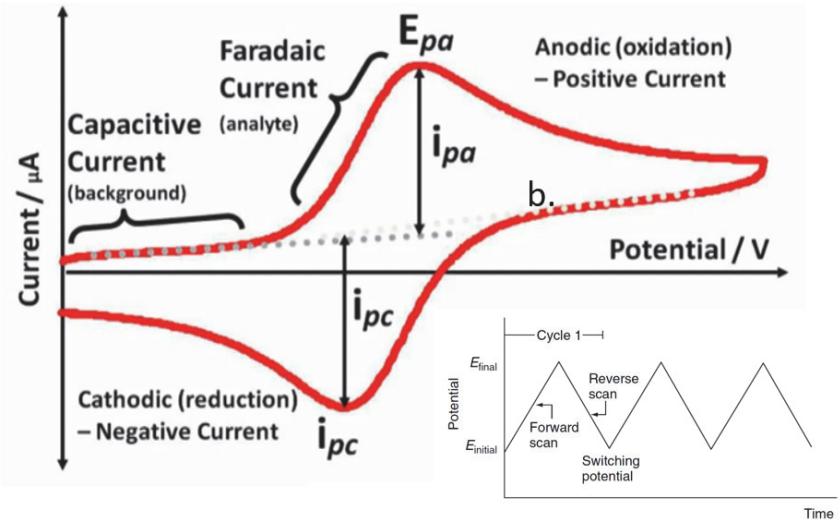


Fig. 9. Typical cyclic voltammetry (CV) for a reversible redox process**,** the insert plot shows the potential vs. time signal in a CV experiment. (Patel, B. University of Brighton).

The increasing demand for highly sensitive and selective glucose sensors in industrial, clinical and pharmaceutical sectors, especially for the effort to develop artificial pancreas, sensing materials for glucose detection have attracted much attention worldwide.

The first generation glucose biosensor is based on monitoring the generation of H_2O_2 and the O_2 consumption such as Clark glucose biosensor, (Fig.11). The reaction is restricted by the O_2 because the normal O_2 concentration is 1 order of magnitude lower than the normal physiological level of glucose, [10]. The second generation uses additional mediators to enhance the electron transport from the GOx enzyme active sites to the surface of the electrode, (Fig.11). Flavin adenine dinucleotide (FAD) is a redox-active cofactor associated with GOx at his active site. The biocatalytic reaction of GOx includes the reduction of Flavin group in the enzyme GOx_{FAD} by reacting with glucose to form the reduced state of the cofactor GOx_{FADH2} (Fig.10).

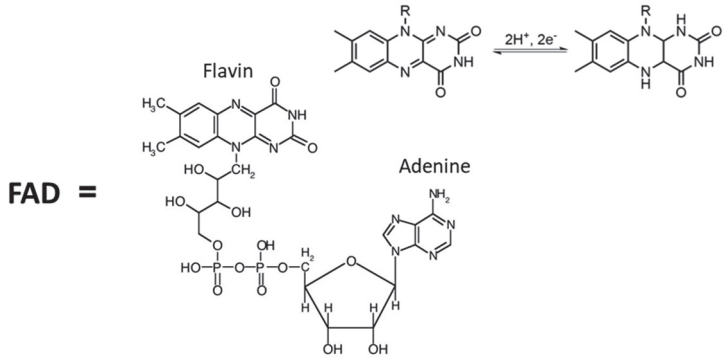


Fig. 10. Cofactor flavin adenine dinucleotide (FAD) in two state of redox process.

The electrons transfer from substrate to FAD to form $FADH_2$ and from $FADH_2$ to the $Mediator_{ox}$ to form the reduction state of it and from $Mediator_{red}$ to electrode to apply current that is proportional to the concentration of the glucose. This is one cycle of the redox cascade process that transfer two electrons to anode and reproduce the $Mediator_{red}$ and $FADH_2$ to the original oxidation state, FAD and $Mediator_{ox}$ respectively. The current signals of the oxidation of mediator are used to detect the concentration of glucose.

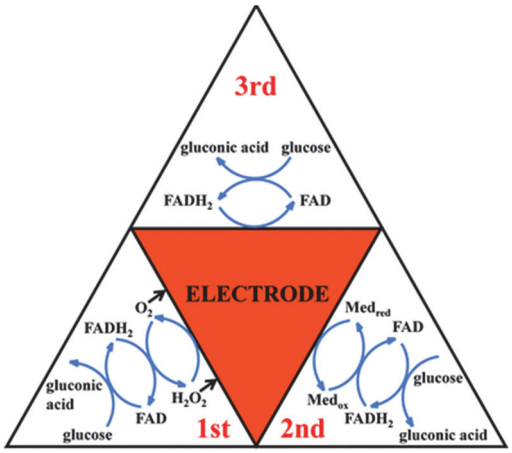


Fig. 11. Scheme describe the first, second and third generation enzymatic glucose sensors. [12]. Copyright (2013) Royal Society of Chemistry.

The third generation glucose sensor is based on the direct electron transfer between GOx active site and the electrode surface in the presence of mediators, (Fig.11), [12]. In a native reaction oxygen oxidize the $FADH_2$ to produce FAD and H_2O_2, while in the third generation glucose sensor the anode electrode replaces the oxygen in the oxidation process. For this process the system should evacuate the oxidation from the solution by dissolving other inert gas such as nitrogen or Argonne.

Direct electrical activation of redox enzymes, to stimulate the bioelectrocatalyzed oxidation (or reduction) of the enzyme substrates provided electron transfer between the electrode and the redox enzyme. This electron transfer is fast, resulting current corresponds to the turnover rate of the electron exchange between the substrate and the biocatalyst. Hence, the transduced current reflects the substrate concentration in the system.

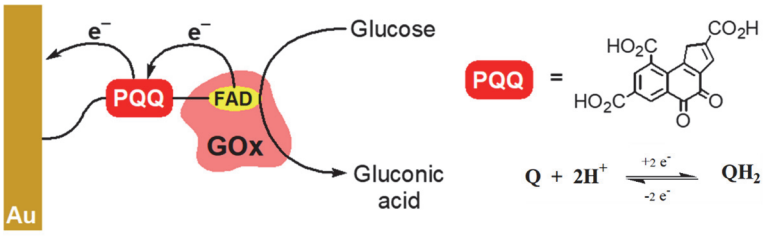


Fig. 12. Reconstitution of GOx on Quinone - Q (Pyrroloquinoline Quinone –PQQ) mediator monolayer, assembled on a Au electrode.

Direct electrical communication between redox, proteins, and specifically redox enzymes, and electrodes is, however, usually prohibited, [13]. The donor and acceptor separation distance is a major factor that controls the electron transfer rates. In fact, for most redox proteins (diameter 80-150Å) the redox centers are electrically insulated. That is, redox enzymes (or proteins) that form a donor-acceptor pair with an electrode lack electrical contact with this electrode, [7]. The tailoring of bioelectronics systems requires the assembly of biomaterials on electrode and the design of the appropriate electronic communication between the biological matrices and the support element. The tailoring the mediator and the redox enzyme onto the surface of the electrode is enforce liner transfers of the electrons from the active site of the redox enzyme to the electron, as describe in figure 12. The concentration dependent CV and the calibration curve corresponding to the amperometric responses of the enzyme electrode at different concentrations of glucose is shown in figure 13.

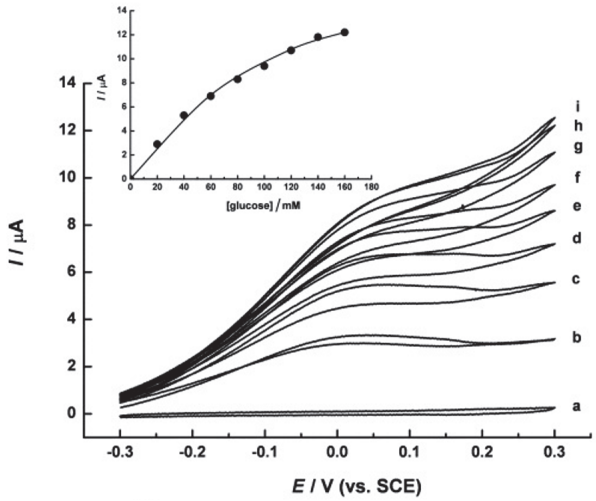


Fig. 13. Cyclic voltammograms of the GOx on an Au-electrode and the calibration curve corresponding to the amperometric responses of the enzyme electrode at different concentrations of glucose. (adopted from [17])

Microgravimetric quartz-crystal-microbalance

Microgravimetric quartz-crystal-microbalance (QCM) analyses based on mass change of the crystal. QCM is determining the mass variation per unit area by measuring the change in frequency of a quartz crystal resonator (Fig.14).

$$\Delta F = -\frac{2f_0^2}{A\sqrt{\rho_q \mu_q}}\Delta m \qquad (9)$$

The frequency shift induced by a thin sample which is the rigidly coupled to the crystal, is described by the Sauerbrey equation 9, where f_0 is resonant frequency, ΔF is frequency change, Δm is the mass change, A is the active crystal area, ρ_q is the density of quartz and μ_q is the shear modulus of quartz.

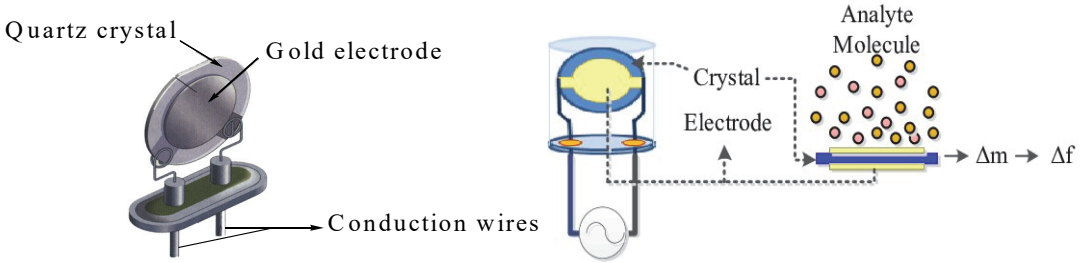


Fig. 14. Quartz Crystal Microbalance, (adopted from [16])

The QCM based biosensor have been developed for immunosensors and DNA sensors, Fig.15. The immobilized oligonucleotide probe onto the QCM electrode for the micro-gravimetric, piezoelectrical analysis of the complementary oligonucleotide analyte is provide a sensitive DNA sensor. Figure 15b shows the sensing process of the analyte. The hybridization of the oligonucleotide analyte with an oligonucleotide probe on top of the QCM electrode results in a mass change and causes the crystal frequency de-crease [14,15].

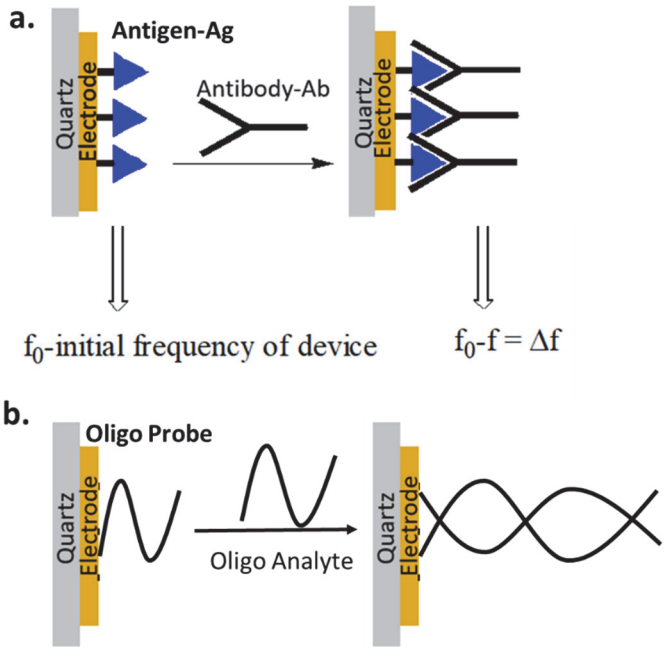


Fig. 15. Quartz Crystal Microbalance sensor: a. immunosensor and b. DNA sensor.

The extent of the crystal frequency decrease is enhanced as the bulk concentration of oligonucleotide analyte increases, Fig.16.

$$[Probe] + [Analyte] \leftrightharpoons [Probe - Analyte] \qquad (10)$$

$$k_a = \frac{[Probe-Analyte]}{[Probe]\times[Analyte]} \qquad (11)$$

The concentration dependent respose of the biosensor is based on the equilibrium constant of formation the complex of Probe-Analyte on the electrode surface, equation 10. The amount of the recognition location in the matrices probes is fixed while the concentration of the analyte changes. Since association equilibrium constant, k_a, and probe concentration is constant, the creation of the Probe-Analyte complex is proportional to the analyte concentration according equation 11.

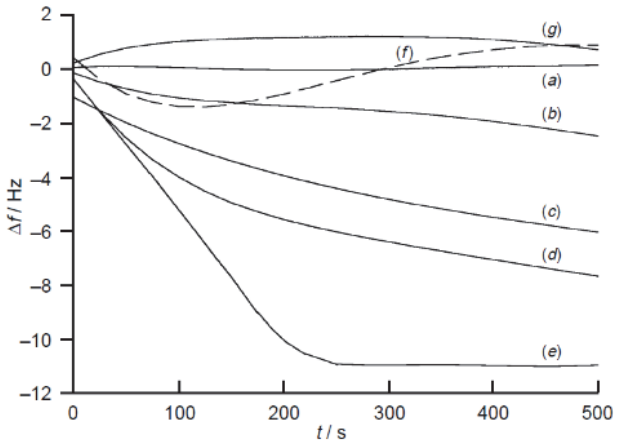


Fig. 16. Time-dependent frequency changes of a functionalized Au–quartz crystal by oligonucleotide probe upon addition of: (a) buffer solution;(b-e) increase the oligonucleotide complementary analyte respectively; (f) high concentration of non-complementary oligonucleotide, [14].

5 Epilogue

This chapter described a few methods and techniques for biosensing based on electrical signal. There are viruses of techniques and methods of electrical biosensing that are not described in this chapter while discussed, classified and presented in figure 1 and table 1. In addition, the miniaturization, nanostructure based biosensor, biosensors embedded on biochips (Lab on chip-LOC) and microfluidic related to sensor bio implant are not discussed. Furthermore, Biosensing based on optoelectronic and bio photonic such as Surface Plasmon Resonance (SPR) should be discussed. All these topics should be covered in the other chapter.

References

1. Willner, I., Katz, E.: Bioelectronics from Theory to Applications. Wiley-VCH, Weinheim, Germany (2005).
2. Hoffmann, K.-H.: Coupling of Biological and Electronic Systems. Springer, Berlin (2002).
3. Bullen, R. A. Arnot T. C., Lakeman J. B., Walsh F. C.: Biofuel cells and their development. Biosens. Bioelectron. 21, 2015-2045 (2006).
4. Gizeli, E. Lowe C. R.: Biomolecular Sensors. Taylor and Francis, Cambridge (2002).

5. Willner, I., Katz, E., Willner, B.: Layered Functionalized Electrodes for Electrochemical Biosensor Application. In Biosensors and Their Applications, (Eds: V. C. Yang, T. T. Ngo), pp. 47–98. Kluwer, New York (2000).
6. Nimse, S.B., Keumsoo, S., Sonawane , M.D., Sayyed, D.R., Kim, T.: Immobilization Techniques for Microarray: Challenges and Applications. Sensors 14(12), 22208-22229 (2014).
7. Willner, I., Katz, E.: Integration of Layered Redox Proteins and Conductive Supports for Bioelectronic Applications. Angew. Chem. Int. Ed. 39, 1180-1218 (2000).
8. Bardea, A., Burshtein, N., Rudich, Y., Salame, T., Ziv, C., Yarden, O., Naaman R.: Sensitive Detection and Identification of DNA and RNA using a Patterned Capillary Tube. Analytical Chemistry 83, 9418-9423 (2011).
9. Reddy, B., Salm, E., Bashir, R.: Electrical Chips for Biological Point-of-Care Detection. Annu. Rev. Biomed. Eng., 18:329–55 (2016).
10. Pauls, A., Moharil, P., Ravichander, P., Veerappapillai, S.: First Generation Amperometric Glucose Biosensor for Determination of Glucose at Room Temperature Using Glucose Oxidase. International. Journal of Pharmaceutical Sciences Review and Research 38(2), 102-105 (2016).
11. Wang, J.: Analytical Electrochemistry. 3rd edn. John Wiley & Sons, Germany (2006).
12. Zhang, C., Zhang, Z., Yang, Q., Chen, W.: Graphene-based Electrochemical Glucose Sensors: Fabrication and Sensing Properties. Electroanalysis 30, 2504–2524 (2018).
13. Chen, Q., Kobayashi, Y., Takeshita, H., Hoshi, T., Anzai, J.: Avidin-Biotin System-Based Enzyme Multilayer Membranes for Biosensor Applications: Optimization of Loading of Choline Esterase and Choline Oxidase in the Bienzyme Membrane for Acetylcholine Biosensors. Electroanalysis 10, 94-97 (1998).
14. Bardea, A., Dagan, A., Ben-Dov, I., Amit, B., Willner I.: Amplified Microgravimetric Quartz-Crystal-Microbalance Analyses of Oligonucleotide Complexes: A Route to a Tay-Sachs Biosensor Device. Chemical Communications, 839-840 (1998).
15. Bardea, A. Katz, E. Willner I.: Probing Antigen-Antibody Interactions on Electrode Supports by the Biocatalyzed Precipitation of an Insoluble Product. Electroanalysis 12 (14), 1097-1106 (2000).
16. Rivai, M. Arifin, A. Agustin, E. I.: Mixed vapour identification using partition column-QCMs and Artificial Neural Network. Conference: 2016 International Conference on Information & Communication Technology and Systems (ICTS) (2016)
17. Yehezkeli, O. Yan, Y.M. Tel-Vered, B.R. Willner, I.: Integrated Oligoaniline-Cross-Linked Composites of Au Nanoparticles/Glucose Oxidase Electrodes: A Generic Paradigm for Electrically Contacted Enzyme Systems. Chemistry A European Journal 15, 2674-2679 (2009).

Photoelectric measuring transducers in environmental and objects monitoring systems

Oleg Subotin[1][0000-0002-6095-5840], Vladislav Rudenko[1][0000-0002-2336-6609]
Anton Cherniavskyi[1][0000-0001-6594-1600], Andriy Kovalenko[1][0000-0003-3379-2000],
Serhii Dobriak[1][0000-0002-1568-4359]

[1] Donbass State Engineering Academy, Kramatorsk, Ukraine
app@dgma.donetsk.ua

Abstract. The paper presents the theoretical foundations for calculating the primary measuring transducer of the optical type, operating as an optical locator – the design model of the optical type transducer and analytical calculation. The results of a study of a photoelectric measuring device and a signal propagation medium control system based on it are shown. An increase in the reliability of control when using measuring transducers based on optical signals can be achieved by increasing the signal power or the length of the pulse sequence. This increases energy consumption and decreases control performance. Therefore, the development of methodological, structural, algorithmic and hardware means to increase the reliability and speed of the measuring channels of photoelectric converters in order to increase the efficiency of control systems is an urgent task. It is solved by using digital methods for processing optical information signals, for example, by coding signals and minimizing the code length. This increases the reliability of monitoring the parameters of objects and the environment, as well as the speed of information processing based on them. A converter has been developed and investigated for receiving and processing primary information from sensors for monitoring systems of object parameters and environmental monitoring. The analysis of modular structures of photoelectric measuring transducers for utilization in monitoring systems of objects and signal propagation medium is carried out.

Keywords: Photoelectric Primary Measuring Transducer, Optical Signal Propagation Channel, Measuring Channel, Signal Coding, Reliability of Control, Information Measurement System.

Introduction

There were different analyses of automation systems for monitoring the parameters of objects, the environment and information support, as well as external optical and electromagnetic interference that impede the operation of primary measuring transducers. These analyses showed that the amount of necessary primary information is determined by the technical tasks of monitoring: the presence of objects and their position, geometric and dimensional parameters, speed, temperature, environmental pollution.

Their implementation is carried out using specialized primary converters based on various physical operation principles. This determines their large number and variety [1,2].

The utilization of analog informational signals in measuring transducers is hampered by the presence of predominantly continuous disturbing influences. Using single pulses, the energy of the transmitted signals is equivalent, and in some cases significantly less than the energy characteristics of the medium of their propagation. When such devices operate under interference, the probability of receiving false information or missing a useful signal is significantly increased. The appearance of these control errors is related to the low noise immunity of the measuring transducers and, as a consequence, the low reliability of the received primary information [3].

On the other hand, the automation of measurements of fast-flowing processes determines increased requirements for the speed of the information system. It is known that the efficiency of the control system is defined by the speed and quality indicators of the used primary converters.

However, the existing optical (photoelectric) converters of analog and pulse type are largely subject to the influence of external optical and electromagnetic interference, leading to a modification of the information signal. It has been found that the signal-to-noise ratio is increased by synchronous accumulation of a "package" of pulses, which enlarges the reliability of the measurement information. At the same time, a large number of information pulses in this "package" (several hundred) significantly reduces the speed of the information system.

Thus, increasing the speed and reliability of the primary means of monitoring the technological parameters of objects is an important task.

1 Theoretical basis of the calculation of the primary measuring transducer of optical type, operating as optical locator

1.1 The mathematical model of the primary measuring transducer of optical type

Theoretically considered primary measuring transducer (PMT) of optical type is generally based on establishing the connection between the input variable x_{in} and the light emission power distribution P(x) at a specific propagation medium [3,4].

The length of the optical signal propagation channel L_c is much more than the light wavelength, so it is accurate enough for practical use in the method of calculating the light distribution of these communication channels and may be obtained by using the laws of geometric optics. However, using an optical signal in the near infrared wavelength range and analytical relationships for visible light when calculating the static characteristics of the PMT as the optical type primary measuring transducers information is incorrect.

According to the general theory of optical transducers [5,6] for the mathematical description the concept of the generalized forces F and speeds υ is used. In theory, the

selection of F and υ may be arbitrary, but is practically determined by available functions. In this case, it is advisable to move from concept of the light flux to the radiation power – $P(x)$. Then, basing on the well-known expression

$$P = U \cdot I = I^2 \cdot R, \tag{1}$$

where P - power, W; U - voltage, V; I - current intensity, A, R - resistance, Ohm.

It is convenient to choose the current I as the generalized force, and the generalized speed is the power P.

By analogy with the expression of geometrical optics, the binding force of the light flux J to propagating Φ within an angle Ω is

$$\Phi = J \cdot \Omega. \tag{2}$$

Generalized set of parameters will be interpreted as follows:

$I^2 \sim J$ – generalized force of the light flux (intensity);

$P \sim \Phi$ – generalized radiation power;

$R \sim \Omega$ – optical resistance, $\Omega = 1/S$, where S – the area bounded angle Ω.

Calculation of the distribution of radiation power in the optical signal propagation channel (OSPC) with a focused radiation source is a very complicated problem, because $P(x)$ depends on the channel length (L_c), the solid angle of radiation (Ω), the reflection coefficient (K_{ref}) from the source parameter, its installation space and other parameters. PMT design scheme with a concentrated light source is shown in Fig. 1.

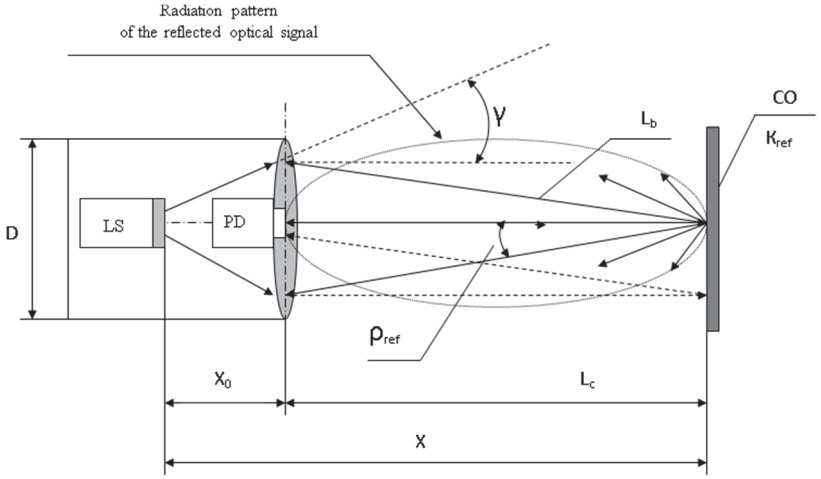


Fig. 1. Diagram of the PMT of optical type with a concentrated light source. The following designations are taken: LS – light source (emitter), PD – photodetector, CO – controlled object, Kref – the reflection coefficient of the object surface, γ - beam exit angle, D – the diameter of the lens aperture, Lc, Lb – the length of the channel and the beam, respectively, X0, X – distance from the emitter to the lens and object, respectively.

For the PMT of optical type it is convenient to consider separately the power of the direct reflected radiation $P_d(x)$ and reflected at an angle to the lateral surface of the

control object $P_a(x)$, which are components of the total power $P(x)$ that goes to photo detector (PD)

$$P(x) = P_d^{ref}(x) + P_a^{ref}(x). \tag{3}$$

Based on the mathematical model of the optical type PMT and the above arguments, the distribution of the radiation power in the OSPC is given by

$$P_d^{ref}(x) = P_o \exp(-K_x \cdot x) K_{ref} = I^2 \frac{S}{x^2} \exp(-K_x \cdot x) K_{ref}, \tag{4}$$

$$K_x = \frac{1}{4\pi}.$$

Where K_x – natural losses in the OSPC, X – optical channel length, m.

Somewhat more complicated stays the problem of determining the power of the reflected angle of the light flux $P_{ref}(x)$.

The radiation source is a point, placed on the axis and giving a light beam with a small aperture – in the range of the angle γ (see Fig. 1).

When reflected, the beam is partially absorbed or passes into the indicated object. Any object has a value of reflectance determined by reflection coefficient K_{ref}. Moreover, the reflection coefficient is a function of angle incidence. Practically, the change of K_{ref} is determined by the increase of the angle of reflection (incidence), because of which the beam path length is increased, and this leads to a partial loss of radiated power.

The length of the beam path, which was released at an angle γ to the axis, is equal to:

$$L_b = \frac{(L_c + x_o)}{\cos\gamma} = \frac{X}{\cos\gamma}. \tag{5}$$

Then the change in the reflection coefficient is given by:

$$\rho_{calc} = K_{ref} \cdot \exp\left(-K_x X \int_{\gamma}^{0} \frac{1}{\cos\gamma} d\gamma\right), \tag{6}$$

where ρ_{calc} - the estimated reflectance.

Then, a sufficiently accurate formula for practical calculation determining the power of the reflected angle γ light emission can be obtained from the expression (4) with considering (6).

$$P_a^{ref}(x) = P_d^{ref}(x) \cdot P_{calc}. \tag{7}$$

$$P_a^{ref}(x) = P_o \exp(-K_x X) \cdot K_{ref} \exp\left(-K_x X \int_{\gamma}^{0} \frac{1}{\cos\gamma} d\gamma\right). \tag{8}$$

Then, in the PMT, realizing optical location, the total power of radiation incident on the PD is equal to:

$$P(x) = P_o \exp(-K_x X) \cdot K_{ref} + P_o \exp\left(-K_x X \int_{\gamma}^{0} \frac{1}{\cos\gamma} d\gamma \right) \cdot K_{ref} = \\ = P_o K_{ref} \left[\exp(-K_x X) + \exp\left(-K_x X \int_{\gamma}^{0} \frac{1}{\cos\gamma} d\gamma \right) \right]. \tag{9}$$

Obviously, the main power of the radiation reflected from the object surface is focused on the optical axis of the system.

According to the calculation scheme the angle γ is determined from the expression:

$$tg\gamma = \frac{D/2}{X_o}. \qquad \gamma = arctg\frac{D}{2X_o}, \tag{10}$$

where D - diameter of the lens, m; X_0 - installation position of the lens relative to the radiation source, m.

For the case of optical location the variable X, which characterizes the length of the path of the beam incident on the displayed surface and reflected from it, is set equal to $2L$.

1.2 Analytical calculation of the primary measuring transducer of optical type

The proposed interpretation of the mathematical model of the optical type PMT allows to make the calculation of power distribution radiation $P(x)$ at the photo detector at different values of the reflection coefficients K_{ref} and channel length, as well as to determine the optimum length of the channel in which the reflected optical signal is surely fixed at the receiver of the converter. During the simulation of the PMT by varying generalized methods and variable elements the new technical solutions according to the requirements were searched.

Thus, the expression (9) is converted to the form

$$\frac{P(x)}{P_o} = K_{ref} \left[\exp(-K_x X) + \exp\left(-K_x X \int_{\gamma}^{0} \frac{1}{\cos\gamma} d\gamma \right) \right]. \tag{11}$$

It is the main expression (11) in the development of the method for calculating the converter and can be considered static characteristic equation of the optical type PMT. Static characteristic defines the basic parameters and the converter operation.

Due to the complexity of the integral expressions (11) describing the mathematical model of the optical type PMT, it is expedient to solve it by a numerical method on a computer. Thus, an important issue is to determine the actual values of the reflection coefficients for the materials of different type and thermal condition. A comparison of

the obtained coefficients with the calculated ones will test the adequacy of the proposed model.

Fig. 1.2 shows the dependence of the relative value of the total power $P(x)/P_o$ at the photo detector on the total length of the optical channel X (where P_o - radiation power value at $x = x_o$).

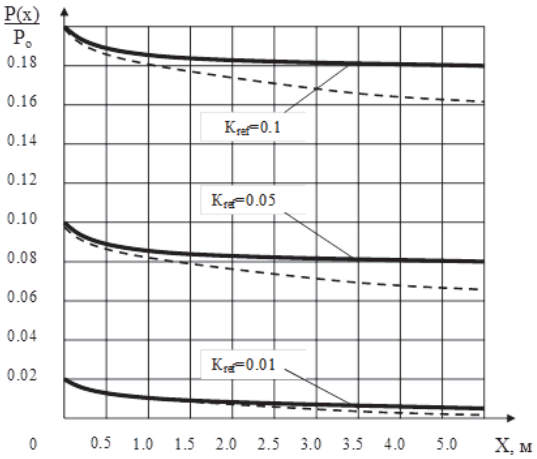


Fig. 2. Static characteristics PMT of the photoelectric type depending on the reflection factor (straight line – received coefficients, dashed line – calculated ones).

Fig. 1.3 shows a diagram of the change in the signal power at the photo detector $P(x)/P_o$ depending on the reflection angle γ, which determines the calculated reflection coefficient ρ_{calc}.

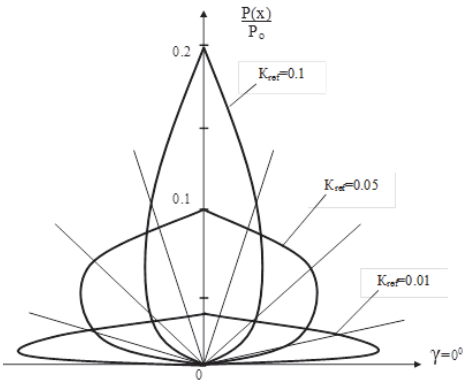


Fig. 3. Power distribution of the reflected signal depending on the viewing angle and K_{ref}

Calculations were carried out on a computer numerical. Data for calculation are taken from the real optical PMT (photoelectric) with the following parameters: K_x = 0.08, X_o = 0.005 *m*, D = 0.005 *m*, K_{ref} = 0.01...0.5.

According to the analytical calculation, the following conclusions were obtained:

1. Based on the above studies of the processes of transmission and reception of optical signals calculated dependences were obtained which describe the variation of the radiation power on the length of optical channel and magnitude of the external attenuating factors.
2. These results allow us to determine the optimum length of the optical channel (in general or for a particular type of object material) for maximum reliability of the control.
3. Theoretical studies have shown that the diagram of the reflected signals has a narrow focus, for the PMT it must be installed exactly perpendicular to the controlled object, that will give an opportunity to get maximum possible power of the reflected and attenuated optical signal on the photodetector.
4. Decreasing the angle of view of a photoelectric transducer to a value greater than five degrees is energetically favorable, since the power of the reflected signals is drastically reduced, due to the specificity of mixed reflection (the main energy is distributed according to the laws of directed reflection, and the rest - the diffusely scattered).
5. Due to the operation of photoelectric transducer as an optical locator the problem of studying the attenuating effect of the optical signal propagation channel, in particular the reflectivity of displayed objects K_{ref} is urgent. This value has the greatest effect on the reflected optical signal, more than 10 times of weakening its energy.
6. The main optical signal power loss is due to low reflectivity of the displayed object. Thus, the main way to increase the channel length is to increase the sensitivity of the photo detector.

As a result, it became necessary to experimentally the reflection coefficients of some materials in order to determine the actual value of attenuation of the signal energy information, and correcting the optimum length of the channel.

2 Experimental studies of the photoelectric transducer and the system based on it

2.1 The purpose and methodology of the experiment

A photoelectric transducer (sensor), as a primary element of an information-measuring system, can be installed at any test object, where the reflectivity of the material varies within wide limits, for example, due to the effect of temperature. Consequently, a necessary requirement for the normal functioning of the control system is the knowledge of the value of the reflection coefficient of the optical signal, both depending on the temperature of the material and its type. The reflectivity is also in-

fluenced by the quality of machining of the controlled object and the angle of the optical-type PMT relative to this object [3,7,8].

The tests of the digital photoelectric transducer prototype to control the presence and position of the object in the control zone were carried out in order to check its operability, to assess the speed and accuracy indicators in the formation of primary measurement information for the information-measuring system.

The methodology of the experiment is as follows.

1. Determining the optimal distance from the PMT to the object, wherein the reflected signal is received surely – determining the channel length.
2. Measurement of the maximum deviation angle of the optical sensor from 90 degrees, wherein the further signal is fixed – determining the angular error control.
3. Knowing the distance from the sensor to the object (see pos. 1) and a maximum beam deflection angle (see pos. 2) the magnitude of the reflected beam reliable reception zone is determined – determining a linear error control.

2.2 Results of experimental studies

Experimental studies of the photoelectric transducer and the system based on it led to the following conclusions.

1. The angular deviation of the line of sight transducer, wherein the reflected signal was recorded still does not exceed 3 degrees.
2. Linear control error does not exceed 2.5% of the total length of the channel.
3. Attenuation of the reflected beam increases by an average of 7% per meter of the channel during the propagation of an optical signal in real conditions of OPC in comparison with the calculated values.
4. With a decrease the reflection coefficient K_{ref}, the power of the reflected signal significantly decreases and acquires a narrow directivity, which makes the requirements for the orientation of the transducer at an angle of 90 degrees to the controlled object more stringent.
5. The transducer is capable of reliably operating at a distance of at least 0.1 m from the monitored object. The maximum distance to the monitored object is determined by its reflectivity and the intensity of optical interference in the location area ($1.1m$ according to the test).
6. The reflection coefficient of the object K_{ref}, at which its optical location is possible, is not less than 0.01.
7. The speed (response time) of the converter is $t_r \leq 1\ ms$.
8. The use of special modular means of obtaining primary information on the basis of photoelectric measuring converters with digital filtering reduces the time for information processing by more than 2 times.
9. The signal-to-noise ratio due to the introduction of additional measures to increase the noise immunity and the organization of a discrete measuring channel was 16.7 (with a theoretical value of 18), which determines a greater degree of attenuation of the optical signal in comparison with the calculated one.

According to Fig. 2, the theoretical dependences reflect the experimental data quite well (shown by the dotted line), which indicates a good convergence of the experimental and analytical data. It is obvious that under the control of objects with reflectance $K_{ref} \geq 0.5$ experimental curve will tend to the theoretical one.

The studies carried out prove that the developed photoelectric digital measuring transducer can be used in information systems for monitoring the parameters of objects of animate and inanimate nature, to monitor their presence and composition in any conditions, for example, the action of high temperatures, intense optical and electromagnetic interference of various nature.

The tests of the converter confirmed the reliable reception of reflected signals at a distance of the control object from the photo detector up to 1.1 *m*, the absence of gaps in the controlled objects and the reliable transmission of primary information about the presence of the object to the registration unit (simulator of the information system). A significant decrease in the length of the optical channel is due to the strong weakening effect of the object surface.

As a result, it was found that when an object appeared in the control zone, the photo detector was triggered only when the reflected beam was deflected by an angle of no more than 1.5 degrees, which corresponded to a linear control error of less than 2% for the specified length of the optical channel.

3 Investigation of noise immunity of a photoelectric converter

It is known that the most appropriate way for monitoring parameters of objects is photoelectric which provides desired noise immunity, reliability and control accuracy. This method is realized in the photoelectric transducers. Control objects based on the detected binary optical signal, and the formation of the control signals are carried out by the controlled object itself – both when working on clearance and optical location [8,9].

However, existing photoelectric converters of analog and pulse type when operating in various conditions are largely exposed to optical interference and electromagnetic perturbations, leading to the distortion of the information signal.

The use of analog information signals in the transducers is hampered by the presence of predominantly continuous disturbances, due to the specificity of the existing equipment. In the single pulse transmitted signals energy is equivalent and in some cases much less than the energy of the medium of their propagation. When such devices operate in extreme conditions the probability of errors of the I-st and II-nd kind – a false alarm P_{fa} and missing signal P_{ms} is significantly increased.

$$P_{fa}(S) = P_S(0) = p\left[\sum_{i=1}^{n} \xi_i \rangle \; X_L\right]; \tag{12}$$

$$P_{ms}(S) = P_0(S) = p\left[\sum_{i=1}^{n} S + \xi_i \langle \; X_L\right], \tag{13}$$

where ξ - active interference; X_L - limit signal level.

When low-power infrared LEDs are used as emitters, the reception should be carried out with the accumulation, when one low-power signal should be repeated several times, and the results obtained at the receiver, must be summed up in a certain way. Such method can significantly raise the signal-to-noise ratio. Since the deterministic signal is summed arithmetically, its amplitude increases n times (n – is the number of pulses in the signal, it can be a few hundreds), and the energy – increases n^2 times. In this case, the fluctuating interference is added in terms of power. Thus, when the control objects in the optical location mode are based on the synchronous accumulation of "package" of pulses the signal-to-noise ratio increases significantly (by more than 10 times of magnitude), which improves the accuracy of the measurement information:

$$\sum_{i=1}^{Tp} \int_{Tp}^{Tp+trec} \left\{ \left[S(t) + \xi(t) \right] \times h_f(t) \right\} dt \neq X_L, \tag{14}$$

where $h_f(t)$ - filter response to signal; T_p - pulse time; t_{rec} - time of the reception.

However, a large number of data pulses in a "package" (several hundred) significantly reduces the speed of the control system. Thus, there is a need to develop an optimal structure of the information optical signal that satisfies the certain energy requirements, conditions of speed and noise immunity.

The results of the study of industrial and other interference made it possible to formulate the energy and frequency requirements for the information optical signals. At the same time information coding can significantly increase the noise immunity of the primary transducer.

Noiseproof photoelectric converters (transmitters) typically operate with an active process control and special emitters, which creates an optical signal with properties significantly different from those of optical interference.

The analysis made it possible to identify the main trends in the development of transducers, as well as to offer new approaches to increase the credibility and reliability of the primary information.

The methods of receiving single signals and the pulse sequence, that increase the noise immunity of the transmitter were researched and summarized. To receive the random sequence of discrete transmitter signals (to solve the problem of detection) different receivers can be used, while the number of signal reception techniques is so large that a special systematization was required to study them. The main task of considering reception techniques is to determine the performance of the receiver, which display the dependence of the error detection probability on the energy relations in the communication channel, and other parameters that affect this probability:

$$\alpha = F(P_{er}) \tag{15}$$

$$h_0^2 = f(\alpha) \tag{16}$$

where $h_0^2 = E/N_0$ – the signal-to-noise ratio energy, E – signals energy, N_0 – two-sided interference spectral density; α – the argument of the integral probability $\Phi(\alpha)$.

A method is proposed for the probabilistic assessment of the reliability of receiving code-pulse optical signals, based on a joint analysis of the operating characteristics of the receiver and the energy parameters of the transmitted signals.

According to the proposed method, the desired probability of receiving the code-pulse signal is set – P_{er}. The argument of the probability integral $\Phi(\alpha)$ is determined from the dependence $\alpha = f(P_{er})$ (see Fig. 4).

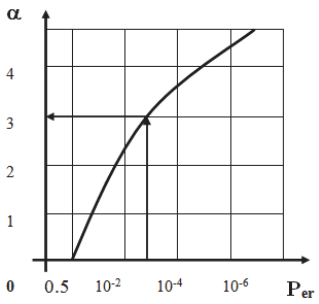


Fig. 4. Dependences of the argument of the probability integral on the reception error $\alpha=F(P_{er})$.

From the set value α, according to the dependence $h_0^2 = f(\alpha)$ (see Fig. 5) the minimum value of the signal-to-noise ratio energy h_0^2 is determined.

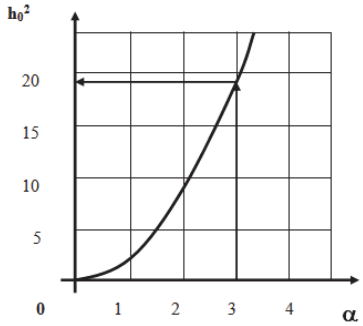


Fig. 5. Value of the signal-to-noise energy ratio depeding on the obtained argument α of the probability integral $h_0^2 = f(\alpha)$.

On the other hand, an analytical method is proposed for determining the optimal weight w of the code sequence by the energy indicators of the operating interference and transmitted signals [9]. The ratio of the signal output power to its power at the receiver input is determined by the 2nd degree of the impulse response of the sequence.

$$\frac{P_{m.out}}{P_{m.in}} = \frac{\sigma_{out}^2}{\sigma_{in}^2} = \sum_{n=0}^{b} h^2(nT) = w^2. \tag{17}$$

The greater the weight codeword w, the greater the ratio $P_{m.out} / P_{m.in}$ and hence the signal-to-noise ratio.

It is obvious that $w^2 \approx h_0^2$. For signals with a passive pause, the value h_0 characterizes the minimum energy consumption for transmitting a unit of information with a predetermined error probability P_{er}. And value w, in turn, determines the minimum weight value of pulse code sequence the energy indicators of which are equivalent to h_0.

Then, the accuracy of the correct reception of the code sequence is determined by its weight w (quantity of single pulses) for a given error probability (see Fig. 6).

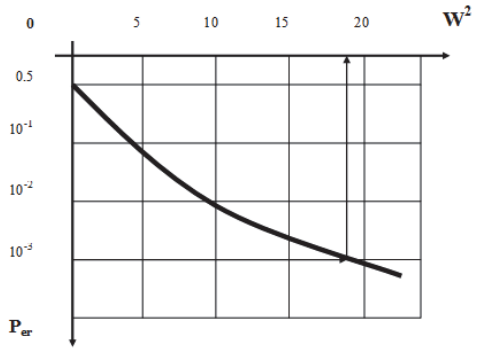


Fig. 6. Receiver performance $P_{er} = f(w^2)$ for passive paused signals.

An increase in the codeword weight improves the output signal energy by a factor of w^2, fluctuating interference is added in power, which leads to an increase in signal-to-noise ratio only by a factor of w. It should be noted that the atomicity of code b (its total length) does not directly affect the noise immunity of the code, but only determines the sampling step for any coding method:

$$max[e] < 2^{-b-1} = T/2 , \tag{18}$$

where e – error; T - sampling step.

It is analytically proved that the optimal value of the weight codeword is $w \geq 5$.

The developed method makes it possible to determine the optimal weight w of the code sequence for a given value of the probability of an error, without referring to the energy calculations of the acting interference and transmitted signals.

Consider an example for devices with one transmitter and one receiver. In this case element doubling code is the most effective and maximally resistant to interference. It is characterized by the introduction of additional symbols for each symbol of the combination of information part, with the one is complemented by zero and zero – by one. Then the initial combination 11111 is converted into a kind of 1010101010. The code allows you to detect all errors, except when there are two errors in the paired elements (combination of type 00 or 11) – a binary error. The noise immunity of the code is high due to increased redundancy and the use of a stable structure of the

source code, where there is no interleaving of elements. In turn, this allows you to simplify the encoding and decoding process in the information processing system.

Methods of digital filtering of the optical signal in such photoelectric transducers have also been investigated. When using digital filters, a high accuracy of information conversion is achieved, and the speed of the control system is increased. Based on the structure of the proposed code-pulse information signal *x(nT)*, a non-recursive digital filter is developed, the equation of which is

$$y(nT) = \sum_{k=0}^{9} b_k x(nT - kT) = b_9 x((n-9)T) + b_7 x((n-7)T) + \\ + b_5 x((n-5)T) + b_3 x((n-3)T) + b_1 x((n-1)T) \quad , \tag{19}$$

where b_i – filter coefficient matrix; $b_k = x(nT) = \{0,1,0,1,0,1,0.1,0,1\}$; k – code length.

At the information level, the problem of detecting the object of control is solved using the theory of statistical decisions (statistical hypothesis testing method *H*) [9]. The receiving device based on the received realization of *y(t)* determines which of the signals it contains (S_1 or S_2), and hence which element of the digital sequence was transmitted (X_1 or X_2) – a two-alternative solution. For a channel with constant parameters and additive noise *n(t)* according to the implementation $y(t)=S_i(t)+n(t)$, where $0<t<\tau_0$ the statistical hypotheses H_i are checked.

$$H_1: X_1 = 1 - "yes", \tag{20}$$

$$H_2: X_2 = 0 - "no". \tag{21}$$

Because of the presence of interference $S_i(t)$ can take either the value of the signal $S_1(t)$ with a probability $p(S_1)$, or the value of $S_2(t)$ with a probability $p(S_2)$. The probability of errors when receiving exactly known binary signals:

$$P_{er} = 0{,}5\left\{1 - \Phi\left[\sqrt{\frac{E(1-\chi)}{4N_0}}\right]\right\}, \tag{22}$$

where $\Phi(\chi)$ – the correlation integral,

$$\Phi(\chi) = \frac{1}{\pi}\int_o^x e^{-\xi^2/2}\, d\xi; \tag{23}$$

χ – cross-correlation coefficient between the signals $S_1(t)$ and $S_2(t)$; ξ – acting hindrance.

The choice of a hypothesis is carried out on the basis of the value Δ, which characterizes the likelihood of a particular hypothesis – the likelihood ratio:

$$\Delta = \frac{p(S_1)}{p(S_2)} \exp\left\{\frac{1}{N_0}\int_0^{\tau_0}\left[y(t) - S_2(t)\right]^2 dt - \frac{1}{N_0}\int_0^{\tau_0}\left[y(t) - S_1(t)\right]^2 dt\right\}. \tag{24}$$

Expression Δ fundamentally solves the problem of optimal discrimination between two signals according to the criterion of the minimum probability of an error, since in this expression all known and unknown quantities are functionally connected. Therefore, knowing Δ and acting in accordance with the Bayesian rule, one can decide which of the two possible signals (S_1 or S_2) was transmitted, and therefore, which of the message elements x_1 or x_2 must be reproduced at the channel output: $\Delta>1$ – the signal $S_1(t)$ is received; $\Delta<1$ – signal $S_2(t)$ is received.

On the basis of this, the criterion of optimality of the converter structure was adopted – high noise immunity, which ensures reliability and accuracy of control in extreme conditions (criterion of minimum error).

The existing designs of photoelectric transducers, in spite of their simplicity and reliability, do not fully satisfy the ever-increasing requirements for the accuracy and reliability of the information obtained about the object or signal propagation medium. The use of a discrete information channel makes it possible to reduce hardware costs, as well as reduce errors in signal conversion in the electronic path. Discretization of messages allows you to get rid of random analog (low-frequency) noise, some of the impulse ones that do not fall into the frequency bandwidth, therefore, significantly reduce the total accumulated error during reception.

The structure of an optical signal receiver, optimal according to the proposed criterion, has been developed.

As a result of an integrated approach in the development of the structure of the measuring transducer, a mathematical model of a discrete photoelectric converter with high noise immunity and speed was developed (see Fig. 7), which reflects the essence of the processes of converting an information signal during its passage in the electronic path of the converter [9,10].

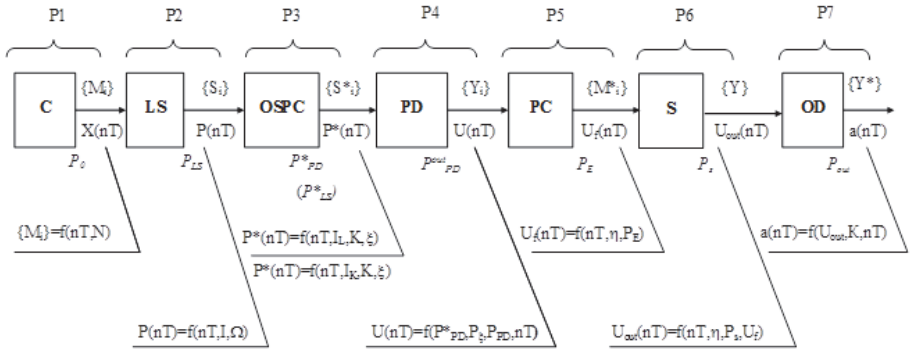


Fig. 7. The mathematical model of discrete measurement channel. The following devices are indicated: C - encoder, LS – light source (emitter), OSPC - optical signal propagation channel, PD – photodetector, PC - information signal processing circuit, S - solver, OD - output device, *Pi* - operators for converting optical information signal in the measuring channel.

Such a model of a discrete information channel reflects the structure of an optimal photoelectric measuring transducer with digital filtering of a pulse-code optical signal.

Timing diagrams shown in Fig. 8, explain the operation algorithm of the pulse code receiver.

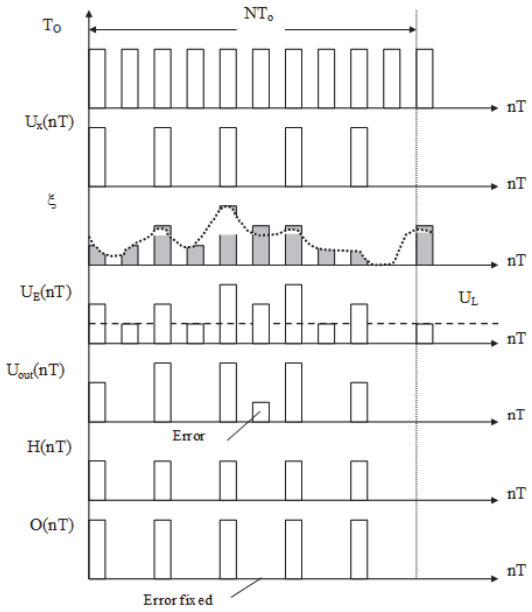


Fig. 8. Timing diagrams explaining the algorithm of the receiver of pulse code signals. Fig. 8 denotes: $U_x(nT)$ – code sequence $\{x\}$; ζ – the current interference among the optical signal; $U_E(nT)$ – voltage at the input of the PD; $V_{out}(nT)$ – voltage at the output of the PD, $H(nT)$ – the impulse response of the filter, $O(nT)$ –message generation for the output device (display), U_L – threshold signal, T_O - discreteness period.

The use of digital filtering in the photoelectric devices requires a fundamentally new structural and algorithmic information processing techniques and methods of evaluation of reliability of the control. Acquisition of new approaches for the detection and location of the optical signal, as well as optimization of existing photoelectric converters has identified a new trend in optical location – the transition to digital information processing. This increases the accuracy of the operation and structure of the transducer is optimized, the inverter provides interaction with the information-measuring system, which will allow to realize more accurate signal processing algorithm and increase the speed of the control system, and make it possible to interact promptly with the technological process.

Digital photoelectric PMT has been developed and experimentally investigated as a part of an information measurement system (IMS) for monitoring environmental parameters.

Experimental studies have confirmed the adequacy of the converter analyzes to the real processes occurring in photoelectric PMT and compliance with the actual expected results. The technical effect was manifested in increasing the reliability of

monitoring objects while increasing the speed. This allows you to increase the productivity and efficiency of IMS in general.

As a result, the conducted studies made it possible to obtain scientifically substantiated theoretical and experimental results, which, when applied together, significantly increase the quality of the functioning of the information-measuring system (IMS) due to a significant increase in performance and reliability of control. At the same time, the introduction of such photoelectric measuring transducers in various industries will reduce the cost of developing a wide range of specialized devices for monitoring and measuring individual parameters, increase the quality and reliability of measurement information, and simplify the construction of object control systems and parameter control.

4 Development of modular structures for photoelectric measuring transducers

When designing photoelectric measuring transducers as a means of primary control of object parameters, attention should be paid to the tasks performed by a specific converter, the possibility of integrating it into a local information control system, as well as the possibility of combining control functions. Photoelectric primary measuring transducers (PPMT) must be easy to operate, reliable, have a unified set of blocks (repairable) [10].

The considered PPMT structures can be used in measuring systems to control, for example, displacements, speed, temperature, as well as air, liquid or solid pollution due to the photocell's response to a change in the light flux (direct or reflected) [11, 12,13].

The assessment of the composition of the air environment is carried out by transducers that work as an optical locator. It is used in medicine to assess the efficiency of the respiratory process or to determine the pulse rate. Such measurements are carried out by photometric measurement of the percentage of oxyhemoglobin in peripheral arterial blood, which is possible due to the difference in the spectral characteristics of light absorption by reduced hemoglobin and oxyhemoglobin. It is known that at a light wavelength of 620 ... 680 μm the absorption coefficient for hemoglobin is several times higher than for oxyhemoglobin [14,15].

Sensors for such measurements are made in the form of a clip and are put on the earlobe in such a way that on one side there is an emitter, and on the other side there is a photodetector. A change in the light flux on the photodetector will indicate a change in the quality of the blood [16].

On the other hand, a change in the luminous flux can be caused by a change in the degree of absorption of light due to a change in the thickness of the tissue when it is filled with blood. Since the blood supply changes in time with the contractions of the heart, the frequency of the change in the light flux can be judged on the pulse rate. Pulse sensors are usually mounted on the earlobe or on the nail phalanx of the fingers.

The commercially available photo-electro-colorimeter is used to measure in separate sections of the wavelength range 315 ... 980 *nm*, emitted by light filters, the

transmittance T and optical density D of liquid solutions and solids. Such device also serves to determine the concentration of substances in solutions by constructing calibration graphs. The colorimeter is used in water supply enterprises, in the metallurgical, chemical and food industries, in agriculture, in medicine and other areas of the national economy.

The principle of operation of such devices is based on the absorption of light as it passes through a substance [17]. According to the Bouguer - Lambert - Beer law, the intensity of the light flux Φ passing through a layer of matter with a thickness d and the intensity of Φ_0 incident on it are related by the relation

$$I(l) = I_0 e^{-k_\lambda l} \tag{25}$$

where $I(l)$ is the intensity of light that has passed through a layer of the substance with a thickness, I_0 is the intensity of light at the entrance to the substance, k_λ is the absorption index for the wavelength λ.

The absorption index is determined by the properties of the substance and in general case depends on the wavelength λ of the absorbed light. This dependence is called the absorption spectrum of the substance.

The principle of measuring the transmittance consists of the fact that the light fluxes of full Φ_0 and Φ passed through the medium under investigation are alternately directed to the photodetector and the ratio of these flows is determined - an analogue of the differential type PPMT.

Let us make a brief description of a number of modular structures of PPMT based on pulse-code modulation of an optical signal, proposed as a means of primary control of parameters of objects with extreme external conditions.

Fig. 9 shows the structure of a code combined PPMT (optical locator). In Fig. 9 denotes: Ci – code generators; LS – light source (emitter); SPC – signal propagation medium (channel); PD – photodetector; CS – comparison scheme; RC – reference code generator; *{i}* – signal conversion; Y – output signal. The principle of operation schemes is sufficiently simple and clear from the following schemes: the generated code sequence is transmitted to the photodetector, after which the received code sequence is restored in the control circuit and is compared with the reference code value.

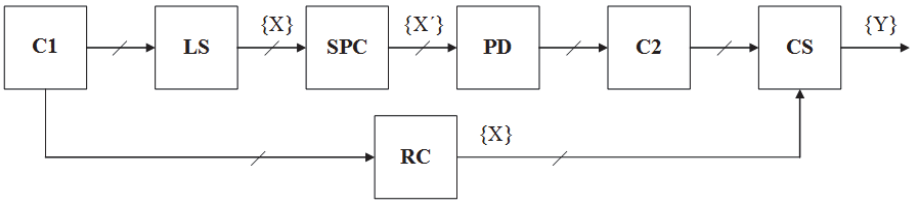


Fig. 9. Block diagram of the code combined PPMT.

Fig. 10 shows the structure of a code distributed PPMT. In Fig. 10 denotes: C – code generators; LS – light source (emitter); SPC – signal propagation medium (channel); PD – photodetector; MCS – main control scheme; Δt – sampler; DCS –

digital comparison scheme; RC – reference code generator; *{i}* – signal conversion; Y – output signal.

Code distributed PPMT is more preferable due to mechanical and electrical isolation of the transmitter and receiver. However, such a transducer is rather complicated to implement due to the difficulties of synchronizing the transmission and reception signals.

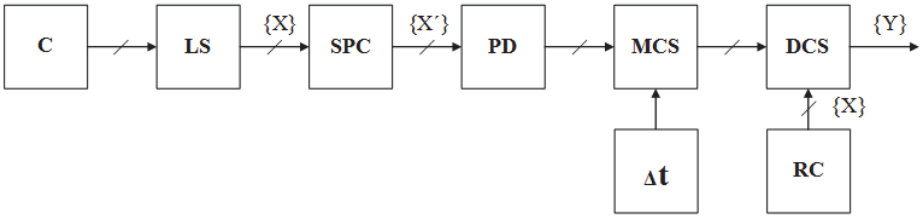


Fig. 10. Block diagram of the code distributed PPMT.

The simplest example of a self-tuning PPMT can be seen in the standard system of the remote control of home audio and video. The receiver implements a simple digital filter with time and frequency selection of the information optical signal, which gives a sufficiently high level of noise immunity. In this case the receiver operating frequency depends on the frequency of the transmitter and is configured by transmitted code.

This type of PPMT is possible to allocate in a particular way due to the fact that it has a standard element base and provides the possibility of revising the electrical circuit and adapting to specific conditions of use. In this case, the electronic transmission path is not associated with the electronic reception path.

Of particular interest is the compensation PPMT based on digital filtering of the received code. Complete information on the principle of operation of such information devices is available in the works [9,10].

Analyzing the circuit of the developed compensation converter, the block diagram of which is shown in Fig. 11, we can conclude that it allows you to determine the mismatch of the transmitted and received signal, as well as restore the received code.

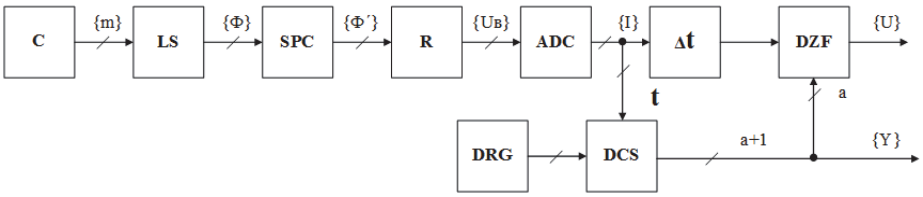


Fig. 11. Block diagram of the compensation PPMT.

In Fig. 11 denotes: C – code generators; LS – light source (emitter); SPC – signal propagation channel; R – receiver; ADC – analog-to-digital converter; *Δt* – sampler; *t*

– time delay; DZF – dynamic zeroing digital filter; DCS – digital comparison scheme; DRG – digital reference generator; *{i}* – signal conversion; Y – output signal.

The proposed structure of the PPMT can be used in cases where it is necessary to determine the magnitude of interference (mismatch) acting in the signal propagation environment and affecting the transmitted code, when making a logical conclusion about the correctness of the control – "yes" or "no". The received code is compared by the comparator with the reference subtraction path, and the difference is fed to the dynamic zeroing filter, where this difference is subtracted from the received sequence. As a result of zeroing, a restored sequence of signals is obtained, and the difference is a reception error, which is used to determine the degree of "pollution" of the environment.

A universal tracking type PPMT is proposed, the block diagram of which is shown in Fig. 12. In Fig. 12 denotes: T – clock generator; C – code generator; LS – light source (emitter); OSPC – optical signal propagation medium (channel); PD – photodetector; CO – control object; DCS – digital comparison scheme; CC – control circuit; Co – counter; Reg - code storage register; *{i}* – signal conversion; Y – output signal. It is designed to solve complex problems of calculating the length of the controlled product, its speed (with known dimensions) of movement, object detection, as well as determining the number of products and is intended for adaptive control systems.

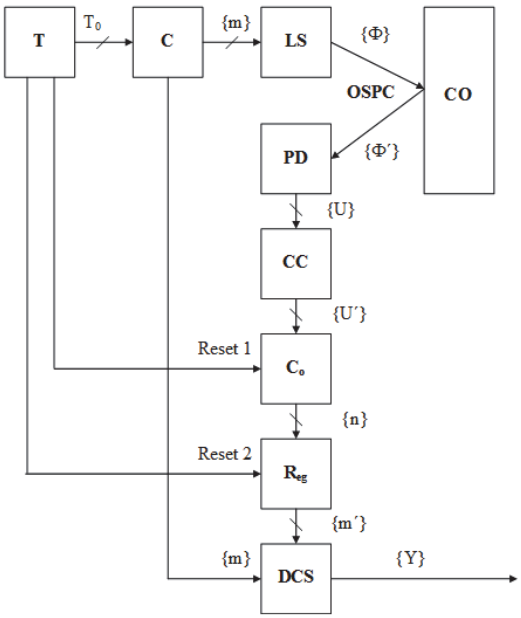


Fig. 12. Block diagram of the universal tracking type PPMT.

Such a PPMT works equally reliably both with a set of pulses and with code sequences, which is achieved due to the original circuit design. The generated code is

simultaneously fed to the transmitter and to the digital comparator of the receiver, where these codes, being compared, show the degree of signal distortion in the optical signal propagation channel (OSPC). Depending on the amount of code distortion, a message about the result of the control is issued. Digital filtering of the received code is carried out by time selection of the code sequence and at the same time by controlling the «weight» of the code combination.

As a result, the general classification of PPMT by the type of transmitted signal is shown in Fig. 13.

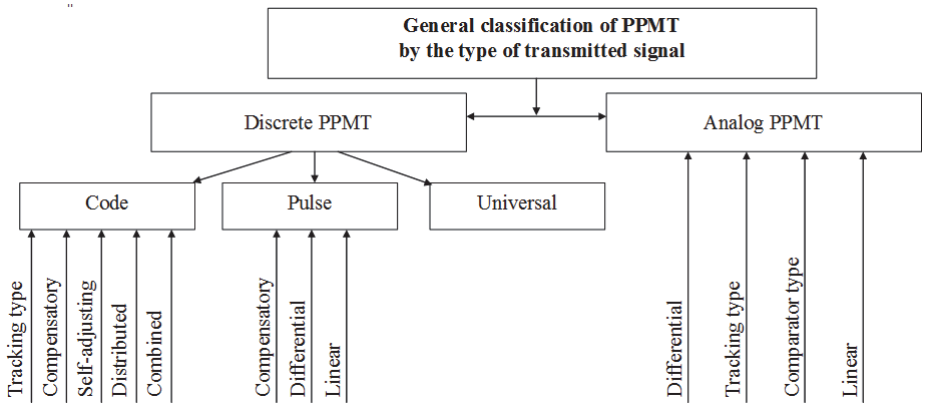


Fig. 13. The general classification of PPMT by the type of transmitted signal.

Analog and discrete PPMT were studied in detail in [9,10] and do not require additional consideration.

Conclusion

The most promising direction in the creation of PPMT as a means of primary control is the design of control systems for objects using several control channels, when one central transmitter sends control signals to several spatially separated receivers. It is possible to transmit messages only over one channel, then the problem of separating the transmission-reception channels arises. There are three possible ways of dividing information channels for PPMT with pulse code modulation of an optical signal.

Time division of control channels is implemented in the following way. From one transmitter, commands or messages are sent alternately to each receiver in a strict sequence and with a regulated transmission time. This method is typical both for spatially distributed channels and for single-channel reception and transmission of several messages. The only and significant disadvantage of such channel separation is the low speed, which prevents the build-up of channels, reduces the overall efficiency of the information-measuring system with the PPMT.

The level separation of channels consists in varying the power of the transmitted signal when transmitting signals to different receivers or when transmitting different

messages. This method is ineffective due to the complexity, and sometimes the impossibility of changing the power of the optical signal, for example, when the emitter operates in a pulsed mode.

Pulse number separation is most expedient in application. With this method of dividing channels, each of them is encoded in a certain way, which excludes the possibility of erroneous channel selection. Also, the performance (processing speed) does not decrease with any number of channels. This method is easy to implement in hardware and without unnecessary design time.

In the proposed primary transducers of the photoelectric type, with a slight modification of the sources and receivers of code-pulse optical signals, the possibility of number-pulse (code) channel separation is provided. This makes it possible to create a multichannel control system of the object, provided that the measuring transducers are conveniently located at the object of control.

As a result of a comprehensive analysis of all available types of discrete PPMT with different kinds of optical information signals, their structural organization, the specifics of their functioning and the tasks being solved, three main groups of PPMT can be distinguished (see Fig. 13) that have a specific field of application.

References

1. Talanchuk, P., Skripnik, Y., Dubrovny. V.: Measuring instruments in automatic and control systems: A textbook for university students who are trained in the specialty "Automation of technological processes and production". Raduga, Kiev (1994).
2. Freiden J.: Modern sensors. Handbook. Technosphere, Moskow (2005).
3. Butsenko, A., Yakovenko, V., Sagaida, I., Lutsik, J.: Sensors and methods for improving their accuracy: tutorial. New in science and technology – for students and pupils. Issue 4. High school, Kiev (1989).
4. Goldstein, A.: Physical basis for obtaining information: tutorial. TPU Publishing House, Tomsk (2010).
5. Miroshnikov, M.: Theoretical foundations of optoelectronic converters (OEC). Mechanical engineering, Leningrad (1977).
6. Azimov, R., Shipulin, Y.: Optoelectronic transducers of large displacement based on hollow fibers. Automation library. Issue 664. Energoatomizdat, Moscow (1987).
7. Borysewicz, J.: Requirements, test methods and evaluation criteria of optical signaling devices (OSD). In: Borysewicz, J., Stępień, P., Chołuj, L. Technique and technology, vol. 44, issue 4, pp. 159-164. BiTP (2016).
8. Goldstein, A.: Physical foundations of measuring transformations: tutorial. TPU Publishing House, Tomsk (2008).
9. Subotin, O.: Increasing the reliability of primary information in the systems of integrated automation of rolling mills. Material pressure processing 3(24), 205-211 (2010).
10. Subotin, O.: Development of a number of modular structures of photoelectric measuring converters. Tool reliability and optimization of technological systems 42, 80-86 (2018).
11. Vasiliev A.V., Grinenko E.V., Shchukin A.O., Fedulina T.G.: Infrared spectroscopy of organic and natural compounds: Textbook. SPbGLTA, SPb (2007).
12. Lucassen G.W., Caspers P.J., Puppels G.J.: Infrared and Raman spectroscopy of human skin in vivo II Handbook of Opt. Biomed. Diagnostics. V. PM107, Chap. 14, SPIE Press, Bellingham (2002).

13. Rosenfeld L.G., Wenger E.F., Kollyukh A.G.: Matrix semiconductor photodetector of infrared radiation and its application in biotechnology. Electronics and communications. Biomedical devices and systems 2, 27-29 (2007).
14. Mamilov S.O., Usman S.S., Stelmakh N.V., Veligotskiy D.V.: CO-OXIMETER. Biological and medical attachments and systems. X International Scientific and Technical Conference “Physical Processes and Fields of Technical and Biological Objects” 13, 94-95 (2010).
15. Veligotskiy D.V., Stelmakh N.V., Mamilov S.O., Osman S.S.: Mathematical modeling of non-invasive diagnostics of carboxyhemoglobin in blood streams. Science newsletter of the Kremenchutsk University of Economics, Information Technologies and Management "New Technologies" 2 (32), 55 (2011).
16. Chemist's Handbook 21. Chemistry and Chemical Technology, https://chem21.info/info/8530/, last accessed 2021/01/05.
17. Encyclopedia of Physics and Technology, http://femto.com.ua/articles/part_1/0388.html, last accessed 2021/01/05.

Implantable biotechnical systems for orthopedics and dentistry

Serhii Tymchyk[1][0000-0003-2977-1602], Elena Sorochan[2][0000-0003-3686-7856], Oleksandr Hrushko[3][0000-0001-5551-375X], Inna Vishtak[4][0000-0001-5646-4996]

[1,3,4]Vinnytsia National Technical University, Khmelnytsky highway 95, Vinnytsya, 21021, Ukraine
[2]Priazovskyi State Technical University, vul. Universytets'ka, 7, Mariupol, 87555, Ukraine
[1]tymchyk@vntu.edu.ua
[2]sorochanen777@gmail.com
[3]grushko1alex@gmail.com
[4]innavish322@gmail.com

Abstract. This section is devoted to development and design of implantable biotechnological systems (IBTS). Biotechnical levels, characteristics and requirements that modern IBTS must meet are determined. The choice of standards and ranges for medical radiofrequency communication is justified, which made it possible to determine implant modules for the basic set of medical radio frequency telemetry. The design features of IBTS for orthopedics (for bone osteosynthesis) and dentistry (intelligent dental implants) are formulated, the basic configurations of IBTS, implant modules and software of IBTS are considered. An integrated classification of means of dental implantation is proposed. It is shown that the level of biotechnicality and state of interfaces of biotechnological systems are objective criteria for the quality of restoration of lost organs and human functional capabilities.

Keywords: osteosynthesis, implant, dental implantology, biotechnical system, biotelemetried module.

1 Introduction

Ensuring the rehabilitation and restoration of the functional capabilities of the biological object is an important task aimed at improving the standard and living conditions, the ability to fully perform their duties, adaptation to living conditions in society. Despite the fact that the damage to the body can be both external and internal, the materials used for rehabilitation are subject to a number of conditions that they must meet.

In the last years, there has been an active trend towards the development of materials aimed at creating artificial tissues that replace damaged skin, muscle tissue, blood vessels, nerve fibers, and bone tissue. Such materials are called biomaterials. This work is successfully applied in the treatment of pathology of the musculoskeletal system, which was facilitated by the development of the industry of endoprosthetics of large

joints. However, the use of biomaterials has a wider application in various areas of human health.

Biomaterials that are used as implants, replace a portion of the bone or as temporary fixators for a broken bone (plate bones, intramedullary rods): biotolerant materials, bioinert materials, bioactive materials [1].

In modern traumatology, there are two fundamentally different approaches to solving the problem of treating injuries and diseases: 1) simple replacement of the damaged area of the bone with an implant, up to the creation of a huge bioengineering structure, replaces the bone and adjacent with it joints, or 2) creating conditions for the regeneration (restoration) of the bone in the damaged area with an implant.

2 Osteosynthesis: types, requirements, application

To osteosynthesis is an operative connection of bone fragments in bone fractures in order to create and obtain optimal positive conditions for the fusion of fragments, the implementation of the process of bone regeneration. Osteosynthesis, as a rule, is used in the treatment of fresh fractures or those that have not healed after treatment with traditional methods (for example, with a plaster cast) [2]. The purpose of osteosynthesis is the creation and further provision of reliable mutual fixation of the matched fragments of the damaged bone, provision of conditions for their further fusion and restoration of the integrity and normal functioning of the bone.

The main prerequisite for osteosynthesis in the treatment of fractures is sufficient reduction, reliable fixation of bone fragments.

Currently, such a distribution of the main types of osteosynthesis is generally accepted:

1) submersible (intraosseous or intramedullary);
2) bone;
3) transosseous (compression-distraction).

Submersible (or internal) osteosynthesis is usually called the method of joining bone fragments by their surgical treatment by fixing them using various structures made of metal, plastic, ceramics, bone grafts, synthetic materials, etc. Today, structures made of special materials and alloys (stainless steel or titanium) are widely used for internal osteosynthesis. In intramedullary osteosynthesis, the fixing structure is inserted into the medullary cavity of the damaged bone. The osteosynthesis type of osteosynthesis carries out mutual fixation of bone fragments by fixing the fixing structure on the bone surface.

Internal osteosynthesis is divided into conventional and compression. Using conventional osteosynthesis (static) with immobile fragments, they are not compressed together at the fracture site. Within compression (or dynamic) osteosynthesis, compression is carried out with a certain force of the fixing structure (for this, screws, bolts, compression plates are used). Sometimes special devices are used for this purpose called contractors. The use of a static or dynamic type of osteosynthesis is determined by the nature of the damage and the choice of treatment technology by a traumatologist.

Compression of broken bone fragments due to the influence of tonic muscle contraction and axial load on the limb is called physiological compression.

Constructions of metal osteosynthesis are made of stainless materials and alloys, the main requirement for which is bioinertness (the ability not to cause harmful reactions (metalosis, necrosis, inflammation, etc.) when interacting with tissues that are nearby). Poor quality metal, heterogeneous in composition, with harmful impurities as a result of corrosion, reduces the strength of the structure and is the cause of various complications (inflammation, suppuration, resorption of bone tissue, necrosis, metalosis, delayed consolidation, etc.). The main technical means of osteosynthesis are: metal and polymer screws, pins, knitting needles, bone onlays, plates, wires, staples, ceramic elements or fixators, which dissolve on their own.

One of the most widespread and accessible types of osteosynthesis in our time is bone osteosynthesis, which does not require complex and expensive surgical equipment and apparatus. Osteosynthesis is one of the methods used for the treatment of nonunion fractures, pseudoarthrosis and long bone defects. This type of osteosynthesis allows you to combine the period of fusion of fragments of the damaged bone, the period of restoration of the functions of the operated limb, is one of its significant advantages.

Exposed osteosynthesis, like any other type of fixation of damaged bone fragments, also has some advantages and disadvantages. The main advantage of this type of osteosynthesis continues to be its availability, low cost, relative simplicity, the possibility of performing a traumatologist of not the highest qualifications, and the absence of expensive X-ray television equipment. However, there is frequent destruction of the extraal fixators, thinning of the cortical substance of the bone in the place where the extraal plate is installed. As a result of the latter reason, repeated bone fractures often occur.

In our time, polymeric materials cannot effectively compete with metal structures in their physical and mechanical properties. However, the rapid pace of development of chemistry, the development of new polymeric materials, the emergence of new high-molecular compounds with unexpected, previously expected properties, the production of substances with wide capabilities of various compounds and compounds are causing a new impetus to the interest of scientists and engineers in the study implementation of these promising materials in medical practice. A large number of polymeric materials are characterized by good chemical resistance in various aggressive environments. These materials do not cause biological rejection, metabolic disorders, metabolism, retain their physical and mechanical properties for a fairly long time[3,4].

3 Simulation of load on bones.

The bone model was fixed in the clamps of the proximal end, the weight suspension was attached to the distal end of the bone, after which two dial indicators were installed. The indicators were fixed in two mutually perpendicular planes at the distal end of the model, then it was loaded with weights from 1 to 4 kg (10-40 N). The loading in all cases was carried out at equal distances from the place of pinching of the bone model. The measurement of the deflection was carried out in 4 mutually perpendicular planes

(dorsoventral, ventro-dorsal, lateromedial, mediolateral). Parallel to this, deformations were observed in the horizontal plane, the magnitude of which was only a small fraction (hundredths of a percent) of the deflection in the vertical plane. The load was carried out to the limit when residual deformations can occur, at which Hooke's law no longer holds.

Researches have shown that deflection occurs in both planes: vertical and horizontal. From this it follows that the fold will not be flat, but oblique. This phenomenon is explained by the shape of the bone itself, as well as by a change in its section along its length.

To determine the stress-strain state of the femur during oblique bending, a study was carried out on wooden models that meet all the geometric parameters of a natural bone. In fig. 1 is a photograph of a wood femur model.

Fig. 1. Life-size wooden model.

Primarily, the study focused on bending bone models. As a result, it was found that deformations occur not only in one plane, but simultaneously in two, that is, the phenomenon of oblique bending is observed - one of the types of complex loads [5].

Table 1. Average data of the dependence of the deflections of bone models in the dorsoventral plane under bending load.

Bending plane, deflection value (mm))								
№	P(kg)	M (Н·м)	dorsoventral					
			$f_В$	M±m	$f_Г$	M±m n=10	$f_Σ$	M±m n=10
1	0	0	0	0	0	0	0	0
2	1	6	0,678	0,03	0,029	0,001	0,679	0,03
3	2	12	1,426	0,07	0,062	0,003	1,43	0,07
4	3	18	2,567	0,09	0,144	0,007	2,58	0,11
5	4	24	3,661	0,12	0,21	0,01	3,68	0,15

4 Application of biotechnical systems in osteosynthesis.

A biotechnical system (BTS) (Fig. 2) consists of several autonomous modules, a multifunctional database, a microcomputer, software and hardware-software interfaces, means of recording, storing and displaying data.

The difference between the developed BTS is the presence of an implanted biotelemetric module in its structure (Fig. 3). The implanted module, structurally made in the

form of a separate functionally completed structure with an autonomous power source, is fixed directly on the metal plate of the retainer using a mechanical (screw) connection or glue.

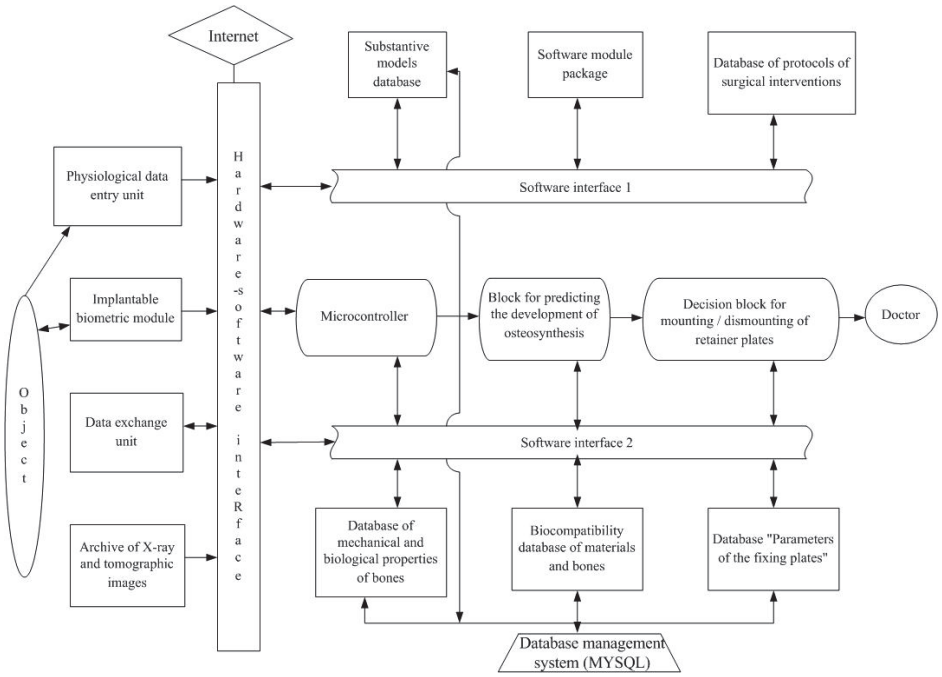


Fig. 2. Block diagram of a biotechnical system for monitoring object parameters.

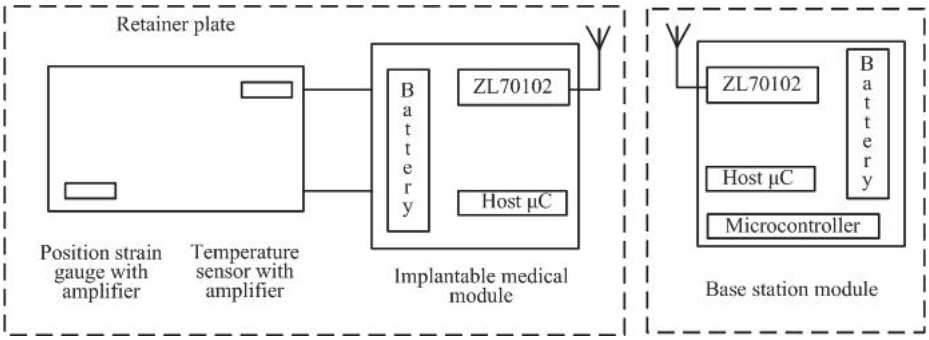


Fig. 3. Block diagram of the implanted biotelemetried module (IBM).

The functioning of the implanted module has some characteristic features. The main one is its functioning in the near inductive radiation zone under specific conditions of an internally organized environment. It defines the basic requirements for the methods of transferring information and the implanted structure. First of all, this concerns the

choice of the radio range, since the use of the low-frequency range of radio waves has significant difficulties for the technical implementation of the equipment.

However, this range also has a limiting factor for its increase, due to the fact that radiation with a frequency of more than 100 MHz negatively affects the human body.

The developed implanted biotelemetried module meets the basic requirements and criteria for devices of this class [6] and has corresponding characteristics:

- frequency range, MHz – 87,5-98,5;
- power consumption, mW– 2;
- weight, g – 14;
- dimensions, length, width, height, cm– 3,2*1,8*0,8;
- compatibility of radio signal parameters with a biological organism – ensured;
- compatibility of the material of the IBM case with the biological tissue of the body - ensured;
- remote control of IBM operation – yes;
- signal compatibility with the communication channel - ensured.

The choice of microcontroller 1 was carried out by comparing the parameters and characteristics of microcontrollers from 5 companies: Atmel, Silicon Labs, Microchip and Texas Instruments, Analog Devaices. As a result, the PIC12F683 microcontroller from Microchip was selected with the following characteristics:

1) type – PIC12F683;
2) company manufacturer– Microchip;
3) current consumption in active mode (IA, μA) – 100@2V, 1 MHz;
4) quiescent current consumption(Is, μA) – 0,001@2V;
5) minimum supply voltage(Umin, V) – 2;
6) speed, MIPS, MHz – 0,25;
7) ADC (bit, width / number) – 10/4;
8) case type (name / number of pins) – SOIC/8.

The main elements of the BTS for extramedullary osteosynthesis, in addition to the implantable biotelemetric module (IBM) (Fig. 2), are: the unit for the input of physiological data for the patient (UIPD); data exchange unit (DEU); software and hardware-software interfaces (SI) and (HSI); software module package SolidWorks; databases: meaningful models; mechanical properties of bones; biocompatibility of materials and bones; database on the parameters of plate-fixing materials, modulus of normal elasticity, Poisson's ratio, permissible stress; knowledge base of protocols of surgical interventions; block for predicting the development of osteosynthesis; block for making decisions on mounting and dismounting of the retainer plate; microcontroller 2, monitor, printer. The BTS also includes a communication modem via the Internet and an archive (actually a database) of tomographic and radiographic images. Database management is carried out through the MSKL DBMS.

The developed biotechnical system functions within the limits determined, first of all, by the subject area and target function, limitations, the main of which are:

- the system should give only insignificant minimum inaccuracies of research results;
- the system must maintain the integrity of the obtained experimental data under the influence of negative external influences and internal failures;
- the system must meet the price / quality criterion;

- the system should be able to quickly change the overall configuration of the system, including changes in the structure and function of its hardware;
- the system should provide personalized access with personal identification control;
- the system should be able to expand or change the structure and content of databases in accordance with new tasks.

The functionally developed BTS are consists of 2 modules: hardware and software, the first of which is made according to the classical configuration and includes a block for input of patient physiological data, an implanted biotelemetric module, a data exchange block, a hardware-software interface and a microcontroller hardware. The structural and functional organization of the BTS software module is shown in Fig. 4 and includes an intelligent interface, the central core of the system, databases and data stores.

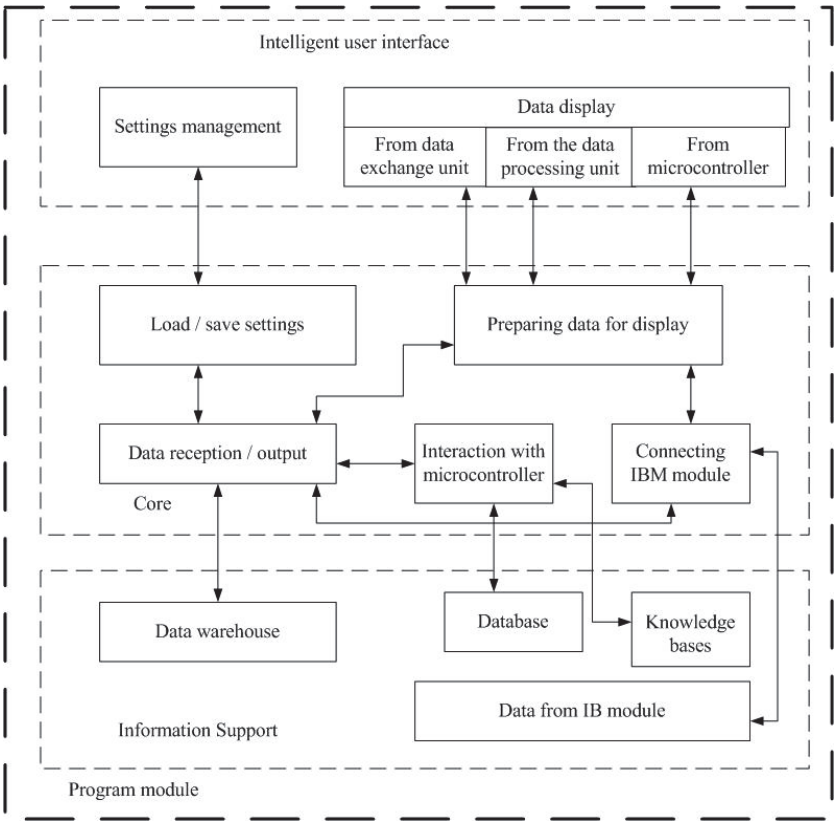


Fig. 4. Generalized diagram of the BTS software module.

In BTS are provided constant or periodic monitoring of the temperature of the plates and its location relative to the bone being treated. Such control allows timely detection of inflammation, bone displacement, etc. and take adequate measures to eliminate and

prevent them. The received telemetry data is sent to the microcontroller unit for processing in order to form a prediction of the development of bone osteosynthesis, check it for adequacy and reliability

5 Dental implantology and biotechnical system in implantology

5.1 Types of dental implants

Types of dental implantation: by the time of installation of the implant from the moment the tooth or its root is removed, they are distinguished: 1 - direct (immediate) implantation, in which dental implantation is carried out simultaneously with the operation to remove the tooth, that is, the implant is inserted into the tooth socket immediately after its removal; 2 - delayed implantation, in which dental implants are tightly installed in the created bed of intact bone tissue when the holes of the previously extracted teeth are no longer traced on the x-ray of the jaw, that is, the teeth or their roots were removed about a year ago[7].

On the basis of communication of a dental implant (its part) with the oral cavity: 1 - one-stage implantation, in which the dental implant is installed in one stage; 2 - two-stage implantation, in which at the first stage the body of the implant is installed and its "engraftment" occurs in the submucosal-periosteal area of the alveolar arch of the jaw, because after the introduction of the root part of the implant into the bone tissue of the jaw, the soft tissues are sutured over it. By the second stage, after “engraftment” of the implant body (about 3-6 months), its head is installed (external alveolar part of the DI)[7].

The indications for choosing the type of implantation, shape, and size of the implant are the topography of the dentition defect, the thickness and height of the existing bone tissue at the implantation site, topographic and anatomical features of the jaw structure, and many others.

5.2 Assessment of the condition of intraosseous implants

At the present stage of development of dental implantology, the main common methods for assessing the state of intraosseous implants next to the clinical one are X-ray, torque testing using a torque wrench when installing implants, gnatodynamometry, frequency-resonance stability analysis implants. Recently, a new trend has emerged in implantology - direct loading of implants with a temporary prosthesis immediately after the placement of the implants. At the same time, the requirements for identifying the degree of implant stability are increasing, which became the basis for the development of a new method for determining the stability of implants - resonance frequency analysis (RFA)[8].

Equally important is the assessment of the stability of the implants - when carrying out a two-stage implantation after the development of full osseointegration, which is

the fundamental condition for the long-term success of prosthetics with an emphasis on dental implants. The process of osseointegration depends on a significant number of external and internal factors: the structure and structure of the jaw bone tissue; the volume of bone tissue at the implantation site; characteristics of tools for forming a site for an implant; surgical technique; the term of the functional load of the implants; material and features of the implant surface; properties and methods of using osteopathic materials; osteoreparative potential of the patient's body..

In implantology, there are several possibilities for indirectly assessing the degree of osseointegration and stability of implants: clinical (manual control, implant stability); X-ray (including determination of bone density around the implant); echoosteometry; periotestometry; gnatodynamometric, orc-test using a torque wrench; resonance frequency analysis.

5.3 Biotelemetric complex for the study of physiological processes

The development and improvement of implantable medical devices (IMD), including dental implants (DI), is currently one of the priority directions in the development of medical technology. The main attention is paid to expanding the functionality of DIs, reducing their weight and size characteristics, increasing reliability, increasing the time of continuous autonomous operation and ensuring biocompatibility [9]. It was these components that ensured the rapid development of dental implantation, which allowed it to occupy a prominent place in the system of complex rehabilitation of patients with dental defects.

The creation of a class of intelligent implants is aimed, firstly, for solving the problems of monitoring the stability of implants, predicting their life cycle and preventing the occurrence of unwanted complications and inflammatory processes, and, secondly, for solving problems in the field of analysis and synthesis of the interaction of biological and technical elements. as a unified biotechnical system, including for pairing biosensors and sensors with dental implants.

This provides for the organization of microcontroller systems (MCS), focused on solving problems of transformation and primary processing of biosignals; study of the principles of constructing computational functional information converters for biotechnical systems (BTS) methods of synthesis of neural network converters of biosignals with a learning structure; development of the ability and skills to develop adapters - converters and separate autonomous microcontroller modules for communicating sensors with digital diagnostic systems.

The implanted biotelemetric complex is designed to study physiological processes occurring in internal organs, in which the parameters of these processes are measured by radiotelemetry. The complex consists of an implanted radio transmitter, a radio receiver, the information from which is fed to a computer, and software necessary for visualization and further processing of the obtained biophysical data[8].

In the general case, the biotechnical system for monitoring the status of physiological parameters includes two key components [9]:

- an implantable telemetried module operating inside the object directly measures physiological parameters and transfers the results using a digital radio channel;

- a distributed system for collecting and processing experimental data, wich are provide collection of information from several geographically distributed laboratories, storage of results in a unified format, providing researchers with remote access to the accumulated data.

The structural diagram of the system for monitoring the state of physiological indicators is shown in Fig. 5 [9].

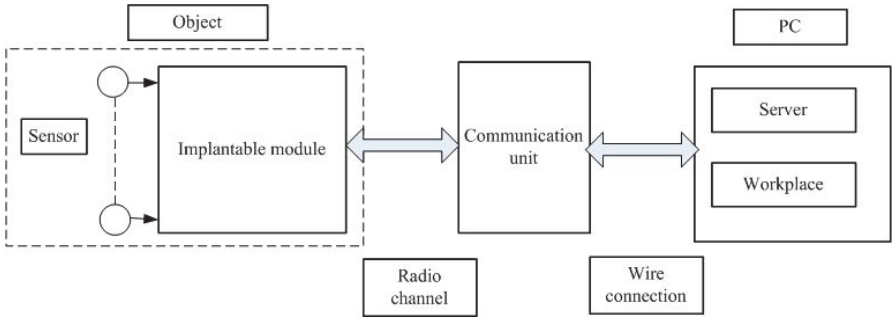


Fig. 5. Block diagram of the monitoring system of physiological indicators [9].

In fig. 6 shows a block diagram of the transmitting module of the biotelemetric complex.

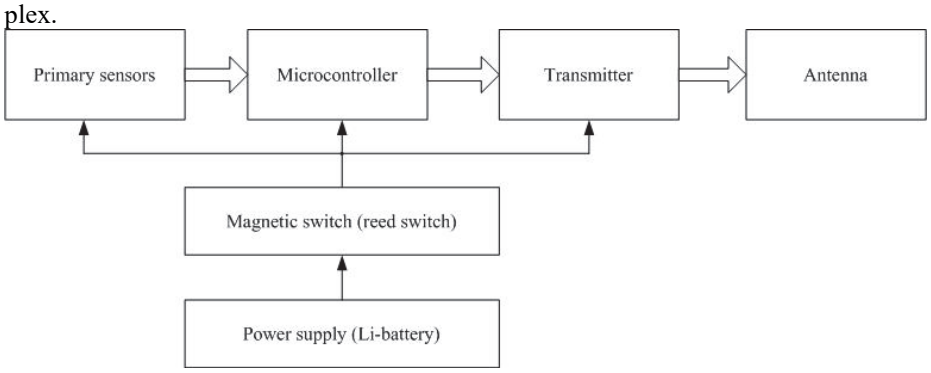


Fig. 6. Block diagram of the implantable transmitting module of the biotelemetric complex.

The implantable module will allow to carry out measurements of physiological indicators, provides two-way digital communication with the outside world, and processes commands to change the measurement mode. The communication unit receives digital information from the implantable module via a radio channel and transmits it to a personal computer via the USB interface. The communication unit provides two-way digital radio communication with the implantable module, ensures the safety of all data, retransmits control commands received from the PC that are implanted to the module, provides the PC, upon request, with the data obtained before the implantable modules[10].

The use of optical radiation for wireless transmission of energy through the skin is a promising direction in the development of implantable medical equipment. The use of directional channels for the transmission of energy and information in artificial sensors, and primarily in devices (such as a visual prosthesis), will significantly reduce the size of the energy receiver, avoid the influence of medical and household emitting devices on the transmission of energy and information, and minimize exposure to surrounding tissues, including brain tissue[9].

The operation of implantable systems has a number of features, the main of which is functioning in the near (inductive) radiation zone in specific environmental conditions, which determine the basic requirements for the transmission methods and the design of the part of the system that is implanted. When a radio signal passes through living tissue, it is significantly weakened - the more, the shorter the wavelength, therefore, low frequencies are used to transmit information from the internal parts of the body. However, the use of the low-frequency range of radio waves is associated with significant technical difficulties, in particular, with the difficulties of creating highly efficient emitters with strict restrictions on the dimensions and weight of the transmitter. The latter circumstance is of particular importance for systems being implanted. Therefore, for such systems that are located close to the receiving antenna, the frequency range of 0.05 ... 100 MHz is used. The limiting factor for increasing frequencies here is medical and biological requirements [10, 11].

6 Conclusions

The development of materials science, engineering methods of mathematical modeling significantly accelerates this process, makes it possible to make significant further significant steps towards improving and qualitatively improving methods and means of treating patients.

Solving the problem of the development, implementation and use of intelligent sensors for assessing the stability and biocompatibility of implants in dentistry will provide the user with new opportunities and provide him:

- means of monitoring the load on the implants immediately after their installation with subsequent identification and assessment of the degree of stability of the implants;
- monitoring control (continuous or periodic) for the stability and condition of the implants during two-stage implantation after full osseointegration, which is the basic condition for early successful prognosis based on dental implants;
- the biotechnical system and the implanted bone osteosynthesis module allow timely identification of possible deviations (inflammation, edema, swelling), thereby preventing possible complications.

References

1. Artiukh, G., Sorochan. E., Karlushin, S., Sergienko, Yu. and Kondrashin, S.: «On the issue of increasing the strength of clubs», Protection of metallurgical machines from breakdowns, № 11, PP. 189-192, 2009.

2. Klimovitskiy, V. and Borodin, D.: «Minimally invasive osteosynthesis in fractures of the proximal femur in elderly patients», Injury, V.14, №1, PP.4-9, 2013
3. Rublenik, I. and Bilyk, S.: «Minimally invasive bone osteosynthesis in the treatment of humeral fractures», Orthopedics, traumatology and prosthetics, №2, PP .111-113, 2002
4. Shaiko-Shaikovskiy, A. , Oleksiu, I., Bursyk, E., Azarkhov, A. , Sorochan, E. and Pastukhova, T.: «Comparative biomechanical assessment of the stability of the osteosynthesis of transverse diaphyseal fractures of the femur using various intramedullary and extramedullary structures», at the International Symposium: Reliability and Quality, Penza, 2016, PP.269-271
5. Bilyk, S., Gutsuliyk, K., Rublenik, I. and Shaiko-Shaikovskiy, A.: «Experimental determination of the attachment force of extra-bone fixators during osteosynthesis of fractures of tubular bones of different localization levels», at the International Symposium: Reliability and Quality, Penza, 2003, P.478.
6. Med-Net Microsemi multi-frequency modules for the development of implantable medical devices. [Electronic resource]. Access mode: http://www.icquest.ru/?article=270.
7. Dental implants. Access mode: https://dentaltechnic.info/index.php/obshie-voprosy/klinicheskayaortopedicheskayastomatologiya/2679-stomatologicheskie_implantaty.
8. Ramazanov, S. R.: Determination of implant stability as an objective method for predicting the effectiveness of treatment in dental implantology. Thesis for the degree of Ph.D. Specialty 14.00.21. – Moscow, 2009. – 79 p.
9. Danilov, A., Masloboev, Yu.P., Tereschenko, S., Titenok, S.: Simulation of transcutaneous wireless energy transmission using infrared radiation/ Medical equipment, 2011, №6(270). – PP. 18-21.
10. Grushko, A., Sheykin, S., Rostotskiy, I.: Contact pressure in hip endoprostethic swivel joints. Journal of Friction and Wear. – 2012. - 33(2), PP. 124-129..
11. Biotechnical Systems: Information Portal on Biomedical Engineering [Electronic resource]. – Access mode: http://ilab.xmedtest.net.

Architectural Characteristics of Biomedical Software Applications

Galyna Tabunshchyk [0000-0003-1429-5180], Olga Petrova [0000-0002-6499-6017], Tetiana Kapliienko [0000-0001-8192-2397], Peter Arras[2][0000-0002-9625-9054]

[1] National University "Zaporizhzhia Polytechnic", Zaporizhzhia, Ukraine
[2] KU Leuven, Campus De Nayer, Belgium

galina.tabunshchik@gmail.com, petrovaoa353@gmail.com, bragina.zntu@gmail.com , peter.arras@kuleuven.be

Abstract. In biomedical applications and the integrated IoT there can be distinguished similar structures in the development and construction of the software parts. As the data which is monitored, transmitted and used is very sensitive data concerning the patient's privacy, it is of the outmost importance that the design of the application and the data should be safeguarded.

Working on a medical device also means that the development process has to comply to the rules and standards which apply to biomedical devices.

In this chapter the architectural characteristics for the software development for biomedical devices is explored and directions for good practices in development with software patterns/templates is given.

As the design process for such a software design is known, the quality can be tested. The design process can be cut into templates which can be reused in further developments. This is illustrated in two different examples.

Keywords: architectural characteristics, IoT-M, biomedical application, Mg-alloy, informational infrastructure, voice navigation.

1 Introduction

When developing IoT applications it is very important to pay attention to the integration of data, interactions and the physical world. Interaction coupled with data transcends the laptop and mobile device, and it becomes literally embedded into any object, infrastructure or interaction.

When developing for biomedical applications, the rigid rules in the certification of applications puts an extra stress on the design. And the reliability of the software for the application should be near 100%.

This makes that the development process also closely needs to be monitored and systemized.

We should be aware that what is considered to be a medical device is strictly ruled and has to comply to standards and directives. The regulations give a number of requirements the medical device has to comply too. Whenever we design, construct, put

into something which is described as a medical device, the rules of the regulations apply.

In the new EU Medical Devices Regulations 2017/745 (as from May 2020 [1] (with a two years transformation period)) the definition is:

- 'Medical device' means any instrument, apparatus, appliance, software, implant, reagent, material or other article intended by the manufacturer to be used, alone or in combination, for human beings for one or more of the following specific medical purposes:
    - diagnosis, prevention, monitoring, prediction, prognosis, treatment or alleviation of disease,
    - diagnosis, monitoring, treatment, alleviation of, or compensation for, an injury or disability,
    - investigation, replacement or modification of the anatomy or of a physiological or pathological process or state,
    - providing information by means of in vitro examination of specimens derived from the human body, including organ, blood and tissue donations,

and which does not achieve its principal intended action by pharmacological, immunological or *metabolic means, in or on the human body, but which may be assisted in its function by such means. (EU (2017a)).*

Also software is as such considered to be a medical device and has to be processed according to the same directives. The quality and the reliability of the software is very important.

A unified development process with the use of templates is put forward to reuse proofed designs.

2 Classification of Biomedical Applications

Biomedical applications often involve the integration of a synthetic device, whether for interrogation or manipulation, with a living, biological organism. The resulting system of a synthetic device working with an organism can be considered a hybrid system, a system that combines traditionally dissimilar technologies to form a superior, multifunctional system [2].

The most common classes of Biomedical Applications:

1. Biomedical sensors. Currently, the primary focus of technological development is on various healthcare measurement instruments and monitoring devices which seek to improve the quality of healthcare monitoring. Biomedical sensors are miniature and lightweight devices that are used to monitor human vital signs e.g. heart rate, body motion, blood pulse pressure, blood glucose levels (Patel, Park, Bonato, Chan,

& Rodgers, 2012), process physiological data, and transmit the information to monitors. Wireless Sensor Networks (WSNs) technology in healthcare is a recent paradigm which incorporates wireless data communication and miniaturized sensors that measure these quantities, allowing portable health monitoring (Lai, Begg, & Palaniswami, 2011). Generally, wireless radio propagation is used to transmit data among the multiple sensors to a decision marker (hub node) or hub node to an actuator which is attached to the human body. The critical issue with radio frequency (RF) propagation is that it consumes battery life quickly and decreases the usefulness of portable monitoring devices (Seyedi, Kibret, Lai, Faulkner, 2013) [3].

2. Hybrid systems. One of the directions in biomedical application development are hybrid systems involving microfabrication and living organisms. Microfabrication has traditionally been a synonym with semiconductor manufacturing, the process used to mass produce electronic devices including everything from individual transistors to highly complex integrated circuits. In the past two decades, the once singular application of microfabrication has been expanded to include devices for micro-electromechanical systems (MEMS), micro-photonics, and microfluidics. The development of the resulting resonators, waveguides, and lab-on-a-chip devices has required new processes for novel substrates and materials and importantly, has redefined the original notion of the applications for microfabrication [2].
3. Data Mining Applications (biomedical applications with usage of data classification with machine learning methods). Over this decade loads of data have been generated, which has been per application individually analyzed and statistical evaluated to produce results. With current technology it is possible to generate information from data, knowledge from information, wisdom from knowledge and make decisions based in this generated wisdom with Artificial Intelligence methods such as machine learning [4]. The focus of biomedical research has shifted from data generation to data usage, or more specifically, biological data mining, to come to new scientific knowledge discovery. The emphasis is no longer on enabling the biologists and clinicians to generate more data rapidly, but on converting the data into useful knowledge through data mining and real time processing. In this scenario, effective biological data mining is crucial for developing a better understanding of disease mechanisms to discover new drugs and to develop data-based clinical decision making and support that is beneficial to the patients. In order to translate the vast amount of biomedical data into useful insights for clinical and healthcare applications, several data mining difficulties have to be overcome: handling noisy and incomplete data, processing compute-intensive tasks, integrating various data sources and exploiting biomedical data with ethical and privacy protection. All these issues pose new challenges for data mining scientists working in this data-intensive era [5].

Biomedical Data Mining Applications can be divided into [5]:

a. text mining in biomedicine and healthcare applications, which include entity recognition, entity linking, relation extraction, and co-reference resolution of essential text-mining technologies that can be applied in biomedicine and healthcare (for example, a case study of a biomedical text mining database

system that automatically recognizes and collects cardiovascular disease related genes);

b. learning to rank: ranking biomedical documents with only positive and unlabeled examples, applies positive and unlabeled learning algorithms (e.g. to an RNA-protein binding dataset collected from PubMed [6]) . Great variety of classification methods could be applied here.

c. applications for Automated mining based on biomedical literature. (E.g. of disease-specific protein interaction networks: Elucidation of protein interaction networks is essential for understanding biological processes and mechanisms.)

3 Requirements Analysis and Test Model for Biomedical Applications

Biomedical applications are domain dependent so requirements analysis is one of the most important stages which should be carried with UML (Unified Modeling Language) diagrams.

The general design process of the IoT Application is presented in the figure(**Fig. 1**).

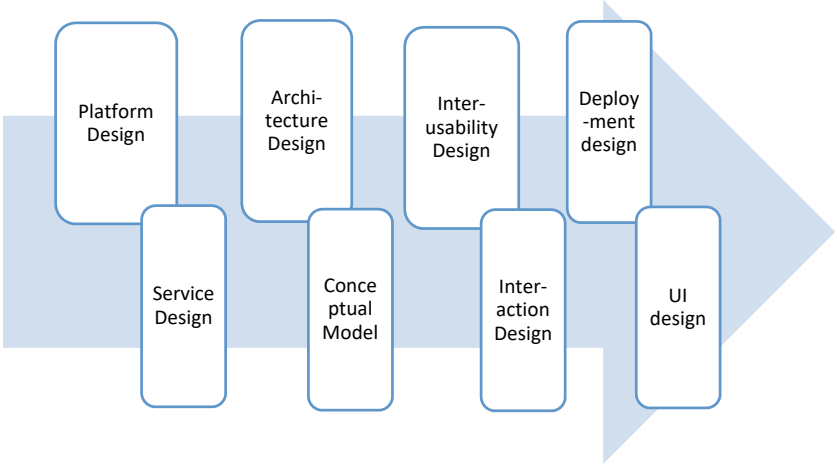


Fig. 1. General design process

The UML diagrams are used to represent the analysis, design and implementation phases in software engineering [8]. Among the various diagrams in UML, the use case diagrams play a prime role in software development because all the other diagrams in UML like class diagrams, activity diagrams, interaction diagrams are generated based on the use case flows in the use case diagrams [9]. In the analysis phase of a software

development, they are used to analyze the requirements and to represent the relationships between different processes involved in the project under development. The number of modules to be developed, the responsibilities of the actors, and the relationship between the actors are decided from the use case diagrams. They are also used to design the database at the backend. But, in case of very large projects in real time environment, the analysis and decision making process through use case.

In [7], an automated use case slicing model for object oriented design is proposed which is useful for designing medical expert systems. This proposed model uses graph generation and slicing techniques proposed by Jaiprakash et al. [11] for effective slicing of use case diagrams. Moreover, in the described case, the model helps the medical doctors for providing a treatment plan to the asthma patients. Thus it enhances the use case diagrams with artificial intelligence by applying rules for effective decision making.

The graph generation and slicing module consists of two phases namely the identification phase and slicing phase. The identification phase identifies the use case diagrams pertaining to an application and helps to bring them from the use case repository. The slicing phase uses the existing graph generation approach proposed by Jaiprakash et al. [11] to slice the use case diagram into actors and use cases and the use cases are used for graph generation. For demonstrating the proposed work, a medical registration application has been developed and implemented using UML modeling and also with the proposed model. In the medical registration module, the patient details, doctor details and the medicine and treatment details are stored. A knowledge base consisting of general rules for problem solving and domain rules for finding and treating asthma are stored. This knowledge base can be updated by adding more rules provided by the classifier through training. In [7], the use case diagram present in the use case repository is considered for analysis.

For the development of user-machine interactions there obviously exist a lot of challenges: complexity, manual design, reliability. IFML is one of the solutions for the designing front-end of the IoT applications. It is strongly integrated with modelling languages as UML, BPMN, SysML, SoaML [12].

On the other hand, Biomedical Applications reliability is one of the most critical parameters. The emphasis is on the analysis of failure data. In software reliability, one is interested in the failure mechanism. Most software reliability models are analytical models derived from assumptions of how failures occur. The emphasis is on the model's assumptions and the interpretation of parameters.

In order to develop a useful software reliability model and to make sound judgments when using the models, an in-depth understanding is needed of how software is produced; how errors are introduced, how software is tested, how errors occur, types of errors, and environmental factors can help us in justifying the reasonable-ness of the assumptions, the usefulness of the model, and the applicability of the model under a given user environment [13].

To use the data obtained during the compilation of the use-case, to increase the reliability of biomedical applications, the V-model can be used. The V-model is an Software Development Life Cycle (SDLC) model where execution of processes happens in a

sequential manner in a V-shape [14]. It is also known as Verification and Validation model.

The V-Model is an extension of the waterfall model and it is based on associating a testing phase to each corresponding development stage. This means that for every single phase in the development cycle, there is a directly associated testing phase. This is a highly-disciplined model and the next phase starts only after completion of the previous phase.

Under the V-Model, the corresponding testing phase of the development phase is planned in parallel. So, there are Verification phases on one side of the 'V' and Validation phases on the other side. The Coding Phase joins the two sides of the V-Model.

The following illustration depicts the different phases in a V-Model of the SDLC [14].

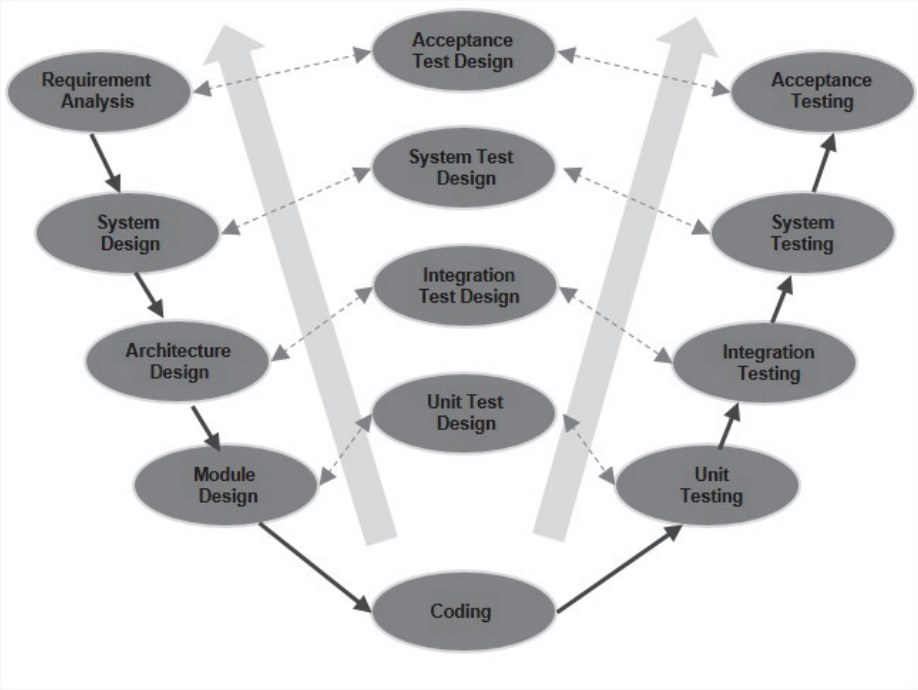


Fig. 2. Typical V-model

The V-Model method has as an advantage that it is very easy to understand and apply. The simplicity of this model also makes it easier to manage. The disadvantage is that the model is not flexible to changes. Whenever there is a requirement change, it becomes very expensive to make the change. Unfortunately today it is very common in to have this change requests.

V-Model method advantage are:

1. It is a highly-disciplined model and Phases are completed one at a time.

2. Works well for smaller projects where requirements are very well understood.
3. Simple and easy to understand and use.
4. Easy to manage due to the rigidity of the model. Each phase has specific deliverables and a review process.

V-Model method disadvantages:

1. High risk and uncertainty.
2. Not an appropriate model for object-oriented and complex projects.
3. Oversimplified model for long and ongoing projects.
4. Not suitable for the projects where requirements are at risk of changing.
5. Once an application is in the testing stage, it is difficult to go back and change a functionality.
6. No working software is produced until late during the life cycle.

4 Classification of patterns for biomedical applications

A design pattern is a repeatable architectural construction that represents a solution to a design problem within a frequently encountered context.

The main benefit of each individual pattern is that it describes the solution to a whole class of abstract problems. Also, the fact that each template has its own name makes it easier for developers to discuss abstract data structures, since they can refer to known templates. Thus, due to the templates, the terminology, names of modules and project elements are unified. A well-formulated design pattern allows you to use it again and again after finding a good solution. Design patterns can speed up the development process by providing tried and tested development paradigms. Effective software design requires considering issues that may only appear later in the implementation phase. Reusing design patterns helps avoid subtle nuances that can cause serious problems and improves code readability for programmers and architects familiar with patterns.

Typically, a template is not a complete sample that can be directly converted into code; this is just an example of a problem solution that can be used in various situations. Object-oriented patterns show the relationships and interactions between classes or objects, without specifying which final classes or application objects will be used [15].

Design patterns fall into the following categories: generative patterns, structural patterns, and behavioral patterns.

- Generative/Creational patterns: the aim of these patters is to provide possibilities for the creation of new objects while hiding the creating logic. Implementation of such an approach gives the programmer the flexibility in decision which object needs to be created in a given use-case. They are: factory method, abstract factory, builder, prototype, singleton.
- Structural patterns: these patterns deal with class and object composition. They are: adapter, bridge, composite, decorator, façade, flyweight, proxy.
- Behavioral patterns: these patterns are dealing with organizing communication between objects. They are – chain responsibility, command, iterator, mediator, memento, observer, state, strategy, visitor.

Considering the architectural patterns of design in the development of applications for Android, we should highlight popular patterns as MVC (Model-View-Controller), MVP (Model-View-Presenter), MVVM (Model-View-ViewModel), MVI (Model-View-Intent) [17-21].

The UI design patterns for the IoT systems can be grouped into three categories: Set Patterns, Get Patterns, and Event-based Patterns [23].

5 Case study

In the following case studies we look at different architectural characteristics when defining the software for biomedical related applications.

The first case describes the setup for a material test for Mg-implants where the data of in-vitro tests and in-vivo tests are combined to create a diagnostic supporting prediction model for the solubility of the Mg in the human body.

A second case concern an indoor navigation tool controlled by voice. The idea is that it can help visually impaired people to navigate through buildings.

5.1 Monitoring system for Mg implants

For monitoring of the new Mg alloys, which could be used for the implants there is developed a new monitoring system, which is integrating the biomedical data, material science data and IoT [22].

The monitoring system is built in a three layer architecture containing a device level, a server level and an application level (**Fig. 3**). All levels are developed independently and when it was designed there were used such patterns as façade, proxy, factory, builder.

The device level serves to collect data from different sources In general it can be divided into two types of data: sensor-information and images. Both types can be collected at in-vitro tests. In in-vivo tests mostly only imaging data (X-ray, tomography, scans..) can be collected. In the case of surgical extraction of (worn out) implants other data (like the loss of material, degradation of the material …) can be measured also.

At the figure 4 there is a picture of the device level for the simulation of the blood circulation.

At sever level, the technique to classify real-time data from several different sources is implemented the following steps: data normalization; the prediction of future points; analysis of residuals; probability calculation; conflicts definition; data fusion; classification; estimation of classification accuracy.

At the application or end user level several tools are developed for the collection of data, the analysis, and the (visual) representation of results of the experiments.

Data which is received from sensors during the tests is also used for the manipulation of the test control system.

Data resulting from in-vivo tests is reused afterwards for defining further in-vitro tests. There is a need for the continuous update of all the information and the matching of the different types of data.

A single model is implemented of combining all data which makes it possible to develop a high quality system for supporting medical trial decisions. In the system we integrate medical, biomedical, physical and clinical data .

Device level	Server level (Cloud level)	Application level (End users applications)
Data classification and aggregation Control system adjusting	Data collection/Stream publication/	Web and mobile applications for collection, analysis, and presentation results of the experiments

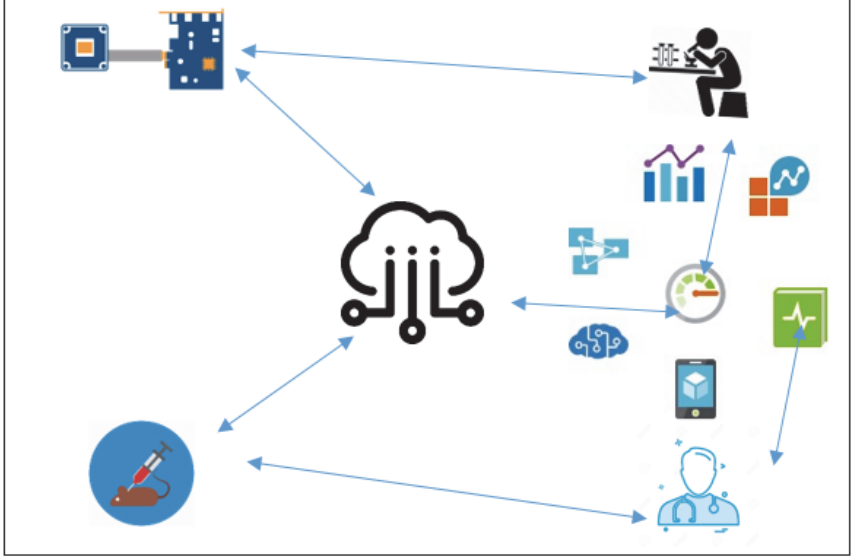


Fig. 3. Architecture of the monitoring system

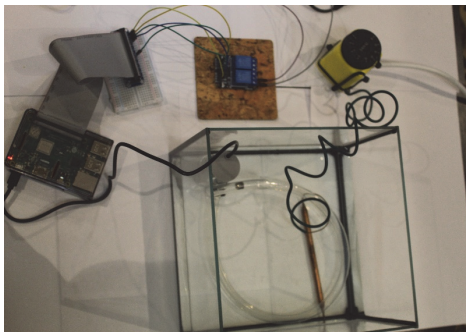


Fig. 4. Device level

The database infrastructure is developed and data from the experiments is collecting. The work in progress is the development of the method for real time detection degradation based on the convolutional neural networks [24].
For the software architecture there were used RESTful API[25], which allows to create reliable systems which could provide variety of client applications.

5.2 Voice recognition system for indoor navigation

A voice navigator was developed to facilitate the adaptation of visually impaired people to social benefits. The developed navigator has been integrated into the Smart-Campus system, which provides support for students, staff and university visitors. Use of voice for navigation systems allows to provide access to information in navigation systems, to interconnect objects and events, and to support new user interaction systems, sensors, mobile devices, devices and applications.

A voice navigator should help a person navigate the building using only the voice. For the correct exchange of information between the user and the application, a module must be developed that can recognize speech signals and can convert a speech signal into a text stream.

This process is often used in technologies that allow a computer or mobile device to be controlled by human voice and allows users to control the device by talking rather than by typing [26].

Modern language recognition systems are based on the principles of recognition of repeating patterns. Methods and algorithms for pattern recognition can be divided into four main classes: discriminant analysis methods which are based on Bayesian discrimination; hidden Markov models; dynamic programming - time dynamic algorithm (DTW); neural networks [26-28].

Developed by authors the voice-based navigation model (fig.4), based on the speech recognition method [26].

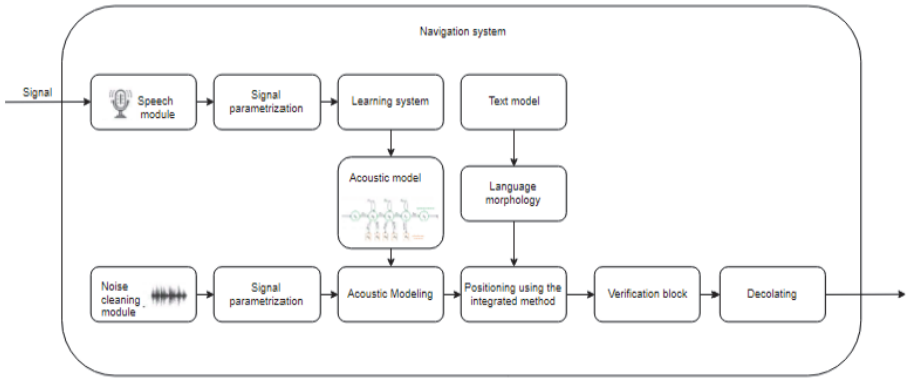


Fig. 5. Voice recognition model

Model and methods were implemented within the Smart-Campus application[24], which allows user with the voice command to display in the mobile application the current location inside the building and the shortest path to a specified position connected to the beacon.

Smart-Campus, is the name of a system with Bluetooth Low Energy devices and a backend database with dedicated content management system (CMS) [30]. An interactive tour around the campus or to guide visitors to their specific location of interest is automated by making a route to find the location from one beacon to the next and so on. To provide navigation, a map of the building should be provided or developed. Next is shown the logical path to another next beacon location. So the developed solution consists of two main parts: a map editor and path detection.

The map editor makes it possible to create a map of a floor of a building. A background picture of a known area can be used as a template or the map can be developed from scratch with the straight forward editor. The app-user is the client of information related to a certain beacon at a certain location and the app allows users to get stored information in a visual attractive format on the smart phone through a dedicated application. The app downloads the information from the server, linked by the unique user ID (UUID) of the beacon. The beacon broadcasts its UUID on regular basis.

On the server the information is stored and maintained by the beacon owners in the content management system. The app user can acknowledge or discard on groups of beacons to view their information.

The voice navigator is helping a person to find the location of the room and the body in which it is located. After the audience is found, the navigator will answer the question about the building in which the audience is located, on which floor and construct the map-device from the current position of the user to the required body

The interaction diagram for audio navigation is shown at fig.6.

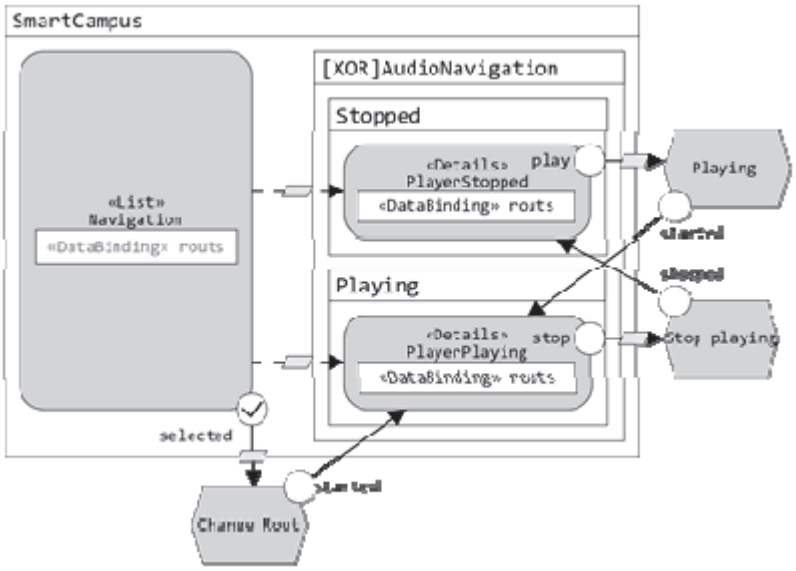


Fig. 6. Interactions in the voice navigator

6 Conclusions

When developing for IoT-M, the software/hardware combination has to comply to the rules set on biomedical devices.

In order to document and streamline the development process an approach with patterns is selected. The p, acting as proven building blocks are reused in future developments.

After the classification of the characteristics 2 case studies are highlighted which look at different aspects of the software development. By checking and certifying the solutions in a case study, a library of templates is constructed.

In the Mg-material test a 3 layer data model is put forward as a solution for the integration of different data types to make a prediction model.

In the voice navigator tool a databinding model with dynamically changing information which has to be presented to users is described.

The reuse of the proven building blocks shortens the development process for future applications.

7 Acknowledgement

This work is carried out partly with the support of Erasmus + KA2 project 586114-EPP-1-2017-1-ES-EPPKA2-CBHE-JP «Innovative Multidisciplinary Curriculum in Artificial Implants for Bio-Engineering BSc/MSc Degrees» [BIOART].

References

1. REGULATION (EU) 2017/745 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 5 April 2017 on medical devices, amending Directive 2001/83/EC, Regulation (EC) No 178/2002 and Regulation (EC) No 1223/2009 and repealing Council Directives 90/385/EEC and 93/42/EEC
2. Bhansali, S., Vasudev, A.: Mems for Biomedical Applications. Woodhead Publishing Limited, Cambridge (2012).
3. Seyedi, M., Kibret, B., Salibindla, S., Lai, D. T. H.: Encyclopedia of Information Science and Technology, Third Edition. An Overview of Intra-Body Communication Transceivers for Biomedical Applications. Information Resources Management Association, USA (2015).
4. Vardasca, R., Magalhaes, C., Mendes, J.: Biomedical Applications of Infrared Thermal Imaging: Current State of Machine Learning Classification. In: 15th International Workshop on Advanced Infrared Technology and Applications (AITA 2019), pp. 1–7. Florence, Italy (2019).
5. Li, X., Ng, S., Wang, J. T. L.: Biological data mining and its applications in healthcare. World Scientific Publishing Co. Pte. Ltd., Singapore (2014).
6. PubMed. Online access: https://pubmed.ncbi.nlm.nih.gov/ last accessed 2020/10/10.
7. Lalitha, R., Latha, B., Sumathi, G.: A software design technique for developing medical expert systems through use case analysis, https://www.biomedres.info/biomedical-research/a-software-design-technique-for-developing-medical-expert-systems-through-use-case-analysis.html, last accessed 2020/10/10.
8. Booch, G, Maksimchuk, RA, Engle, MW, Young, BJ, Conallen J.: Object oriented analysis and design with applications. Pearson Edu (3rd edn.) (2010).
9. Wu, CS, Khoury, I.: Web service composition: From UML to optimization. In: Fifth International Conference on Service Science and Innovation, pp. 1–2. IEEE, Kaohsiung, Taiwan (2013).
10. Zhu, L.Z, Kong, F.S.: Automatic conversion from UML to CPN for software performance evaluation. In: International Workshop on Information and Electronics Engineering (IWIEE), pp. 2682-2686, Elsevier Ltd (2012).
11. Jaiprakash, T.L., Mall, R.: A dynamic slicing technique for UML architectural models. In: IEEE Transactions on Software Engineering, pp. 735- 771, IEEE (2011).
12. Tabunshchyk, G.V. Interaction simulation for IoT systems. In: Internet of Things for Industry and Human Application. In Volumes 1-3. Volume 2. Modelling and Development [Text] / G.V. Tabunshchyk // eed. by V. S. Kharchenko. – Kharkiv: National Aerospace University KhAI, – PP. 110-134. (2019)
13. Software reliability models for critical applications, https://www.osti.gov/servlets/purl/10105800, last accessed 2020/10/10.
14. SDLC - V-Model, https://www.tutorialspoint.com/sdlc/sdlc_v_model.htm, last accessed 2020/10/10.

15. Freeman E. Design patterns . Peter, 656 (2017).
16. Behavioral, Creational, Structural Types of Design Patterns https://www.gofpatterns.com/design-patterns/module2/behavioral-creational-structural.php last accessed 2020/10/10.
17. MVC– Access: http://heim.ifi.uio.no/~trygver/themes/mvc/mvc-index.html
18. A cookbook for using the model-view controller user interface paradigm in Smalltalk-80– Access: https://dl.acm.org/doi/10.5555/50757.50759
19. Introduction to Model/View/ViewModel pattern for building WPF apps Access: https://docs.microsoft.com/ru-ru/archive/blogs/johngossman/introduction-to-modelviewviewmodel-pattern-for-building-wpf-apps last accessed 2020/10/10.
20. Android Architecture Components Access: https://developer.an droid.com/topic/libraries/architecture last accessed 2020/10/10.
21. Model-View-Intent (MVI) In Android Access: https://www.novatec-gmbh.de/en/blog/mvi-in-android/ last accessed 2020/10/10.
22. Tabunshchyk G., Shalomeev V., Arras P. Monitoring System for Tests of the Mg Implants, Proceedings of The Third International Workshop on Computer Modeling and Intelligent Systems (CMIS-2020), Zaporizhzhia, Ukraine, April 27-May 1,., pp. 70-78, (2020) http://ceur-ws.org/Vol-2608/paper6.pdf
23. M. Vega-Barbas, I. Pau, J. C. Augusto and F. Seoane, "Interaction Patterns for Smart Spaces: A Confident Interaction Design Solution for Pervasive Sensitive IoT Services," in IEEE Access, vol. 6, pp. 1126-1136, 2018, doi: 10.1109/ACCESS.2017.2777999.
24. K. Simonyan, A. Zisserman, Very Deep Convolutional Networks for Large–Scale Image Recognition, CoRR, abs/1409.1556, 2014. DOI: 10.1.1.740.6937
25. REST API Tutorial [Electronic resource]. – Access mode: https://restfulapi.net/
26. O.Petrova, G. Tabunshchyk, “Modelling of location detection for indoor navigation systems”, IEEE 9th International Conference on Intelligent Data Acquisition and Advanced Computing Systems (IDAACS), 961-964 (2017)
27. D.V. Shpakov, “Voice Recognition in the Sphere of Information Technologies” , Young Scientist, 2017, №29., 8-11. (2017)
28. D. S Kolesnikova, A. K Rudnichenko, E.A Vereshchagina, E.R. Fominova, “The application of modern speech recognition technologies in the creation of a linguistic simulator to enhance the level of linguistic competence in the field of intercultural communication” Primeneniye sovremen-nykh tekhnologiy raspoz-navaniya rechi pri sozda-nii lingvisticheskogo trenazhera dlya povysheniya urovnya yazykovoy kompetentsii v sfere mezhkul'-turnoy kommu-nikatsii], Internet journal "NAUKOVEDENIE, Vol. 9, No. 6. (2017)
29. J.-T. Chien, Linear Regression Based Bayesian Predictive Classification for Speech Recognition. IEEE Transactions on Speech and Audio Processing, vol. 11, no. 1 January (2003)
30. G. Tabunshchyk, D. Van Merode, “Intellectual Flexible Platform for Smart Beacons” In book: Edit by M. Auer, D. Zhutin Online Engineering and Internet of Things, Springer International Publishing, pp. 895-900. (2017) https://doi.org/10.1007/978-3-319-64352-6_83

The Data Dimensionality Reduction for Biomedical Applications

Sergey Subbotin[0000-0001-5814-8268]

National University Zaporizhzhia Polytechnic, Zaporizhzhia, Ukraine

e-mail: subbotin.csit@gmail.com

Abstract. The data dimensionality reduction problem for a model synthesis on instances is addressed. The set of methods for data reduction have been presented. It includes the intelligent methods of stochastic search applied for feature selection and sample selection problems. The feature selection methods based on evolutionary and swarm intelligent search are described. The theoretical estimations of requirements for the feature selection methods are provided. The individual informativity indictors of sample instances are presented. The experimental investigation of these methods has been provided. The set of presented methods allow selecting subsets of most informative features and instances allowing reducing time for model building.

Keywords: instance importance, sampling, diagnosis, feature selection, classification, data dimensionality reduction.

1 Introduction

At the various problems of decision making for biomedical applications, where there is a lack of available expert knowledge to classify the object state as a rule it is necessary to make a classification model based on a data sample, which contains a set of precedents or instances.

For the cases, where the data sample have a big dimensionality caused by the big number of instances and features, the use of such samples for a model synthesis, usually, provides enlargement of a model building time and of a computer memory volume required for problem solving. For acceleration of the model building process it is necessary to shorten the dimensionality of a data sample.

The wide used mean for data dimensionality reduction are methods of feature selection [1, 2]. Such methods select a subset of the most significant features from the original set of input features relative to output feature computation. These methods are divided by the type of new solution formation procedure on the methods implementing deterministic transition from one decision to another (exhaustive search methods that are complex and time-consuming [1, 3]) and the methods implementing random transition from the current to new solutions (stochastic search methods [2], that avoid the

hitting to the local minima, probabilistically taking into account information about concerned solutions, and provides search of suboptimal solutions for a limited time).

As a rule, the exhaustive methods [1, 3] need a big number of considered combinations of features, which a generated on the base of original feature set. It makes impossible the usage of these methods in cases where the initial feature set is big, because it will require a big volume of computational time. So, for the feature selection it is desirable to use the methods of intelligent stochastic search [2, 4], since they fit to obtain solutions combining the best solutions received at previous epochs and reduce the search. On the other hand, the exhaustive methods are tardy, that is caused by the fact that they do not take into account the information specific for the problem in the search process.

Thus, we need a set of intelligent methods for a feature selection able speed-up the receiving of solutions.

The other possible mean to reduce the data sample dimension is sampling methods [11–22].

The methods of sample selection [11–21] striving to select from a large initial sample the small subsample aiming to provide the saving of the most important characteristics from the initial sample to the subsample. Such methods should be split on probabilistic and deterministic methods.

The sample selection based on probabilistic approach [11, 12, 20, 21] using a random instance selection from the initial set of instances. In these methods instances have a non-zero probability to be included to the resulting sample. The merit of probabilistic methods consists in their primality and the availability to evaluate error of a sample. The imperfection of these methods consists in that fact that there is no guaranty that they will obtain the small non-excessive sample able to reflect key properties of the initial sample.

The sample selection based on deterministic approach [13–19] makes picking of precedents taking into account supposition on their information content. Here certain instances may be not choosen, or their selection probability cannot be accurately estimated. These methods are often grounded on the clustering and aimed to provide topological conformity of the initial sample. The imperfection of the deterministic methods is the inability of the evaluation of the error for created samples. The merit of the deterministic methods consists in that they may find the most promising instances for a model building. Such instances should be used for initializing model and to speed-up the training of a model.

Thus, we need to combine probabilistic and deterministic approaches for sample selection. The evolutionary methods by their nature are stochastic (probabilistic) search and are the methods of a global search. These methods do not require any special requirements for the goal function. They consider the solutions formed at previous search stages. So such methods are promising basis for sample selection. At the same time, evolutionary search is slow, that is caused by the fact that it is neglet the information specific for the solved problem in the search process.

So, it is necessary to speed-up the process of solutions generation in a sample selection method based on evolutionary search.

2 Problem formulation

Let we have a a sample of precedents (instances) $X = \langle x, y \rangle$, where $x=\{x^s\}$, $y=\{y^s\}$, $x=\{x_j\}$, $x^s=\{x^s_j\}$, $x_j=\{x^s_j\}$, $s = 1,...,S$, $j = 1,...,N$. Here we denote as x^s_j the value of j-th input feature x_j for x^s, as y^s the output value for x^s, as S the number of precedents, as N the number of input features. For the problems with integer output (classification) $y^s \in \{1,...,K\}$, $K>1$, where K is a number of classes. Let denote a subset of $\langle x,y \rangle$ as $\langle x', y' \rangle$. Then we denote: N' is a number of features in a sample $\langle x', y' \rangle$, S' is a number of precedents (examples) in a sample $\langle x', y' \rangle$, $f()$ a user defined criterion showing a quality of argument for the problem, *opt* is an optimal value of the criterion $f()$ for the corresponding problem.

Then the feature selection problem [22] should be presented as: reduce the given data set $\langle x, y \rangle$ to obtain the $\langle x', y \rangle$, where $x' \subset \{x_j\}$, $S' = S$, $N' < N$, and the criterion $f(\langle x',y \rangle,\langle x,y \rangle) \rightarrow opt$.

The training sample selection problem [23] should be presented as: for the given data set $X=\langle x,y \rangle$ find subsample $X'=\langle x', y' \rangle$, where $X' \subset X$, $y'=\{y^s | x^s \in x'\}$, $N' = N$, $S' < S$, and the criterion $f(\langle x', y' \rangle, \langle x, y \rangle) \rightarrow opt$.

3 The methods of a feature selection based on the intelligent stochastic search

The evolutionary methods [4, 5] and multi-agent methods [6, 7] are more widely used stochastic methods (Fig. 1) at the practice of optimization.

The Fig. 1 d) show an evolutionary method [4, 5, 22] scheme. This method is founded on the idea of the development of solutions inspired by the natural selection. Method starting to set the epoch counter as t=0. Then method creates initial solution set $P_t = \{H_j\}$, where $j = 1, 2, \ldots, N$, N is a number of solutions in a set, $H_j = \{h_{ij}\}$ is a j-th solution, h_{ij} is a i-th bit (gene) value of j-th solution, $i = 1, 2, ..., L$, where L is a solution's length. After that the iterative replacement of one set of solutions by the next is provided to receive more adopted solution. The method evaluate goal function for solutions at t-th population $\{f(H_j)\}$. Then it verify the reaching of the stop conditions. When the stopping conditions are contented then the method finish the search. If the stopping conditions are not contented then for the current solution set method use a selection operator that probabilistically select solutions for crossing over and mutation grounding on the $f(H_j)$ preferring the better adopted solutions, the crossover operator that makes parent pairs and then produce new solutions on their base, and mutation operator to provide random changes in solutions for expanding the solutions diversity.

The new population is formed based on the produced solutions. It contains new and best solutions, whose goal function values are the best in the population. The best crossover solutions provides that the foremost promising regions of a search space are investigated. At the end the solution set will converge to an optimal problem solution.

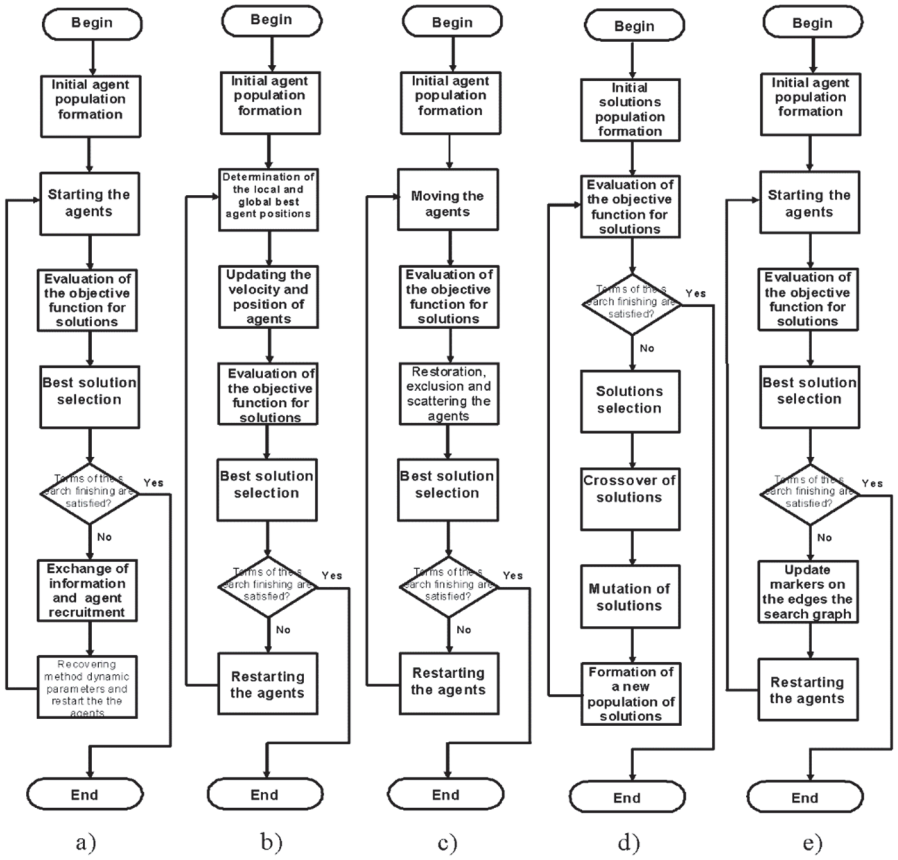


Fig. 1. Schemas of intelligent stochastic search methods:
a) BCO, b) PSO, c) BFO, d) evolutionary, e) ACO

The multi-agent methods [6, 7, 22] is a stochastic methods of multidimensional search. Such methods inspired by the behavior of agent colonies with the collective nature. The ability to solve optimization problems by these methods based on agents' actions for the spatio-temporal location (coordination), expand the actions of the agent and forming the spatial, temporal, and social relations between the agents in the colony's work (specialization), association mechanisms of the agents (cooperation), and reaction of the agents of the colony on the external environment changes (agents collective decision-making).

The multiagent methods may be divided into methods with indirect agent linking and methods with direct agent linking. The direct methods are represented by the Bee Colony Optimization (BCO) [7], which schema is illustrated by Fig. 1 a). The indirect methods are represented by the Particle Swarm optimization (PSO) [4], which schema is shown at Fig. 1 b), by the Bacteria Foruaging Optimization [6] (BFO), shown at Fig. 1 c), and by the Ant Colony Optimization (ACO) [4], which schema is shown at Fig. 1 e).

The stochastic search methods able to optimize the functions of any complexity. They have no extra requirements to the goal function such as continuity, smoothness, unimodality, differentiability, and monotony. Therefore the stochastic search is suitable for the sample and feature selection.

However, the methods of intelligent stochastic search have such shortcomings as method's speed dependence from the search starting point, high iterativity and uncertainty of the convergence time, as well as the search operators disregard the available information in a sample. So, the acceleration of methods of stochastic search through decrease these shortcomings are relevant.

For acceleration of the stochastic feature selection it is convenient to save their stochastic nature, but expand them by using of deterministic search operators. This will make possible to use a priori information on data samples. We propose to use the individual evaluations of feature informativity [9] as a priory information about data [10]. As these methods are classical methods of stochastic search we shall only show their main differences from the basic methods.

3.1 Feature selection method using the entropy

This method based on a canonical evolutionary search [5] with additions [8]. At the stage of solution forming the method set the probability of i-th feature inclusion to the solution: $P_i = 1–e_i$, where e_i is a i-th feature entropy [9, 10]. At the stage of solution crossingover the method use mask $h^*=\{h_i^*\}$, where $h_i^*=1$, if $e_i<e_{\text{tr.}}$ ($e_{\text{tr.}}$ is a threshold), $h_i^*=0$, otherwise. At the stage of solution mutation the method increases mutation probabilities P_M of features with high e_i. For the features with low e_i are the probabilities decreased. At the stage of solution selection stage the method use as a goal function (1):

$$f(H_j)=\left(1+\left(\sum_{i=1}^{N} h_{ij}\right)^{-1}\left(\sum_{i=1}^{N}\sum_{k=1}^{N} h_{ij}h_{ik}d_{ik}\right)E_j\right)^{-1}\left(\sum_{i=1}^{N} e_i\cdot h_{ij}\right)^{-1}, E_j \geq 0, \qquad (1)$$

where H_j is a j-th solution, h_{ij} is an i-th digit of j-th solution, E_j is a model error for j-th solution, d_{ik} is an individual informativity value of i-th feature relative to k-th feature.

3.2 Feature selection method with feature grouping

This method is based on canonical evolutionary search [22]. The stage of solution forming suppose that the set of solutions $\{H_j\}$, $H_j=\{h_{ij}\}$ should be created, where h_{ij} is a digit specifying usage of i-th feature taking into account its individual invormativity value I_i. At the stage of solutions crossingover the method use mask $h^*=\{h_i^*\}$, where $h_i^*=1$, if $I_i>I_{\text{tr.}}$, ($I_{\text{tr.}}$ is a previously specified threshold value), $h_i^*=0$, otherwise. At the stage of solutions mutation the method decrease mutation probabilities P_M of features with high I_i. The method increase mutation probabilities of features with low I_i. At the stage of solution selection the method use as a goal function (2):

$$f(H_j)=\left(\sum_{i=1}^{N} I_i h_{ij}\right)\left(1+\left(\sum_{i=1}^{H} h_{ij}\right)\left(\sum_{i=1}^{N}\sum_{k=1}^{N} h_{ij}h_{ik}d_{jk}\right)E_j\right)^{-1}, E_j \geq 0. \quad (2)$$

3.3 Feature selection method based on eevolutionary search with clustering

This method based on canonical evolutionary search [5]. At the beginning it evaluates the values of individual feature informativities $\{I_i\}$ according to outputs. On their base the indicators of feature similarity are evaluated. Then the conformable features are split into groups. After that in each group the most significant feature is selected. Such feature becomes a corresponding cluster center C_q. At the solution making stage the probabilities of inclusion of the features to a solution calculated grounding on the Euclidean distance from the cluster center $d(x_i,C_q)$, the individual informativities of a feature and cluster center I_i and I_{Cq}, $P_i=I_i+(I_i-I_{Cq})\{d(x_i,C_q)/d_{q\max}|x_i \in C_q, q=1, 2, ...,Q\}$, where $d_{q\max}$ is a maximal distance in q-th cluster.

3.4 Feature selection with search space fixing

The method [22] at the beginning sort features in order of I_i increasing. After that the first $\alpha\eta N$ features are excluded, where α is a coefficient, $0<\alpha\leq 1/\eta$, η is a coefficient of a search space reduction (3):

$$\eta=N^{-1}\sum_{i=1}^{N}\{1|I_i<I_{\text{avg.}}\}, \quad (3)$$

where $I_{\text{avg.}}$ is an average feature significance value. Then for reduced feature space method execute classical evolutionary search [5].

3.5 Feature selection based on multi-agent search with representation of the destination points by features

For the feature selection this method uses a multi-agent search with indirect communication inspired by ACO model [4, 9]. To form the search space the destination points represented as vertices of the search graph, in which the features are placed. The search is organized according the following strategy, when each agent should go through the specified number of destinations N', which defines the number of features that should be leave. For the representation of solutions this method encode the path traveled by the j-th agent is a feature set (solution) H_j, on which the model is synthesized: $H_j = \{h_{ij}\}$, $h_{ij} =$ 1, if the point i is included in the j-th agent path, $h_{ij} = 0$ – otherwise.

3.6 Feature selection based on BFO model

This method uses a BFO model [6] with indirect agent linking as search model with such enhancements. The space for search have special encoding: points encoded by binary strings consisting of digits related to feature importance. The BFO model is modified using operators of simple mutation and proportional selection. To speed-up the search process at the initialization stage for all i-th agents, i=1, 2, ..., H, the initial position

defined as: $h_{ji}=1$, if $rand_{ji}<I_j(\sum_{q=1}^{N}I_q)^{-1}$, $h_{ji}=0$, otherwise, where $rand_{ji}\in[0; 1]$ is random numbers, I_j is a j-th feature individual informativity, $j=1, 2, ..., N$.

3.7 Feature selection based on multi-agent search with feature importance representation of the destination points

This method of a feature selection is based on ACO model [9] with such improvements. For the creation of a search space the destination points as a search graph vertices represented by randomly made binary set $B = \{b_i\}$, $b_i = \{0;1\}$, $i=1, 2, ..., N$, containing N' elements equal to 1. The method provides a search strategy, where agents proceed the way through all destination points. Especial format of encoding of solutions is provided, where the path of j-th agent positions represented by destination points is a binary set of significant features B used to build the model.

3.8 Feature selection based on multi-agent search with direct linking

This method use direct agent linking inspired by BCO model [10] providing some updates. The feature importance is represented by agent collected resource in the $(N'\times N)$-dimensional space. For each solution the goal function evaluated on the basis of usefulness of agent placement at epoch t in a source h, in which $H^h(t)$ agents are placed: $l^h(t)=a_h/H^h(t)$, $h = 1, 2, ..., N'\times N$, where $a_h = \varepsilon^*/ E_h$, E_h is an error of a model for source h, ε^* is a reduction factor. To accelerate the speed of search method splits agents. After the starting of the agents-scouts the one part of scouts randomly distributed in the search space providing global search, and the other part is distributed in the sources proportionally to theit importance providing usage of a priori information.

3.9 The analysis of spatiotemporal complexity of feature selection

To characterize complexity for the above described methods of feature selection the following notations is used: T is a number of epochs of a search method, H is a number of or solutions (for the evolutionary search) or a number of agents (for multi-agent search), F is a simulation complexity of neural model and the complexity of the error evaluation for a training sample.

In [22] the accurate estimates of the complexity are given without suppressing of less power members by greater power members. The rough estimates where greater powers members suppress less power members obtained taking $n = NS \approx N^2$. For the sequential computations the analytical assessments of temporal and spatial complexities for the feature selection methods are given.

For the method 3.1 the accurate time complexity is $O(12NS+N^2+N+THF)$, for the method 3.2 the accurate time complexity is $O(12SN(N+1)+N+THF)$, for the method 3.3 the accurate time complexity is $O(12NS+N+THF)$, for the method 3.4 the accurate time complexity is $O(1.5S(N–1)+C+5N+TFH)$, for the method 3.5 the accurate time complexity is $O(TH(N^2+4N+H+F))$, for the method 3.6 the accurate time complexity is $O(TH(N^2+F+4N))$, for the method 3.7 the accurate time complexity is

$O(8H^2+6H^2N+HF+N+TH(2F+H+N+4))$, for the method 3.8 the accurate time complexity is $O(H+TH(F+9+HN–H))$. The rough time complexity estimate for all these methods is $O(n^2\sqrt{n})$.

For the method 3.1 the accurate spatial complexity estimate is $O(13NS+0.0625HFN)$, for the method 3.2 the accurate spatial complexity estimate is $O(N(S+N+2)+0.0625HFN)$, for the method 3.3 the accurate spatial complexity estimate is $O(NS+2N+0.0625HFN)$, for the method 3.4 the accurate spatial complexity estimate is $O(NS+3N+0.0625HFN+C)$. The rough spatial complexity estimate for all these methods is $O(n^2)$. For the method 3.5 the accurate spatial complexity estimate is $O(NS+H(3N+F+H))$, for the method 3.6 the accurate spatial complexity estimate is $O(NS+3HN+HF)$, for the method 3.7 the accurate spatial complexity estimate is $O(NS+HN+3HF)$, for the method 3.8 the accurate spatial complexity estimate is $O(NS+5H+HN^2–HN+FH)$. The rough spatial complexity estimate for all these methods is $O(n\sqrt{n})$.

4 Evolutionary method of sample selection

During the search to shorten the matching feature combinations it is possible using the data on solutions before analyzed to considerate the new solutions, similar to the previously considered. We also need to provide the chance to be taken into account for each possible solution.

The evolutionary method of sample selection with pseudo clustering [3, 23] should be presented as a consequence of the next stages.

The stage of initialization. Give the initial sample of S instances $<X, Y>$, the maximum number of instances S_f in a resulting sample $<x, y>$, the acceptable value of a criterion I^*, the evolutionary search parameters, the mutation probability $P_м$. Evaluate the criterion I for the $<X, Y>$.

The stage of a solution pseudo-clustering. Consequentially process the features: $i = 1, 2, ..., N$. For the i-th feature sort the original sample instances in a decreasing order by i-th feature. After that for ordered set of instances at i-th feature axis include to the $<X',Y'>$ pair of each two adjacent instances, which belongs to the different classes. Along the i-th feature axis the left and right instances should be included to the $<X',Y'>$.

The stage of initial population generation. Form the H binary decisions**:** set the probabilities of instance inclusion to the k-th solution as (4) for $k = 1, 2, ..., H$, $s = 1, 2, ..., S$:

$$P(h_s^k) = \begin{cases} \dfrac{\lambda + \text{rand}}{2}, X^s \in < X', Y' >; \\ 0{,}5\text{rand}, X^s \notin < X', Y' >, \end{cases} \tag{4}$$

where

$$\lambda = \begin{cases} 1, S' \le S_f; \\ \min\left\{0{,}5; \dfrac{S_f}{S'}\right\}, S' > S_f. \end{cases} \quad (5)$$

h^k_s is an s-th bit of a k-th solution, rand is a random function.

The testing stage to check the reaching search end. Method makes a sample for each k-th solution. Then the values of $I(k)$ evaluated. If the number of executed epochs is more than T, or there is a solution with the number k: $I(k) \ge I^*$, then search should be finished and the sample with the highest value of $I(k)$ should be returned as a result.

The stage of the selection of decisions. The solutions pairs (parents) are produced on the basis of the Monte Carlo rule aiming to reach the goal $I(k) \to \max$ [3].

The stage of solutions crossing. The new solutions are produced using a one-point crossingover [13, 14].

The stage of mutation. For each obtained solution method randomly invert no more than round($P_м S$) bits. After inverting select the solutions having more than S_ϕ the number of bits equal to one. Then randomly invert the rest of bits equal to 1. Then method should delete from the population those solutions that were considered before. After this stage the method should execute the testing stage to ckeck the search end reaching conditions.

The described method joins the random sample selection and the determined search to process the best solutions. It begins the search process from the selection of the most promising instances to include them into solutions. However, for the rest of instances it retains the chances to include them into resulted sample. During the methods' work it aims to improve the taken solutions. The method also ensures for every considered sample to have a size less or equal to S_f.

5 Instance selection for sample forming

To range the instances for sample forming it is possible to use the following set of indicators describing the individual instance importance for sample forming [14, 23].

- The indicator of the s-th instance individual informativity regarding to the interclass border is determined by the formula (6):

$$I^s_{\overline{G}} = \frac{1}{S-1} \sum_{p=1}^{S} \left\{ e^{-\alpha_{y^s, y^p} \sum_{j=1}^{N} (x^s_j - x^p_j)^2} \,\middle|\, s \ne p,\ y^s \ne y^p \right\}, \quad (6)$$

where $\alpha_{k,q} = \dfrac{1}{\dfrac{1}{S^k S^q} \sum_{s=1}^{S} \sum_{p=1}^{S} \left\{ \sum_{i=1}^{N} \left(x^s_i - x^p_i \right)^2 \,\middle|\, (y^s = k, y^p = q) \vee (y^s = q, y^p = k) \right\}}$.

The indicator's value will be in a range between zero and one. The closely located the instance to the interclass boundary the bigger will be indicator’s value.

- The indicator of the *s*-th instance individual informativity regarding to its distance to the class border is determined by the formula (7):

$$I_U^s = 1 - \min_{p=1,2,...,S} \left\{ \mathrm{e}^{-\alpha_{y^s} \sum_{j=1}^{N} (x_j^s - x_j^p)^2} \middle| s \neq p, y^s = y^p \right\}, \tag{7}$$

where $\alpha_k = \dfrac{1}{\max\limits_{\substack{s=1,2,...,S;\\ p=s+1,...,S}} \left\{ \sum\limits_{j=1}^{N} \left(x_j^s - x_j^p \right)^2 \middle| y^s = k,\ y^p = k \right\}}$;

- The indicator's value will in be in the range from 0 to 1. The closer the corresponding instance will be to the class core the bigger will be the value of indicator.
- The indicator of the *s*-th instance individual informativity regarding to the proximity of instance to the cluster center is defined by the formula (8):

$$I_O^s = \frac{1}{S^{y^s} - 1} \sum_{p=1}^{S} \left\{ \mathrm{e}^{-\sum_{j=1}^{N} (x_j^s - x_j^p)^2} \middle| s \neq p,\ y^s = y^p \right\}. \tag{8}$$

The indicator (8) value will be in the range between 0 and 1. The farther the instance will be regarding to the rest instances of a sample the bigger will be it's value. This indicator allows to select instances placed on the classes outlier

The x^s instance's individual informativity integral indicator should be defined as: $I(x^s) = \max\{I_{\bar{G}}^s, I_O^s, I_U^s\}$. This indicator have a value in a range between 0 and 1. The greater the indicator's value, then the more significant is the *s*-th instance relative to the model building. This is becouse this instance is placed at the bound of the classes, or it is a rare, or this instance is located at the class outlier, or instance is near the class core.

- The sample instances group informativity indicator is defined by the formula (9):

$$\bar{I} = \max_{x^s \in X} \{ I(x^s) \}. \tag{9}$$

6 Experimental examples

The experimental investigation of the feature selection methods have been performed for different data sets [22]. The properties of these data are presented in the Table 1.

The experimentally obtained results are presented in the Table 2, where the following designations are used: N_m is a size in Megabytes of used computer memory by the method, t is a time in seconds of a feature selection, $fr = N'/N$ is a fraction of features taken from the initial data set to the reduced data set.

Table 1. The data set properties

Data set	Abbreviation	*S*	*N*
1. Chronic obstructive bronchitis	2. HBR	205	28
3. Simulation the total index of quality of life of patients sick of a chronic obstructive bronchitis	4. KOG	86	106
5. Children's health state modeling taking into account environmental pollution	6. ECO	954	43
7. Agricultural plants recognition [24]	8. PLNT	248	55

Table 2. Generalized estimations for feature selection methods

9. Data set	10. Search space fixing			11. Feature grouping		
	12. *t*	13. $N_{m.}$	14. *fr*	15. *t*	16. $N_{m.}$	17. *fr*
KOG	6984.1	10.10	0.132	6984.1	10.10	0.132
HBR	5429.2	14.75	0.679	5429.2	14.75	0.679
ECO	13832	487.84	0.628	13832	487.84	0.628
PLNT	1538.2	45.09	0.218	1538.2	45.09	0.218
18.	19. Entropy using			20. Feature clustering		
	21. *t*, sec	22. N_{m}, Mb	23. *fr*	24. *t*	25. $N_{m.}$	26. *fr*
KOG	5864	10.65	0.113	6686.6	9.55	0.132
HBR	4700.1	14.46	0.714	4591.6	13.86	0.750
ECO	12867	500.39	0.651	14764	472.47	0.628
PLNT	1334.2	45.94	0.200	1436.6	41.83	0.164
27.	28. Ant colony with features			29. Ant colony with feature informativities		
30.	31. *t*	32. $N_{m.}$	33. *fr*	34. *t*	35. $N_{m.}$	36. *fr*
KOG	12282	1.34	0.123	12792	1.25	0.132
HBR	9327.3	1.96	0.750	9739.1	1.72	0.786
ECO	28000	57.60	0.651	25507	57.30	0.651
PLNT	2922.6	4.98	0.164	2814.9	4.68	0.200
	37. Bee colony optimization			38. Bacteria fouraging optimization		
KOG	39. *t*	40. $N_{m.}$	41. *fr*	42. *t*	43. $N_{m.}$	44. *fr*
HBR	13558	3.73	0.113	14033	1.66	0.142
ECO	9822.1	5.30	0.750	9394.1	2.28	0.714
PLNT	28690	168.21	0.628	27119	76.35	0.651
	3072.3	15.91	0.200	2894.8	6.52	0.200

	45. Canonical evolutionary method		
	46. *t*	47. $N_{m.}$	48. *fr*
KOG	21909	10.00	0.132
HBR	17150	14.12	0.786
ECO	43335	430.40	0.651
PLNT	4714	40.74	0.200

The conducted study is proving the applicability of the proposed methods. The methods of feature selection usage provide acceleration of a search speed relatively to pure evolutionary search [3, 5]. It should be explained by the fact that they use less number of requests of a goal function evaluations and thereafter provide less number of model buildings and adjusting.

For the problems for which there are no strong limitations for available memory resources the evolutionary methods should be recommended. For the problems for which the strong limitations are exists the multi-agent methods should be used.

For the sample selection methods the important property is a search space processing speed exploration. It should be characterized numerically by the λ which designates a fraction of reviewed or eliminated solutions from the whole possible solution set at corresponding method's iteration (epoch). For the developed sample selection method the graph of the λ dependence of the iteration have been analyzed by experimental results averaging.

Fig. 2 demonstrates the obtained graphs for λ in a semi-logarithmic coordinate system relatively to exhaustive search with reduction. Here the following designations are used: T is a number of epochs, H is a number of solutions [11, 12].

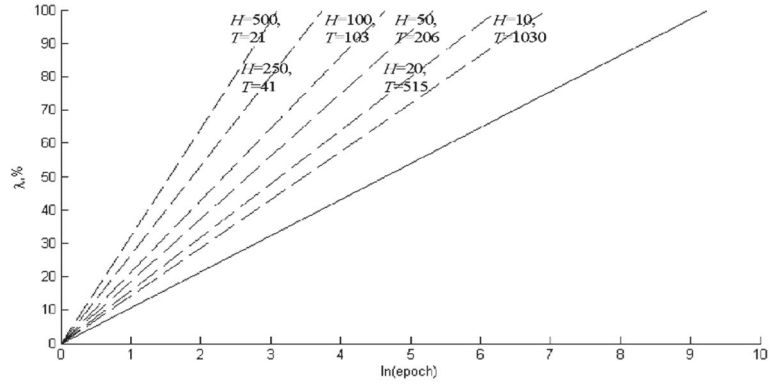


Fig. 2. Search space reduction graphs for the evolutionary search (– –) and exhaustive search with reduction (–) based sample selection methods

The sample selection methods investigation has been made concerning their efficiency. The Fig. 2 illustrates the λ graph for the methods of sample forming relative to the executed epochs. It can be seen from the Fig 2, that the bigger the size of population then the method's speed is increased. However the allowable number of size iterations is significantly reduced. It should be dependable from the original sample size.

To demonstrate the work of the instance informativity selection on the basis of the proposed set of indicators they have been applied for the Fisher IRIS dataset [24]. The part of the obtained experimental results is illustrated by the Fig. 3. Here IRIS classes are marked by dot, cross, and plus markers. The circle around the class markers designate the instances for which the corresponding indicator is bigger than average value for whole data set.

The provided studies have proven the practical usefulness of the proposed indicators of individual informativity of instances. They make possible recommend proposed indicators for usage at problems of sample dimensionality reduction and sample forming.

At the Fig. 3 the instances marked by the circles are located on the boundaries and in the cores of the classes. Fig. 3 shows the applicability of the proposed indicators of individual instance importance.

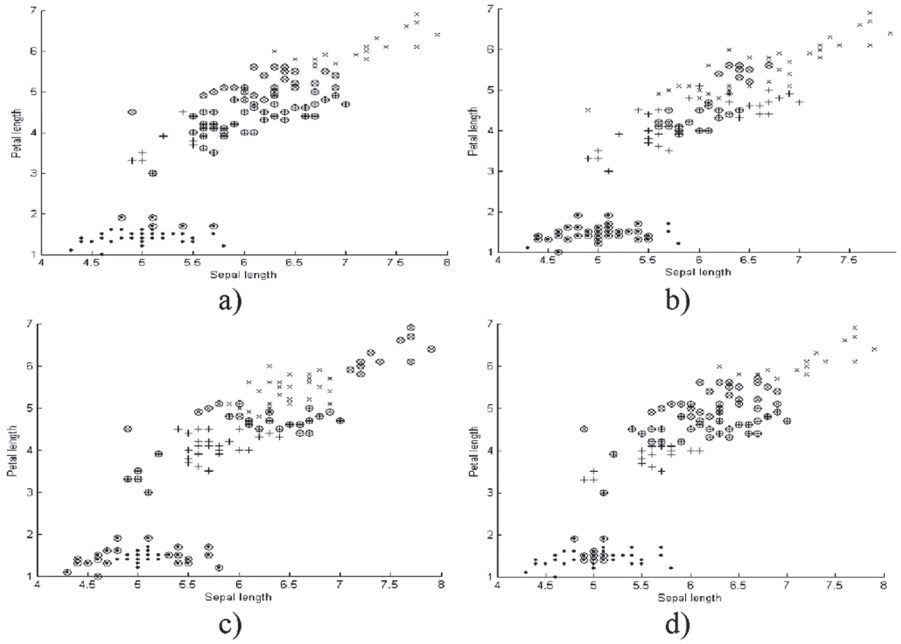


Fig. 3. IRIS problem illustrations for instance selection: a) for the indicator $I_{\overline{G}}^{s}$, b) for the indicator I_{O}^{s}, c) for the indicator I_{U}^{s}, d) for the indicator $\hat{I}^{s}$

7 Practical exercises

To study practically the methods presented in this chapter students need to execute the following steps.

1. Specify the practical problem(s) and take the data sample(s) for the experiments. Students may download data from the repository https://archive.ics.uci.edu/ml/.
2. Take software package or library containing implementation of intelligent stochastic search methods. Then make modified procedures realizing feature selection methods presented in this chapter. Generally, for this we need to compute individual feature informativities and to modify search operators. As an alternative, students may also make the full own program implementation of presented methods too, but it will require a lot of time for this task.

3. For each given data sample run each method of feature selection to obtain subset of selected features. To compare feature selection methods for each task students need to measure the time of each method's work, the number of executed iterations, the maximum used memory volume. Then for each selected feature subset you need to build a model and evaluate its performance (error).
4. Rank the feature selection methods according to the minimum error of a model, the minimum number of selected features, the minimum time and iterations for feature selection, the minimum computer memory volume used for the feature selection. Compare obtained results with presented in the Table 2. Try to explain if they are different from the Table 2.
5. To study sample selection methods make own program implementation of presented method and use it for different tasks. Collect during experiments the time and number of iterations of method, the volume of used computer memory. Compare obtained results with presented at Fig. 1.
6. To study instance importance indicators make own program implementation of indicators presented before and use it for different tasks to compute instance importance indicators. Select some part of original sample instances with best instance importance indicator values. As an alternative select randomly the same number of instances from the original sample. Built the models for both subsamples and for whole original sample. Test all models on the whole original sample also as on subsamples. Compute model errors. Rank samples according the model error. Try to explain obtained results.
7. Combine the best feature selection method and best sample selection method and use them for data sample dimensionality reduction. Try to use different combinations of other feature and sample selection methods. Compare results of data dimensionality reduction. Try to explain obtained results.

8 Conclusion

The data dimensionality reduction problem for a model synthesis on instances is addressed. The set of methods for data reduction have been presented. It includes the intelligent methods of stochastic search applied for feature selection and sample selection problems. The theoretical estimations of a time and space requirements for the feature selection methods is provided. The experimental investigation of these methods has been provided. The set of presented methods allow selecting subsets of most informative features and instances allowing reducing time for model building.

9 References

1. Dash M., Liu H.: Feature selection for classification. Intelligent data analysis, 1,. 131–156 (1997)
2. Jensen R., Shen Q.: Computational intelligence and feature selection: rough and fuzzy approaches. John Wiley & Sons, Hoboken (2008)

3. Subbotin S.A.: Methods of sampling based on exhaustive and evolutionary search. Automatic Control and Computer Sciences, 3(47), 113-121 (2013)
4. Engelbrecht A.: Computational intelligence: an introduction. Wiley, Sidney (2007)
5. Ruan D.: Intelligent hybrid systems: fuzzy logic, neural networks, and genetic algorithms. Springer, Berlin (2012)
6. Liu Y., Passino K. M.: Biomimicry of social foraging bacteria for distributed optimization: models, principles, and emergent behaviors. Journal of optimization theory and applications, 3, .603–628 (2002)
7. Karaboga D., Akay B.: A survey: algorithms simulating bee swarm intelligence. Artificial intelligence review, 31, 61–85 (2009)
8. Subbotin S., Oleynik A.: Entropy based evolutionary search for feature selection. The Experience of Designing and Application of CAD Systems in Microelectronics. Proc. 9th Int. Conf. (CADSM-2007), pp. 442-443. IEEE Press, Lviv (2007)
9. Subbotin S., Oleynik A.: Modifications of ant colony optimization method for feature selection. The Experience of Designing and Application of CAD Systems in Microelectronics. Proc. 9th Int. Conf. (CADSM-2007), pp. 493-494. IEEE Press, Lviv (2007)
10. Oliinyk A.O., Oliinyk O.O., Subbotin S.A.: Agent technologies for feature selection. Cybernetics and Systems Analysis ,2 (48), 257-267 (2012)
11. Lavrakas P. J.: Encyclopedia of survey research methods. Sage Publications, Thousand Oaks (2008)
12. Hansen M.H., Hurtz W. N., Madow W. G.: Sample survey methods and theory. John Wiley & Sons, New York (1953)
13. Bernard H. R.: Social research methods: qualitative and quantative approaches. Sage Publications, Thousand Oaks: (2006)
14. Subbotin, S. The instance and feature selection for neural network based diagnosis of chronic obstructive bronchitis Studies in Computational Intelligence 2015
15. Ghosh S.: Multivariate analysis, design of experiments, and survey sampling. Marcel Dekker Inc., New York (1999)
16. Subbotin S.A.: The training set quality measures for neural network learning. Optical Memory and Neural Networks (Information Optics), 2(19), 126-139 (2010)
17. Smith G.: A deterministic approach to partitioning neural network training data for the classification problem : dissertation. Virginia Polytechnic Institute & State University, Blacksburg (2006)
18. Subbotin, S.A.: The sample properties evaluation for pattern recognition and intelligent diagnosis. DT 2014 - 10th International Conference on Digital Technologies, pp. 321 - 332, IEEE, Zilina (2014)
19. Plutowski M.: Selecting training exemplars for neural network learning : dissertation. University of California, San Diego (1994)
20. Oliinyk, A., Zaiko, T., Subbotin, S.: Training sample reduction based on association rules for neuro-fuzzy networks synthesis. Optical Memory and Neural Networks (Information Optics), 2 (23), 89-95 (2014)
21. Chaudhuri A., Stenger H.: Survey sampling theory and methods. Chapman & Hall, New York (2005)
22. Subbotin S., Oliinyk A.: The dimensionality reduction methods based on computational intelligence in problems of object classification and diagnosis. Recent Advances in Systems, Control and Information Technology. Advances in Intelligent Systems and Computing, vol 543, pp, 11–19. Springer, (2017).

23. Subbotin S., Oliinyk A.: The sample and instance selection for data dimensionality reduction. Recent Advances in Systems, Control and Information Technology. Advances in Intelligent Systems and Computing, vol 543, pp, 97-103. Springer, (2017).
24. Iris Data Set. https://archive.ics.uci.edu/ml/datasets/Iris

Digital signal processing of ECG and PCG signals

David Luengo[1][0000-0001-7407-3630], David Osés[1][0000-0002-8147-1001] and Tom Trigano[2][0000-0002-9912-1499]

[1] Universidad Politécnica de Madrid, 28031 Madrid, Spain
[2] Shamoon College of Engineering, Ashdod, Israel

Abstract. Electrocardiograms (ECGs) were the first biomedical signals where digital signal processing (DSP) techniques were extensively used to remove noise and artifacts, detect the relevant waveforms (e.g., the QRS complexes) and perform waveform segmentation, classify the signals according to their pathological or normal condition, and achieve waveform compression for storage and transmission of long signals with the desired quality. Phonocardiograms (PCGs) are related acoustic signals that provide complementary information about the activity of the heart's valves and where DSP techniques have started to be applied more recently. This chapter provides a brief overview on the acquisition of ECGs and PCGs and the use of DSP techniques in this field. The required background to work in this field, the main challenges that had to be overcome since the beginning, and some of the remaining open issues will also be explored.

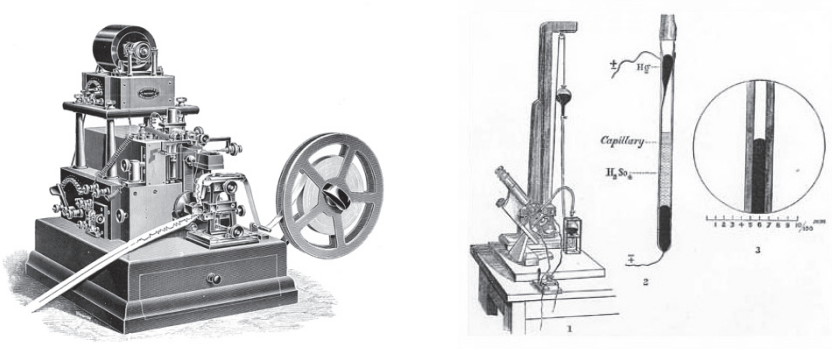
Keywords: electrocardiogram (ECG), phonocardiogram (PCG), digital signal processing.

1 ECG signal processing

1.1 ECGs: acquisition, characteristics and signal processing tasks

The first recordings of the electrical activity of the human heart seem to have been attained by Alexander Muirhead in 1869, using a Thomson siphon recorder designed to receive transatlantic communications signals (Fig. 1a), and Gabriel Lippmann in 1873, using a capillary electrometer (Fig. 1b) [1, 2]. A few years later, electrocardiograms (ECGs) were first recorded in humans by Augustus D. Waller in 1887 using Lippmann's capillary electrometer [3]. However, the quality of the signals at that time was quite poor, and Waller initially failed to recognize the diagnostic importance of the ECG. Soon afterwards, Willem Einthoven substantially improved the recording mechanism, by introducing the so-called string galvanometer, and reported the first reliable recordings of human ECGs in a clinically applicable fashion in 1902 [3, 4]. Einthoven also named the main deflections in the ECG (the so-called P and T waveforms, and the

QRS complex) [5], introduced the tri-axial bipolar system with three derivations (standard leads I, II and III) which are still used nowadays, and showed that several cardiovascular disorders were related to specific features in the ECG [2, 3, 5].



(a) (b)

Fig. 1. (a) Muirhead's Thomson syphon recorder; (b) Lippmann's capillary electrometer. *Source: Wikimedia Commons*.

Einthoven's electrocardiograph required the subjects to have their arms and the left limb immersed in buckets of salt water, weighted 600 pounds, occupied 2 rooms and required 5 operators [1, 2, 3]. Soon afterwards, the English Cambridge Scientific Instrument Company started developing smaller marketable electrocardiographs (see Fig. 2), which only weighed 50 pounds [2]. In 1932, compact and portable direct writing ECG machines had been developed by Duchosal in Switzerland [1], and by 1935, the Sanborn Company had reduced the weight to just 25 pounds [2]. Contact electrodes appeared in the 1920s, and suction electrodes, which allowed the acquisition of the chest leads, appeared in the 1930s [2]. Altogether, these improvements contributed to the widespread use of the ECG to diagnose many cardiac pathologies.

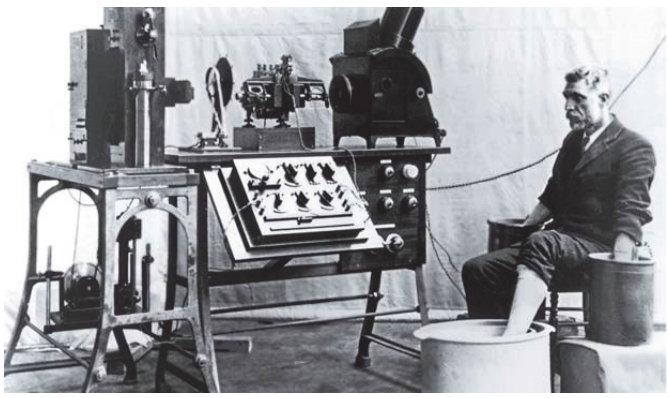
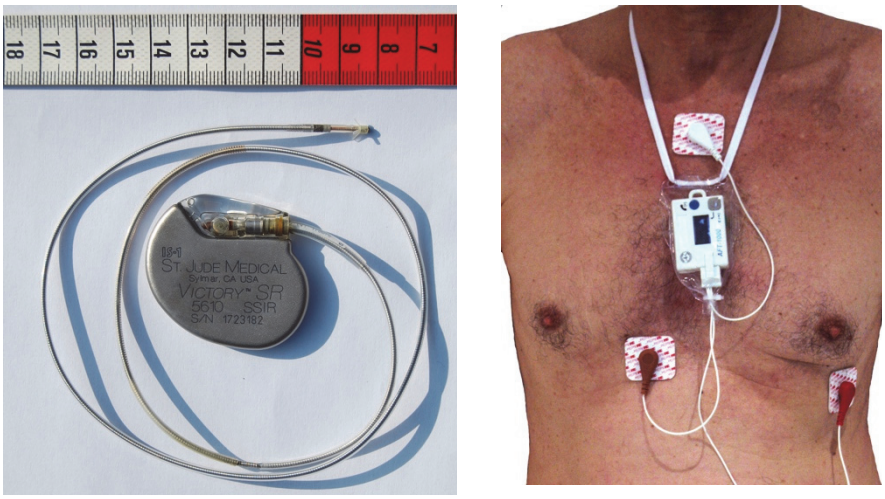


Fig. 2. One of the first commercial electrocardiographs, built in 1911 by the Cambridge Scientific Instrument Company, following the design of Einthoven's 1901 electrocardiograph. *Source: Wikimedia Commons*.

Since then, many advances have occurred in the field of cardiac monitoring. In 1958, the first artificial pacemaker (see Fig. 3a) was implanted in a human patient at the Karolinske Institute in Sweden [6, 7, 8]. Much later, in 1980, the first implantable cardioverter-defibrillator was implanted at John Hopkins Hospital [8]. Since then, substantial improvements have been introduced in both devices, reducing their size, increasing their reliability, and extending their battery lifetime. Single-chamber leadless pacemakers, which promise to reduce the number of long-term complications, have been recently introduced [9, 10]. Another milestone was the commercial release, in 1962, of the Holter monitor (see Fig. 3b), which allowed cardiologists to record the ECG in ambulatory settings [8, 11]. This marked the beginning of telemetric cardiac monitoring, and paved the way for other recent developments, leading to ubiquitous cardiac monitoring using wearable devices and wireless technologies [12, 13, 14].



(a) (b)

Fig. 3. (a) Implantable artificial pacemaker from St. Jude Medical. The body of the device is about 4 centimeters long, and the electrode measures roughly 58 centimeters. (b) Holter monitor with 3 leads. *Source: Wikimedia Commons*.

Nowadays, multi-channel ECGs are still the main tool used by physicians to diagnose and monitor many cardiac disorders [15, 16, 17], whereas single-channel ECGs are also routinely acquired and exploited in other non-cardiac, health-related contexts, like clinical EEG recordings [18, 19] or the monitoring of patients with sleep disorders [20]. The ECG provides a glimpse of the electrical activity occurring within the heart: an electrical impulse originates at the so-called sino-atrial (SA) node in the left atrium, propagating initially through both atria and then through the ventricles, which are electrically connected to the atria through the atrio-ventricular (AV) node. This electrical

signal is responsible for the coordinated contraction of the atria first (thus expelling the blood received from the body to the lungs for oxygenation) and then the contraction of the ventricles (thus pumping the oxygenated blood to the rest of the body). All this electrical activity is reflected in the main peaks that can be seen in the ECG (see Fig. 4): the P waveform (associated to atrial depolarization), the QRS complex (corresponding to ventricular depolarization), and the T waveform (associated to ventricular repolarization).

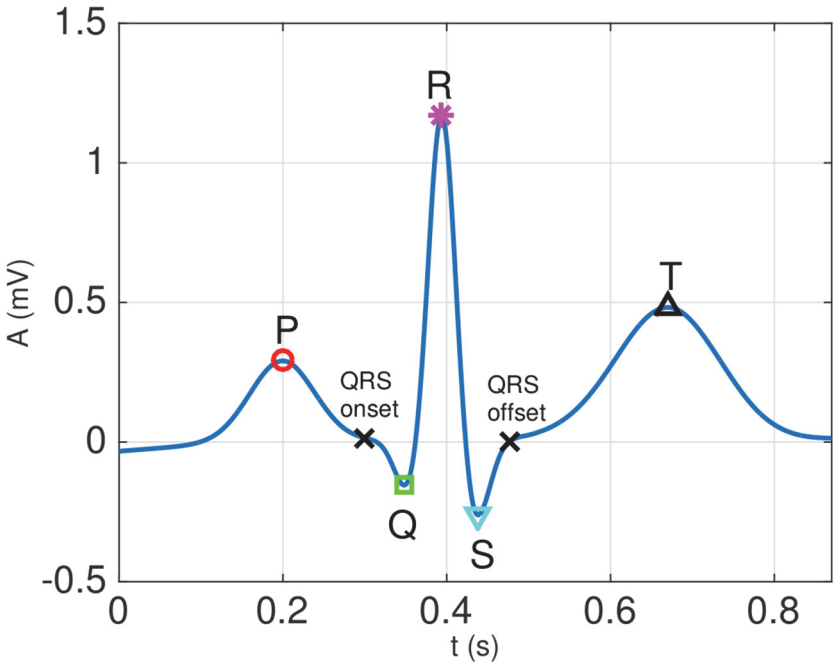


Fig. 4. Example of a synthetic ECG, generated in Matlab using the ECGSYN waveform generator [21], downloaded from Physionet [22], with the following command: `[s, ipeaks] = ecgsyn(1000,60,0,60,1,0.5,1000);`

As previously stated, many cardiac pathologies can be detected by examining the ECG. However, the widespread use of ECG recordings relies heavily on the development of signal processing and machine learning tools for denoising and interference suppression [23, 24, 25], feature extraction [26, 27], and computer-aided diagnosis of diseases [28, 29, 30]. Fig. 5 shows the typical block diagram used to perform ECG signal processing for pathology diagnosis. Usually, the first stage is applying one or several filtering techniques (both linear and non-linear) to **remove noise and interferences**, since the ECG is a weak signal (in the mV range) which is affected by many types of distortions. Then, the repetitive nature of the ECG is exploited: the most prominent waveform (the R peak) is detected, the nearby waveforms are located, and their boundaries estimated (**waveform delineation**). This allows us to extract the relevant information (**feature extraction**), like the morphology, amplitudes and durations of the

different waveforms, as well as the duration of several key intervals (e.g., the ST segment). Finally, the extracted features are fed to a **classifier** that will provide a categorization of the patient's status. Additionally, **ECG compression** may be required in order to reduce the signal size, while preserving its diagnostic power, before storing or transmitting the acquired ECG.

Note that two types of ECGs can be found in the literature. On the one hand, the external ECG has traditionally been obtained non-invasively by attaching a set of electrodes (normally 2-12) to the chest and limbs of the patient [23, 24], although alternative wearable [13, 14] and even wireless devices to acquire the ECG signal have been recently developed [31]. Indeed, low-cost devices that can be easily used in an academic environment to acquire multiple biomedical signals (including ECG signals) with high quality have also been released in the last few years [32, 33, 34]. On the other hand, intracavitary or intracardiac ECGs (a.k.a. electrograms (EGMs)) are obtained by setting one or more (typically 10-20, but can be as high as 64 using basket catheters) electrodes in direct contact with the inner surface of the heart. EGMs are used in implantable devices (e.g., pacemakers and defibrillators) for the identification of ventricular tachycardias, early alert of the presence of acute myocardial infarction, arrhythmia classification for appropriate therapy delivery, etc. They are also acquired during heart surgery performed on patients with sustained atrial fibrillation (AF), which is one of the most common heart disorders, to guide catheter ablation for patients not responding to drug therapies [35].

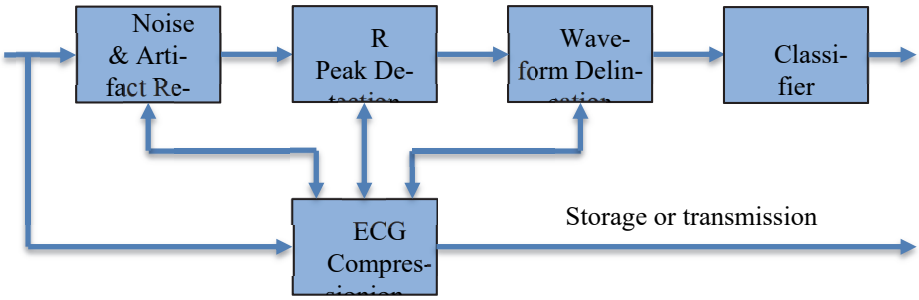


Fig. 5. Typical block diagram used to perform ECG signal processing for pathology diagnosis (adapted from [21]).

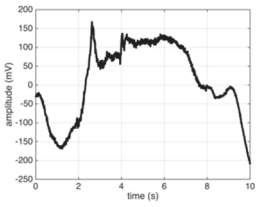
In the following section, we provide an overview of some of the main techniques used for the first two tasks shown in Fig. 5: denoising and waveform delineation (which requires performing R peak detection first). Although EGMs are increasingly used in clinical practice, we will focus on external ECGs, since they are still much more widely used on a daily basis.

1.2 ECG denoising

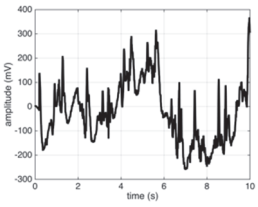
Since ECG signals are non-stationary by nature, noise removal is often a challenging task. Moreover, it is important to notice that all the different noises stem from various

sources and are of different nature [23]. However, some of the developed methods are based on standard techniques described in courses appearing at B.Sc. and M.Sc. levels in most engineering curricula, and can be implemented by students as part of standard assignments in these courses. In this section, we review the most common noise types present in ECG recordings, as well as the techniques used for denoising.

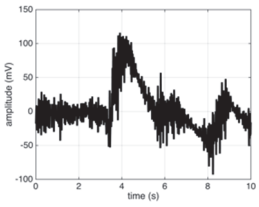
As mentioned for example in [36], we can separate the different noises appearing in ECG recordings between low-frequency noises and high-frequency noises. Among them, we can distinguish 4 types of common noises of interest: baseline wander, electrode noise, electromyographic (EMG) noise and power line interference. **Baseline wander** is a low-frequency perturbation. It is characterized by a slowly-varying signal, on which the actual ECG signal adds up. Partly caused by the respiration and slight moves, it makes the localization of the QRS complexes more difficult. Fortunately, its frequency range generally does not overlap with the P-QRS-T information we wish to retrieve, and is usually below 1 Hz. **Electrode noise** is of different nature. External ECGs are recorded by means of electrodes in direct contact with the skin, creating a closed electrical circuit. During the measurements, any movement leads to a change in the electrode-skin impedance, and introduces additional perturbation, hence the name of electrode noise. This type of noise is more problematic than baseline wander, since its frequency range (up to 10 Hz) overlaps with the spectrum of one of the waveforms of interest. **Electromyographic noise** appears often during stress-tests, when the patient must run on a treadmill with increasing speed until exhaustion. When doing so, muscles located beneath the electrodes contract more often and stronglier than usual, and the electrodes record this as well. This type of noise is called EMG noise. It has a high frequency content, which is mostly higher than the spectral content of interest, but may overlap with the high frequencies of QRS complexes. Eventually, improper placement of the reference electrode and other high-power electrical equipments may also generate a 50 Hz / 60 Hz (depending on the electrical standards) perturbation in the recorded signal. This perturbation, mainly caused by the electrical network, is called **power line interference**. Note that non-portable ECG recording devices are sensitive to any perturbations coming from the electrical network: the presence of additional harmonics and overshoots also have an influence on the power line interference. Fig. 6 shows three examples of noises and interferences (both in time and frequency) taken from Physionet [20], whereas Fig. 7 shows an example of an ECG signal, exhibiting both baseline wander and electrode noise on its left part.



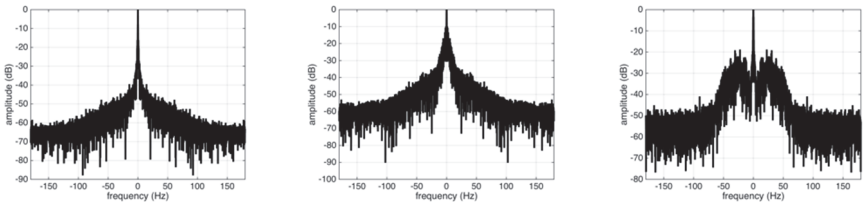
(a) Baseline wander (time)



(b) Electrode motion (time)



(c) Muscle artifacts (time)



(d) Baseline wander (freq.) (e) Electrode motion (freq.) (f) Muscle artifacts (freq.)

Fig. 6. Three examples of noises and interferences (both in time and frequency) taken from Physionet's MIT-BIH stress test data base [22, 37].

As we might infer, the different techniques available for denoising are often noise-dependent. However, we may classify them by their domain of consideration (time or frequency), their flexibility (adaptive/non-adaptive, user-defined or data-driven) and the level of investment required from the students to understand and successfully implement them.

The easiest way to remove noises from ECG recordings is by the use of **linear, time-invariant (LTI) filters** [38]. This line of denoising methods is usually the easiest to implement in a digital framework. For example, baseline wander noise can be easily removed (at least partly) by means of a high-pass filter, whose cutoff frequency is set based on physiological knowledge: considering that the lowest heart rate during bradychardia can drop to 40 beats/minute, a standard choice for the cutoff frequency is $f_c = 0.5$ Hz, as detailed in [39]. Practical implementations include both Finite Impulse Response (FIR) and Infinite Impulse Response (IIR) filters, and a variety of methods are accessible either for undergraduate students in Electrical Engineering, or graduates in other engineering fields [40, 41, 42]. The other important consideration is related to the properties of the group delay response and, consequently, the choice of filter structure. Indeed, a constant group delay (i.e., a linear phase [42]) is highly desirable in order to prevent phase distortion from altering various wave properties of the cardiac cycle, such as the duration of the QRS complex, as shown in [23, 40]. IIR filters, on the one hand, can attain a good selectivity with a low filter's order, but usually have a non-constant group delay, and therefore need to be designed using the zero-phase (forward-backward) filtering idea [43]. This prevents their use for online solutions, but can be useful for offline processing. On the other hand, a constant group delay can be easily attained by designing an FIR filter with a symmetric or antisymmetric impulse response. The price to pay for it is a filter with a (much) higher order than its IIR counterpart. Detailed examples of LTI filters for baseline wander can be found in [36].

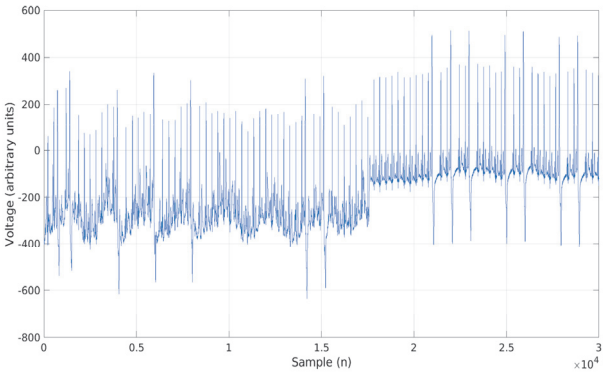


Fig. 7. Example of noisy ECG recording with baseline wander and electrode noise (from Physiobank's MIT-BIH database - [22, 44]).

As aforementioned, the topics related to ECG denoising by means of LTI filters do not necessitate additional knowledge besides what is commonly taught at B.Sc. levels in Electrical Engineering. As a consequence, they make perfect entry level assignments or semester projects in digital signal processing courses, and can be presented to undergraduate students as possible applications. Additional multirate filtering techniques exist for mid-level topics of a DSP undergraduate course, as seen in [22]. A comparison of the performances of different LTI digital filters is provided in [36].

The main advantage of LTI filters is their simplicity, when compared to other denoising techniques. However, the filters' parameters are difficult to tune and to adapt for each kind of measurement. As discussed in Section 1.1, new ambulatory cardiac monitors have been developed for continuous ECG monitoring in recent years. These devices are portable and have an autonomy that is increasing with the advance of low-power micro-electronics. Integration of microprocessors allows performing some embedded signal processing and automatic interpretation. In ambulatory conditions, noise increases with the higher levels of activity of the subjects. **Adaptive filtering** solutions are based on a filter whose impulse response's coefficients vary with time, with the objective of keeping an error criterion minimal. This is particularly useful when the noise is nonstationary, as it is the case with ambulatory motion artifacts. However, a reference signal has to be recorded in addition to the ECG. For power line interference, it is not needed, since the morphology and frequency of the perturbation is known. For muscle noise, several solutions have been proposed, as shown in [45, 46, 47]. Similarly, adaptive nonlinear filtering and extended Kalman filtering have been suggested as an adaptive way to remove power line interference [48, 49, 50].

Since adaptive signal processing is a more advanced topic in the theory of signal processing than standard LTI filtering, the description of these techniques is typically out of the scope of basic DSP courses taken at undergraduate level. However, it is possible to present them, either as part of an undergraduate elective course dedicated to adaptive signal processing, either at graduate level inside a course describing advanced

topics of signal processing. Comparing the performance of LTI and adaptive filters for the removal of different types of noises and artifacts can be a good topic for an M.Sc. thesis [51, 52]. Finally, note that the implementation can be performed, for example, using Bitalino cards [32, 33, 34], which can be equipped with electrodes for ECG measurements, as well as with an accelerometer and/or EMG electrodes.

Another group of ECG denoising methods can be defined as **regression-based methods**. A regression-based method will consider an ECG lead as a noisy curve, and will try to estimate it accordingly. Numerous regression methods exist, whether they be parametric (depending on a basis of known functions) or nonparametric (fully data-driven), see for example [53]. However, since denoising consists in removing parts of the curve without impinging the physiological characteristics of the ECG, most of the regression-based methods developed rely on basis functions. Baseline removal using regression techniques is indeed easy, since it is slowly varying noise when compared to ECG pulses: for example, [54, 55] address it by estimating the baseline with cubic splines, and subtract the obtained estimator from the original signal. Similarly, the technique known as estimation-subtraction method for power line interference removal [23] can be seen as a regression-based method, using the cosines and sines of 50 Hz as basic functions. This method can be used to remove power line interference efficiently, as shown in [56, 57]. More recently, wavelets bases have been introduced to represent ECG signals, and wavelet thresholding methods have been successfully implemented to remove noise from ECGs, as seen for example in [58, 59]. There exists a large literature on signal denoising using wavelets, and we refer the reader to [60] for additional insights on wavelets decomposition and wavelet-based denoising techniques. One of the main advantages of wavelets bases for regression-based methods is their flexibility. Indeed, since wavelets allow for a multiresolution decomposition of the signal at hand, specific levels of the decomposition provide insights on specific types of noises. However, one of the sensitive points in this approach is the type of wavelet used, as well as the thresholding strategy, which must be user-defined.

From the previous discussion, it appears that regression-based techniques are not a good fit to be learned at an undergraduate level, since the material involved in the understanding of these methods heavily relies on estimation theory, and on wavelets, which are more advanced level topics. Therefore, we recommend presenting these methods at a graduate level, and for students whose background in statistics and estimation theory is solid enough to handle the material. However, these methods can be implemented at an undergraduate level, in the framework of a final-year project for example, without putting too much emphasis on the theoretical material involved.

1.3 QRS detection and waveform delineation

After denoising the ECG signal, the next step usually corresponds to the extraction of the relevant P-QRS-T information. For this purpose, the first stage is usually the location of the most prominent waveform in the ECG: the R peak. After locating this peak, we can easily move backward and forward to locate the Q and S peaks, respectively. Then, the P and T waveforms are similarly located, but moving further away before the Q peak and the S peak, respectively. Note that these waveforms are smoother and typically have lower amplitude than the peaks in the QRS complex. Hence, their proper location is more difficult, since they are more sensitive to noise and artifacts. Finally, waveform delineation is performed: the onset and offset of all the aforementioned waveforms are estimated. This is the most difficult task in noisy ECGs, especially for the offset of the T wave, since it shows a very slow decrease towards the isoelectric line [23].

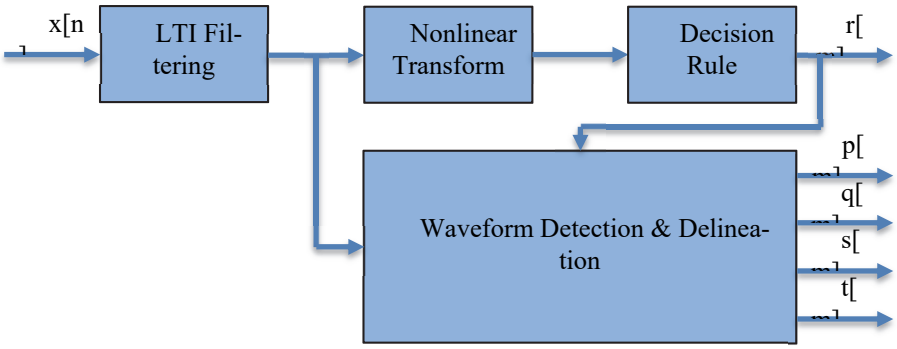


Fig. 8. Typical block diagram used to perform R peak detection and waveform delineation (adapted from [21]).

Fig. 8 in the previous page graphically summarizes the peak detection and waveform delineation process, emphasizing the blocks that perform the critical R peak detection stage:

- An **LTI filtering** stage (often composed of more than one filter), designed to preserve the QRS complexes and remove all the information not required for R peak detection: noise, artefacts and the other low-frequency waveforms (P and T waves).
- A **nonlinear transformation**, designed to maximize the desired signal to interference, noise and distortion ratio (SINAD). This transformation can be either a simple memoryless (e.g., squaring or taking the absolute value) operation or a more complicated transformation with memory.
- A **smoothing** operation (e.g., averaging inside a fixed time window), whose goal is to attain a single positive peak related to the R peak's position, so that the location of the R peak is simplified.
- A **decision rule**, designed to determine whether the detected peaks correspond to R peaks or not. Typically, the decision is taken on a peak-by-peak basis, but more complicated rules that involve several peaks can also be considered.

The first algorithms for R peak detection were developed during the late 1960s and the 1970s. These methods were based only on LTI filtering, without performing any non-linear transformation. Indeed, the first algorithms relied simply on a differentiation filter to emphasize the rapidly changing waveforms in the QRS complex [61, 62, 63]. Since the largest positive and negative slopes (i.e., the largest positive and negative peaks in the derivative) occur in the QR and RS segments, respectively, the zero crossing in the derivative between these two peaks should correspond to the R peak. Unfortunately, differentiation amplifies high-frequency noise, so it was soon realized that a better performance could be obtained by combining a differentiator with a low-pass filter [64, 65]. Note that many of these early QRS detectors are only designed to detect the heartbeats, and precise location of the R peak will require an additional time alignment stage [23].

As an illustrative example of this first category of methods, [64] assumes a sampling frequency $f_s = 250$ Hz, and uses a simple differentiator plus notch filter (with the notch located at $f_c = f_s/4 = 62.5$ Hz), $y[n] = x[n] - x[n-4]$, followed by a low-pass filter defined by the following differential equation:

$$z[n] = y[n] + 4y[n-1] + 6y[n-2] + 4y[n-3] + y[n-4].$$

The sequence $z[n]$ is used as the input of the classifier, which is based on two thresholds with equal magnitudes but opposite polarities ($\pm\gamma$). Firstly, the sequence $z[n]$ is scanned until finding $z[n] > \gamma$ at some n=n*, then a search is performed within the next 40 samples (160 ms), and the following heuristic rule is applied to determine whether the detected sample corresponds to an R peak or not:

- If no other threshold crossing occurs, then the sample is discarded, since it probably corresponds to a baseline shift.
- The sample is classified as belonging to a QRS complex if any of these conditions occur for $0 < j < k < m < 40$:
  - $z[n^* + j] < -\gamma$
  - $z[n^* + j] < -\gamma$ AND $z[n^* + k] > \gamma$
  - $z[n^* + j] < -\gamma$ AND $z[n^* + k] > \gamma$ AND $z[n^* + m] < -\gamma$
- If additional threshold crossings occur, then the sample is discarded, since it probably corresponds to high frequency noise.

This algorithm is implemented in the SQRS function of Physionet's WFDB toolbox for Matlab and Octave (https://archive.physionet.org/physiotools/matlab/wfdb-app-matlab/), and can be easily included as part of a lab. assignment in a B.Sc. digital signal processing or biomedical signal processing course (just like any other method in this class), as it only requires basic programming skills and knowledge of standard LTI filtering methods.

Then, in the late 1970s and the 1980s, more sophisticated algorithms that included memoryless non-linear transformations and smoothing appeared. These methods used either the absolute value or the square of the input samples, followed by a moving average integrator that uses a rectangular [66, 67] or a triangular [68, 69] window. The

well-known Pan-Tompkins algorithm [67, 70] belongs to this class. This method uses first a band-pass filter, implemented using a low-pass filter with a cut-off frequency $f_c = 12$ Hz,

$$y[n] = 2y[n-1] - y[n-2] + x[n] - 2x[n-6] + x[n-12],$$

and a high-pass filter with a cut-off frequency $f_c = 5$ Hz,

$$y[n] = -y[n-1] - x[n] + 32x[n-16] - x[n-32].$$

This is followed by a differentiator, which is defined by the following differential equation:

$$y[n] = \frac{1}{8T_s}(-x[n-2] - 2x[n-1] + 2x[n+1] + x[n+2]).$$

Then, a point by point squaring, $y[n] = (x[n])^2$, is performed to ensure a positive output and emphasize the higher frequencies associated to the QRS complex, and a moving window integration is applied:

$$y[n] = \frac{1}{N}\sum_{k=0}^{N-1} x[n-k]$$

The length of the integration window, $N = 30$ (i.e., 150 ms), is selected to be equal to the maximum expected length of the QRS complex, so that the three waveforms in the QRS complex are converted into a single peak at the output. Finally, the classifier is based on an adaptive threshold that must be crossed simultaneously at the filter's output and the integrator's output. An *eye-closing period* corresponding to the refractory period and *look-back detection mode* are also included in order to reduce the number of false alarms and missed detections, respectively. See [67, 70] for further details on the classifier.

A full implementation of the Pan-Tompkins algorithm is freely available at Matlab Central (https://www.mathworks.com/matlabcentral/fileexchange/45840-complete-pan-tompkins-implementation-ecg-qrs-detector) [71]. This algorithm is more challenging to program (due to the adaptive threshold, the eye-closing period and the look-back detection), but could be included as a final assignment for advanced B.Sc. or M.Sc. students. The available Matlab version can also be easily used out-of-the-box in lab. assignments or B.Sc./M.Sc. thesis which require QRS complex detection.

Methods based on nonlinear memoryless transformations and proper classification rules, like the Pan-Tompkins algorithm, already provide an excellent performance for QRS complex detection in clinical settings. However, their performance can still be improved in ambulatory contexts and most of the aforementioned methods do not address the precise determination of the ventricular onset and offset (i.e., the beginning and end of the QRS complex). More recently, other methods that use non-linear transformations with memory have been developed. As an example of these methods, [72]

applies first an LTI low-pass filter with a cut-off frequency $f_c \approx 16$ Hz (assuming a sampling frequency $f_s = 250$ Hz), described by the following differential equation:

$$y[n] = 2y[n-1] - y[n-2] + x[n] - 2x[n-5] + x[n-10].$$

Then it applies the discrete-time version of the so-called curve length transformation:

$$L_M[n] = \sum_{k=n-M}^{n} \sqrt{T_s^2 + (y[k] - y[k-1])^2}$$

Where $T_s = 1/f_s$ is the sampling period, and $y[k]$ is the output of the LTI low-pass filter. The sequence $L_M[n]$ is then fed to the block that implements the decision rule, which is based on adaptive thresholding and local search strategies, and returns the QRS onset and offset, as well as the R peak location. See [72] for further information.

This algorithm is implemented in the WQRS function of Physionet's WFDB toolbox for Matlab and Octave (https://archive.physionet.org/physiotools/matlab/wfdb-app-matlab/), and can be easily included as part of a lab. assignment in a B.Sc. digital signal processing or biomedical signal processing course.

Note that a huge number of QRS detectors have been developed in the last 50 years. A review of some of the early methods (with an emphasis in ambulatory settings) is provided in [73], whereas a more modern review (2002), which also includes some of the early machine learning learning approaches (e.g., artificial neural networks) that researchers have started applying in this area, can be found in [74].

Regarding waveform segmentation and delineation, their goal is locating all the other relevant peaks (P, Q, S and T) starting from the location of the R peak, and determining their boundaries (onset and offset). ECG segmentation/delineation is more challenging, due to the high variability in the morphology of the QRS complexes, the low amplitude of the P waveform, and the smooth transitions of the low-frequency waveforms (especially at the offset of the T waveform). Typically, segmentation/delineation approaches are based on a battery of LTI filters and transformations which are tailored to each of the specific waveforms to be detected. The interested reader can refer to [23, 75] for an account and comparison of several ECG segmentation/delineation methods.

Note that many of the methods described in this section can be included in lab. assignments in a biomedical signal processing course, and they are also good candidates for B.Sc./M.Sc. thesis that compare several QRS detection or ECG segmentation/delineation methods [76, 77] or as part of other thesis that require fiducial point extraction of ECG signals [78]. The methods in the first two classes simply require a background in basic digital signal processing, while the methods in the latter class may require knowledge of statistical signal processing and machine learning.

2 PCG signal processing

2.1 Introduction

In this section we present an introduction to the study of the heart's sounds. Auscultation with a stethoscope was the traditional way of performing a basic analysis of the heart's activity, just listening to the sounds, in order to detect possible cardiovascular pathologies. From a simple auscultation, a well-trained listener can anticipate abnormal cardiac conditions, for example: arrhythmias, murmurs, different valve diseases (including stenosis and prolapses), etc. So we can consider the study of heart sounds as an essential first step, previous to further diagnostics and heart examinations, to distinguish simple risk factors from more severe heart diseases.

According to the World Health Organization (WHO), cardiovascular diseases (CVD) are the first cause of death in the world, causing about 17.9 million deaths each year [79]. Most disorders of the heart, including coronary and rheumatic heart diseases, can be detected prematurely in a simple auscultation, so the study of the heart's sound is now playing a relevant role in CVD investigation.

The phonocardiogram or phonocardiograph is a recording of the heart's sound that allows us to apply digital signal processing techniques to get the most important features of the sounds. Analyzing the obtained features and applying automated segmentation and classification algorithms it is possible to detect and identify abnormalities and problems in the proper functioning of the heart, thus helping doctors in the diagnosis process [80]. Computer-aided auscultation, along with digital signal processing, is making way for countless improvements in the heart sounds analysis and can help us to identify the presence of diseases that are not detected in a traditional auscultation process, as murmurs related to aortic stenosis and other valve pathologies.

The most common methods for the automated classification are based on Neural Networks [81], Hidden Markov models [82], and Support Vector Machines [83], but in recent years new approaches based on Deep Learning and Artificial Intelligence are continuously being published [84, 85, 86], highlighting the processing of spectrograms. Given the large number of strategies proposed, some papers present reviews [87], in terms of characteristics and performance, of the published algorithms, scoring and ranking them.

2.2 The sounds of the heart

In order to get a clear idea of the problem of analyzing and processing PCG signals, it is very interesting, at the beginning, to spend some time listening to different records corresponding to all kinds of heart sounds, including noisy and pathological ones.

In some databases we can find relatively clean signals with a perfect sound for training health professionals in auscultation skills. It is very interesting to hear this signals in order to prepare our ears for the noisy ones. For example, in the online edition of the book "*Teaching Heart Auscultation to Health Professionals*" we have access to 15 sounds (pathological and non-pathological) in **mp3** format, obtained with a sampling

frequency (fs) of 11025 Hz [88]. If we want to listen to a normal heart, we can play the file labelled as "*adult-case-1*" to hear a very clear sound corresponding to a 22 year old male. In the same way, the University of Michigan has a well-illustrated small database named "*Heart Sound & Murmur Library*" [89], where we can find 23 sounds obtained in three different auscultation areas (apex, aortic and pulmonic), acquired using a sampling frequency fs = 44100 Hz. Other very interesting database can be found in the companion Web site of the book "*Auscultation skills: breath & heart sounds*" [90], where we can find more than 50 sounds of very good quality, recorded with fs = 48000 Hz. The records include voice comments at the beginning with useful information about the pathologies. By listening to the 50 recordings, students can obtain a very good understanding of the complexity of the scenario that we are working on.

In Fig. 8 we can see the representation of the four heart sounds in the time domain, altogether with the power spectrum of the whole signal. In general, only S1 and S2 are considered, because S3 and S4 are very difficult to hear in a normal auscultation. The succession of the S1-S2 cycles form the characteristic lub-dub sound of a heart beating.

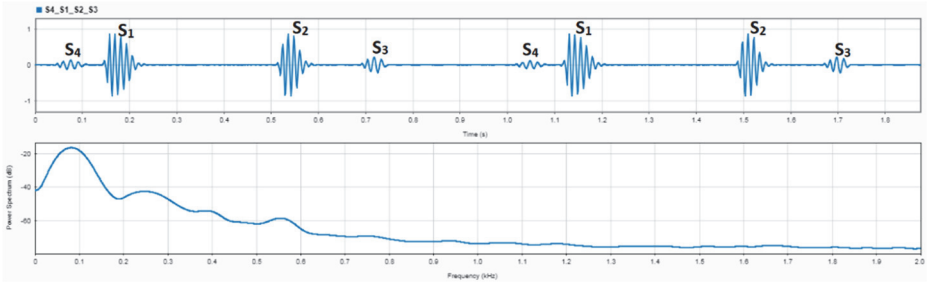


Fig. 8. Above: the four sounds of heart labelled with the corresponding denomination. **Below:** Typical power spectrum of the heart sounds.

In the sequel, we summarize the main characteristics of the PCG:

- *Average heart sound durations*: **S1**: [70–150] ms, **S2**: [60–120] ms, **S3**: [40–100] ms, and **S4**: [40–80] ms.
- *Cardiac cycle period*: 800 ms, ***systolic period***: 300 ms, and ***diastolic period***: 500 ms.
- *Frequency ranges*: **S1**: [50–150] Hz, **S2**: [50–200] Hz, **S3**: [50–90] Hz, and **S4**: [50–80] Hz.
- **S1** appears at the beginning of the systole, and it is associated with the vibrations produced by the closure of tricuspid and mitral valves.

- **S2** occurs at the beginning of the diastole, with the closure of pulmonic and aortic valves, and identifies the onset of ventricular diastole and the end of mechanical systole.
- **S3** occurs 120 ms to 180 ms after the onset of **S2** and is produced by a large amount of blood striking the left ventricle at high speed.
- **S4** occurs 90 ms before **S1** and corresponds to vibrations during the final atrial contraction to force blood into the left ventricle. **S3** and **S4** are low pitched and faint and may be pathological.

2.3 ECG and PCG

The electrocardiogram (ECG) and the phonocardiogram (PCG) are a pair of signals clearly synchronized: the electrical activation leads to the mechanical activity, that is, the opening and closing of heart valves and the contractions of the atria and ventricles in systole and diastole. Combining these two synchronous signals may improve the diagnosis of different pathologies, so a multimodal database including synchronized recordings of both will be perfect to understand the different heart sounds and identify them. The PhysioNet/Computing in Cardiology (CinC) Challenge 2016 includes five heart sounds databases with more than 3000 recordings, lasting from 5 seconds to just over 120 seconds [91, 92]. They were taken in different environments, both clinical (mainly hospitals) and non-clinical (home visits), and correspond to normal and abnormal sounds. The last ones can include different pathologies, such as heart valve defects (mitral valve prolapse, mitral regurgitation, aortic stenosis and valvular surgery) and coronary artery diseases. In ***challenge/2016/training-a*** we can find files of two different formats: ***.wav*** and ***.dat***. The first one only contains the heart sounds sampled at 2 kHz, while the latter format includes both ECG and PCG. It is very interesting to observe the relationship of the two signals (acoustic and electric), so we recommend using the ***.dat*** file to get the signals to process in this section. We can find a third type of files with the extension ***.hea*** containing additional information as, for example, whether the sounds correspond to a normal heart or are pathological (abnormal).

The normal trigger for the heart to beat arises in the SA (sinoatrial) node, the heart's natural pacemaker. The SA node sends out regular electrical impulses, causing the atria to contract and to pump blood into the ventricles. The electrical impulse then passes to the ventricles through the AV node (atrio-ventricular node). This electrical impulse spreads into the ventricles, causing the muscle to contract, pumping blood to the lungs and the body. Synchronized with this electrical activity, the different valves are opening and closing, allowing blood to fill and empty the atria and ventricles, and preventing it from flowing backwards during the isovolumetric contractions and relaxations.

S1 appears 10-50 ms after the R-peak in the ECG, and coincides with the beginning of ventricular systole. The first heart sound results mainly from the closure of the *mitral* and *tricuspid* valves. In Fig. 9, M1 denotes the closure of the mitral valve and T1 the closure of the tricuspid valve. After the closing of the atrioventricular valves, the open-

ing of the semilunar valves occurs: A1 corresponds to the aortic and P1 to the pulmonary. The sound for M1 is much louder than for T1 due to higher pressures in the left side of the heart. S1 precedes a palpable carotid pulse.

The second heart sound, S2, is produced, basically, by the closure of the *aortic* and *pulmonic* valves. A2 denotes the closure of the *aortic* valve, and P2 the closure of the *pulmonary* valve. In general, A2 sounds louder than P2. S2 indicates the beginning of isovolumetric relaxation of the ventricles and it is very close to the T wave of the ECG, that is, the electrical repolarization of the ventricles driving the mechanical relaxation. In normal conditions, the S2 sound onset matches in time the T-wave offset and appears 280-360 ms after the R-peak. Finally, M2 and T2 indicate the opening of mitral and tricuspid valves, respectively.

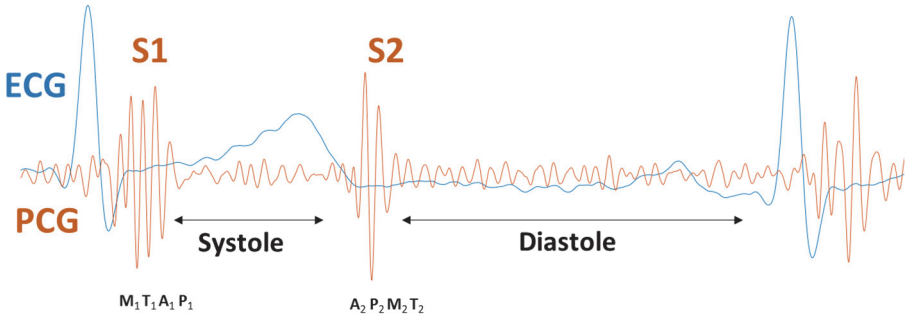


Fig. 9. Joint representation of ECG and PCG signals denoting the relationship between the electrical and the mechanical activity of the heart.

2.4 Automatic localization of heart beats

Over the years, a large number of methods have been proposed for the automatic analysis and segmentation of the heart's sounds, working both in the time and frequency domains. In what follows, we describe a simple approach to the detection of the sounds S1 and S2, which are the most audible and relevant for the segmentation of the PCG signal. To illustrate this point, we use some of the files available in *challenge/2016/training-a* [92]. The initial idea, to become familiar with the position of the sounds, is to take advantage of the ECG signal associated to the PCG. Both signals are provided in the records of the database. If we have the reference of the R peak, the sound S1 must be the first one to appear in the PCG, and it will be relatively easy to identify. The Pan–Tompkins algorithm is a well-known and widely used QRS complex detector in the ECG (see Section 1.3), since it is simple and robust, and it works very well under power line noise and baseline wander artifacts, thanks to the LTI denoising filters and the adaptive thresholds implemented [67, 70].

In Fig. 10 we have represented the waveforms corresponding to the last 10 seconds of the PCG and the ECG included in the 'a0007' recordings of *training-a*. The black dashed lines indicate the position of the R peaks of the ECG detected by running the Pan-Tompkins algorithm in Matlab [71]. We can see that the algorithm has detected all

the R peaks (indicated with yellow lines) in the ECG despite the noisy conditions. Close to the yellow lines, the sound S1 can be clearly identified.

Fig. 10. Joint representation of PCG, ECG, and the position of the R-peaks detected using the Pan-Tompkins algorithm [67, 70].

In order to isolate and identify the S1 and S2 sounds, a typical approach consists in using the five steps process depicted in Fig. 11. To accomplish the initial task of locating the sounds S1 and S2 in the PCG, it is highly recommended to apply an initial preprocessing filtering stage to eliminate unwanted noises. Since the main spectrum of S1 and S2 lies around a central frequency of 110-120 Hz, a band-pass filter with cutoff frequencies at 80Hz and 150Hz will help to clean the original signal, facilitating the following processing steps.

Fig. 11. Basic PCG signal processing for the identification of the sounds S1 and S2.

To avoid the delay and phase distortion introduced by nonlinear phase filters, the output of the band-pass filter shown in Fig. 12 has been obtained implementing a zero-phase filtering process. For example, if we are working with MATLAB, the **filtfilt** function performs this task by processing the input data in both the forward and backwards directions [43].

Fig. 12. PCG signals before (green line) and after (purple line) the band-pass filter.

Once unwanted noises have been removed, the next step, as we see in Fig. 11, is to obtain the envelope of the filtered signal. The **envelogram** of $x(t)$ is defined as the magnitude of the analytic signal $y(t)$ according to the following expression [93]:

$$y(t) = x(t) + j\, x_H(t),$$

where $x_H(t)$ is the Hilbert transform of $x(t)$. In the time domain, the Hilbert transform is defined as the following convolution:

$$x_H(t) = x(t) * \frac{1}{\pi t}$$

In the frequency domain, the corresponding relationships are

$$X_H(\omega) = X(\omega)[-j \cdot Sig(\omega)]$$

$$Y(\omega) = X(\omega)[1 + Sig(\omega)]$$

The algorithm can be seen as a complex demodulation yielding a high resolution envelope. Hence, it is clear that we can obtain $y(t)$ as the inverse DFT of a single band spectrum of $x(t)$ containing only the values for $\omega \geq 0$. In Fig. 13 we can see the envelope obtained after computating the envelogram. The synchronized averaging of PCG

envelopes is commonly used to detect different systolic murmurs due to cardiovascular diseases and defects such as aortic stenosis and mitral regurgitation.

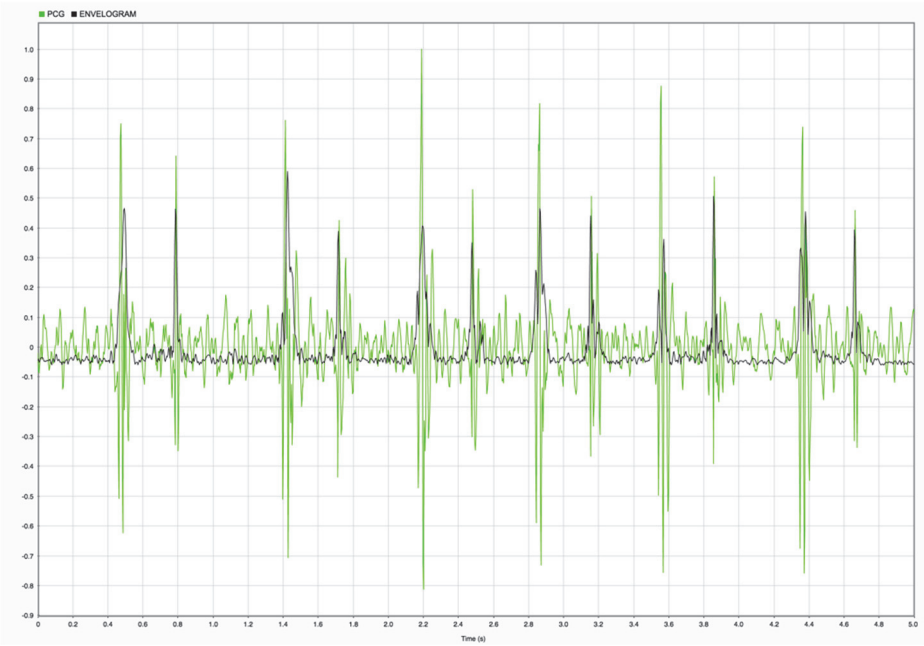


Fig. 13. Envelogram corresponding to the filtered PCG. The black line corresponds to the envelogram and the green one to the original PCG sounds.

Before proceeding to peak detection, we need a smoother signal to avoid mistakes with the maximum values that we can see close together in the envelope of the same sound. At this point, the empirical mode decomposition (EMD) can help us to extract the essential components [94] in order to determine the time location of the first and second heart sounds. EMD decomposes a signal into a set of simple oscillatory components called *Intrinsic Mode Functions* (IMFs). This functions satisfy two main properties: the number of local minima and maxima differs at most by one, and the average value is zero. Once the decomposition process is completed, the EMD algorithm produces a finite number IMFs and a residue, so the original signal can be written in the form:

$$x(t) = \sum_{i=1}^{L} h_i(t) + \ r(t)$$

Here, L corresponds to the number of IMFs extracted and i indicates the IMF's order. High frequencies (fast oscillations) are present in the low order modes (small values of i) while high order modes correspond to low frequencies (large values of i). On the other hand, $r(t)$ is the residual slow trend monotonic function.

Obviously, superimposing all the computed IMFs and the residual we obtain a perfect reconstruction of the original signal $x(t)$, but what we really want at this stage is a smoother signal to perform the peak detection easily. To obtain this new signal at the output of the EMD block of Fig. 11, we simply add a few of the central modes, avoiding IMFs with too high and too low frequencies. More precisely, modes 3 to 8 have been added to obtain the red line signal in Fig. 14. It is clear that we have used the EMD to carry out a filtering process that gives us a very clear signal to identify the sounds S1 and S2.

Fig. 14. The red line (emd_3_8) represents the smooth signal obtained after the addition of 3 to 8 IMFs of the EMD corresponding to the envelogram. The green line is the original PCG.

The last step is quite straightforward: to detect the peaks we can use a simple algorithm to look for local maxima points. The **findpeaks** function of MATLAB works perfectly.

Finally, the discrimination between S1 and S2 must be done. To this end, we have a wide range of strategies [95, 96], but for most signals S2 is closer to the previous S1 sound than to the following S1. The classification of the red peaks (S1) and the black peaks (S2) in Fig. 15 has been carried out according to this criterion of proximity in time. We can see that the peaks in this figure perfectly identify the sounds S1 and S2 of the original PCG signal.

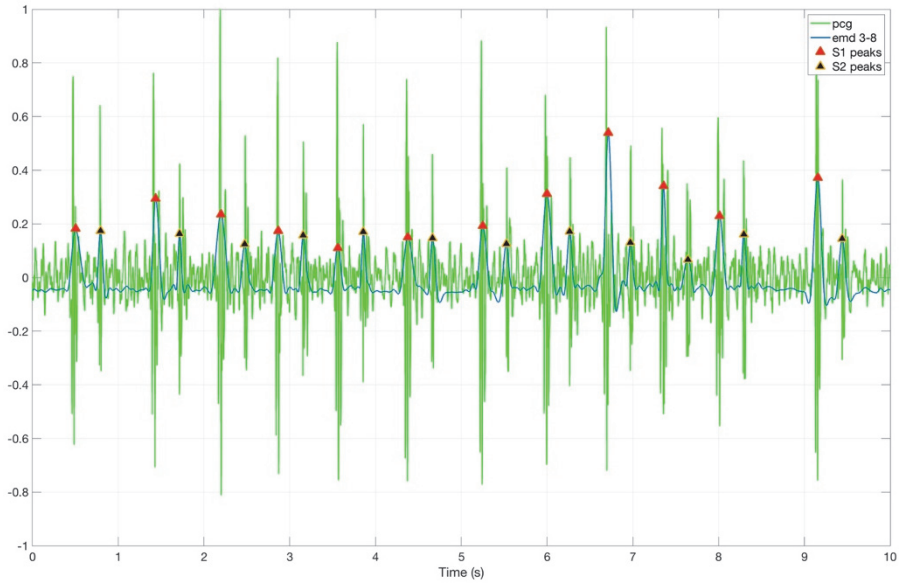


Fig. 15. Representation of the final peaks detected and classified with the algorithm of Fig. 11.

Note that the algorithm described in this point is just a simple first approach to the localization and classification of heart sounds, but the results are good enough as a starting point. Signals especially noisy and full of artifacts will require more complex signal processing techniques. Once more, the methods described in this section are ideal for lab. assignments in biomedical signal processing courses or B.Sc. thesis [97, 98], since they only require some programming expertise and basic signal processing knowledge.

3 Discussion

In this chapter, we have summarized first the main characteristics of the two main biomedical signals that are used to diagnose cardiac pathologies: the electrocardiogram (ECG) and the phonocardiogram (PCG). On the one hand, the ECG is the most widely used biomedical signal and should be a central part of any biomedical signal processing course. On the other hand, the PCG is also widely used, but digital signal processing techniques have only recently been applied to it. However, given its complementary nature with respect to the ECG, we believe that it is also a worthy signal that deserves a place in a biomedical signal processing course. In the case of the ECG, we have focused on the two tasks that must always be performed regardless of the application: denoising and QRS complex detection. Regarding the PCG, we have focused on the automatic location of heart beats and the identification of the S1-S2 sounds. In both cases, we have tried to emphasize the knowledge required by the students in order to perform lab. assignments or develop B.Sc./M.Sc. thesis in this field.

References

[1] E. M. M. Besterman, “The beginnings and development of the electrocardiograph.” *West Indian Medical Journal*, 54(3): 213-215, 2005.

[2] M. Rivera-Ruiz, C. Cajavilca and J. Varon, “Einthoven's string galvanometer: the first electrocardiograph.” *Texas Heart Institute Journal*, 35(2): 174-178, 2008.

[3] S. S. Barold, “Willem Einthoven and the birth of clinical electrocardiography a hundred years ago.” *Cardiac Electrophysiology Review*, 7(1): 99-104, 2003.

[4] P. Kligfield, “The centennial of the Einthoven electrocardiogram.” *Journal of Electrocardiology*, 35(4): 123-129, 2002.

[5] J. W. Hurst, “Naming of the waves in the ECG, with a brief account of their genesis.” *Circulation*, 98(18): 1937-1942, 1998.

[6] G. D. Nelson, “A brief history of cardiac pacing.” *Texas Heart Institute Journal*, 20(1): 12-18, 1993.

[7] K. Jeffrey and V. Parsonnet, “Cardiac pacing, 1960–1985: a quarter century of medical and industrial innovation.” *Circulation*, 97(19): 1978-1991, 1998.

[8] W. Greatbatch and C. F. Holmes, “History of implantable devices.” *IEEE Engineering in Medicine and Biology Magazine*, 10(3): 38-41, 1991.

[9] Tjong, F. V., Reddy, V. Y., “Permanent leadless cardiac pacemaker therapy: a comprehensive review.” *Circulation*, 135(15): 1458-1470, 2017.

[10] Lee, J. Z., Mulpuru, S. K., Shen, W. K., “Leadless pacemaker: Performance and complications.” *Trends in Cardiovascular Medicine*, 28(2): 130-141, 2018.

[11] B. Del Mar, “The history of clinical Holter monitoring.” *Annals of Noninvasive Electrocardiology: The Official Journal of the International Society for Holter and Noninvasive Electrocardiology*, 10(2): 226-230, 2005.

[12] S. Kumar, K. Kambhatla, F. Hu, M. Lifson and Y. Xiao, “Ubiquitous computing for remote cardiac patient monitoring: a survey.” *International Journal of Telemedicine and Applications*, 1-19, 2008.

[13] J. M. Pevnick, K. Birkeland, R. Zimmer, Y. Elad and I. Kedan, “Wearable technology for cardiology: an update and framework for the future.” *Trends in Cardiovascular Medicine* 28(2): 144-150, 2018.

[14] M. M. Baig, H. Gholamhosseini and M. J. Connolly, “A Comprehensive Survey of Wearable and Wireless ECG Monitoring Systems for Older Adults.” *Medical and Biological Engineering and Computing*, 51: 485-495, 2013.

[15] D. Corrado, A. Biffi, C. Basso, A. Pelliccia, and G. Thiene, “12-lead ECG in the athlete: physiological versus pathological abnormalities.” *British Journal of Sports Medicine*, 43(9): 669-676, 2009.

[16] K. Harris, D. Edwards and J. Mant, “How can we best detect atrial fibrillation?” *The Journal of the Royal College of Physicians of Edinburgh*, 42(Suppl 18): 5-22, 2012.

[17] J. Pirinen, J. Putaala, A. L. Aro, I. Surakka, A. Haapaniemi, M. Kaste, E. Haapaniemi, T. Tatlisumak and M. Lehto, “Resting 12-lead electrocardiogram reveals high-risk sources of cardioembolism in young adult ischemic stroke.” *International Journal of Cardiology*, 198: 196-200, 2015.

[18] J. N. Acharya, A. J. Hani, P. D. Thirumala and T. N. Tsuchida, “American clinical neurophysiology society guideline 3: a proposal for standard montages to be used in clinical EEG.” *The Neurodiagnostic Journal*, 56(4): 253-260, 2016.

[19] M. T. Kendirli, M. Aparci, N. Kendirli, H. Tekeli, M. Karaoglan, M. G. Senol and E. Togrol, “Diagnostic role of ECG recording simultaneously with EEG testing.” *Clinical EEG and Neuroscience*, 46(3): 214-217, 2015.

[20] O. Faust, U. R. Acharya, E. Y. K. Ng and H. Fujita, “A review of ECG-based diagnosis support systems for obstructive sleep apnea.” *Journal of Mechanics in Medicine and Biology*, 16(1): 1-25, 2016.

[21] P. E. McSharry, G. D. Clifford, L. Tarassenko and L. A. Smith, "A dynamical model for generating synthetic electrocardiogram signals." *IEEE Transactions on Biomedical Engineering*, 50(3): 289-294, 2003.
[22] Goldberger, A.L., Amaral, L.A., Glass, L., Hausdorff, J.M., Ivanov, P.C., Mark, R.G., Mietus, J.E., Moody, G.B., Peng, C.K., Stanley, H.E.: PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals. *Circulation*, 101: E215-220, 2000.
[23] L. Sörnmo and P. Laguna, *Bioelectrical Signal Processing in Cardiac and Neurological Applications*, Academic Press, 2005.
[24] Clifford, G. D., Azuaje, F., McSharry, P., *Advanced Methods and Tools for ECG Data Analysis*, Artech House, 2006.
[25] Sweeney, K. T., Ward, T. E., and McLoone, S. F., "Artifact removal in physiological signals—Practices and possibilities." *IEEE Transactions on Information Technology in Biomedicine*, 16(3): 488-500, 2012.
[26] Gacek, A., and Pedrycz, W. (Eds.), *ECG signal processing, classification and interpretation: a comprehensive framework of computational intelligence*. Springer Science & Business Media, 2011.
[27] Krishnan, S., and Athavale, Y., "Trends in biomedical signal feature extraction." *Biomedical Signal Processing and Control*, 43: 41-63, 2018.
[28] Goldberger, J. J., & Ng, J. (Eds.), *Practical signal and image processing in clinical cardiology*. Springer Science & Business Media, 2010.
[29] Hagiwara, Y., Fujita, H., Oh, S. L., Tan, J. H., San Tan, R., Ciaccio, E. J., & Acharya, U. R., "Computer-aided diagnosis of atrial fibrillation based on ECG signals: A review." *Information Sciences*, 467: 99-114, 2018.
[30] Schläpfer, J., & Wellens, H. J., "Computer-interpreted electrocardiograms: benefits and limitations." *Journal of the American College of Cardiology*, 70(9): 1183-1192, 2017.
[31] Majumder, S., Chen, L., Marinov, O., Chen, C. H., Mondal, T., and Deen, M. J, "Noncontact wearable wireless ECG systems for long-term monitoring." *IEEE Reviews in Biomedical Engineering*, 11: 306-321, 2018.
[32] da Silva, H. P., Fred, A., and Martins, R., "Biosignals for everyone." *IEEE Pervasive Computing*, 13(4): 64-71, 2014.
[33] da Silva, H. P., Carreiras, C., Lourenco, A., and Fred, A., "Off-the-person electrocardiography." *International Congress on Cardiovascular Technologies*, 2013.
[34] da Silva, H. P., Carreiras, C., Lourenço, A., Fred, A., das Neves, R. C., and Ferreira, R., "Off-the-person electrocardiography: performance assessment and clinical correlation." *Health and Technology*, 4(4): 309-318, 2015.
[35] Zanon, F., *et al.*, "Basic Properties And Clinical Applications Of The Intracardiac ECG." *Journal of Atrial Fibrillation*, 9(4): 1-7, 2017.
[36] Haritha, C., Ganesan, M., Sumesh, E.P., "A survey on modern trends in ECG noise removal techniques." *International Conference on Circuit, Power and Computing Technologies (ICCPCT)*, pp. 1–7, 2016.
[37] Moody GB, Muldrow WE, Mark RG., "A noise stress test for arrhythmia detectors." *Computers in Cardiology*, pp. 381-384, 1984.
[38] Oppenheim, A.V., Willsky, A.S., Hamid, S., *Signals and Systems*. Pearson, Upper Saddle River, N.J (1996).
[39] Bailey J J, Berson A S, Garson A, Horan L G, Macfarlane P W, Mortara D W, Zywietz C., "Recommendations for standardization and specifications in automated electrocardiography: bandwidth and digital signal processing. A report for health professionals by an ad hoc writing group of the Committee on Electrocardiography and Cardiac Electrophysiology of the Council on Clinical Cardiology, American Heart Association." *Circulation*, 81: 730–739, 1990.

[40] Manolakis, D.G., Ingle, V.K., *Applied Digital Signal Processing: Theory and Practice*. Cambridge University Press, New York, 2011.
[41] Proakis, J., Manolakis, D., *Digital Signal Processing*. Pearson, Upper Saddle River, NJ, 2006.
[42] Oppenheim, A., Schafer, R., *Discrete-Time Signal Processing.* Pearson, Upper Saddle River, 2009.
[43] Gustafsson, F., "Determining the initial states in forward-backward filtering." *IEEE Transactions on Signal Processing*, 44(4): 988-992, 1996.
[44] Moody, G.B., Mark, R.G.: The impact of the MIT-BIH Arrhythmia Database. *IEEE Engineering in Medicine and Biology Magazine*, 20: 45–50, 2001.
[45] Romero, I., Geng, D., Berset, T., "Adaptive filtering in ECG denoising: A comparative study." *Computing in Cardiology*, pp. 45–48, 2012.
[46] Raya, M.A.D., Sison, L.G., "Adaptive noise cancelling of motion artifact in stress ECG signals using accelerometer." *Second Joint 24th Annual Conference and the Annual Fall Meeting of the Biomedical Engineering Society*, vol. 2, pp. 1756–1757, 2002.
[47] Afonso, V.X., Tompkins, W.J., Nguyen, T.Q., Michler, K., Shen Luo, "Comparing stress ECG enhancement algorithms." *IEEE Engineering in Medicine and Biology Magazine*, 15: 37–44, 1996.
[48] Ziarani, A.K., Konrad, A., "A Nonlinear Adaptive Method of Elimination of Power Line Interference in ECG Signals." *IEEE Transactions on Biomedical Engineering*, 49: 540–547, 2002.
[49] Mneimneh, M.A., Yaz, E.E., Johnson, M.T., Povinelli, R.J., "An Adaptive Kalman Filter for Removing Baseline Wandering in ECG Signals." *Computers in Cardiology*, 33: 253–256, 2006.
[50] Avendano-Valencia, L.D., Avendano, L.E., Ferrero, J.M., Castellanos-Dominguez, G., "Improvement of an Extended Kalman Filter Power Line Interference Suppressor for ECG Signals." *Computers in Cardiology*, 34: 553–556, 2007.
[51] Mykhailo Romanov, "Comparison of powerline interference mitigation algorithms for ECG recordings", *M.Sc. Thesis*, Universidad Politécnica de Madrid, Oct. 2020.
[52] Vladyslava Tverdokhlib, "Evaluation of baseline wander removal approaches in ECG recordings", *M.Sc. Thesis*, Universidad Politécnica de Madrid, Oct. 2020.
[53] Silveman, B.W., Ramsay, J., *Functional Data Analysis*. Springer Series in Statistics, 2005.
[54] Meyer, C.R., Keiser, H.N., "Electrocardiogram baseline noise estimation and removal using cubic splines and state-space computation techniques." *Computers and Biomedical Research*, 10: 459–470, 1977.
[55] Badilini, F., Moss, A.J., Titlebaum, E.L., "Cubic Spline Baseline Estimation In Ambulatory ECG Recordings For The Measurement of ST Segment Displacements." *International Conference of the IEEE Engineering in Medicine and Biology Society*, vol. 13, pp. 584–585, 1991.
[56] Christov, I., "Dynamic Powerline Interference Subtraction from Biosignals." *Journal of Medical Engineering and Technology*, 24: 169–172, 2000.
[57] Levkov, C., Mihov, G., Ivanov, R., Daskalov, I., Christov, I., Dotsinsky, I., "Removal of Power-Line Interference from the ECG: a Review of the Subtraction Procedure." *BioMedical Engineering OnLine*, 4: 1–18, 2005.
[58] Sayadi, O., Shamsollahi, M.B., "Multiadaptive Bionic Wavelet Transform: Application to ECG Denoising and Baseline Wandering Reduction." *EURASIP Journal on Advances in Signal Processing*, 2007: 1–11, 2007.
[59] Mozaffary, B., Tinati, M.A., "ECG Baseline Wander Elimination using Wavelet Packets." *Proc. of World Academy of Science, Engineering and Technology*, 2005.
[60] Mallat, S., *A Wavelet Tour of Signal Processing: The Sparse Way*. Academic Press Inc, Boston, 2008.

[61] Pryor, T. A., Russell, R., Budkin, A., and Price, W. G., "Electrocardiographic interpretation by computer." *Computers and Biomedical Research*, 2(6): 537-548, 1969.
[62] Holsinger, W. P., Kempner, K. M., and Miller, M. H., "A QRS preprocessor based on digital differentiation." *IEEE Transactions on Biomedical Engineering*, 18(3): 212-217, 1971.
[63] Belforte, G., De Mori, R., and Ferraris, F., "A contribution to the automatic processing of electrocardiograms using syntactic methods." *IEEE Transactions on Biomedical Engineering*, 26(3): 125-136, 1979.
[64] Engelse, W. A. H., and Zeelenberg, C., "A single scan algorithm for QRS-detection and feature extraction." *Computers in Cardiology*, 6: 37-42, 1979.
[65] Usui, S., and Amidror, I., "Digital low-pass differentiation for biological signal processing." *IEEE Transactions on Biomedical Engineering*, 29(10): 686-693, 1982.
[66] Ligtenberg, A., and Kunt, M., "A robust-digital QRS-detection algorithm for arrhythmia monitoring." *Computers and Biomedical Research*, 16(3): 273-286, 1983.
[67] Pan, J., and Tompkins, W. J., "A real-time QRS detection algorithm." *IEEE Transactions on Biomedical Engineering*, 32(3): 230-236, 1985.
[68] Kunt, M., Rey, H., and Ligtenberg, A., "Preprocessing of electrocardiograms by digital techniques." *Signal Processing*, 4(2-3): 215-222, 1982.
[69] Murthy, I. S., and Rangaraj, M. R., "New concepts for PVC detection." *IEEE Transactions on Biomedical Engineering*, 26(7): 409-416, 1979.
[70] Hamilton, P. S., and Tompkins, W. J., "Quantitative investigation of QRS detection rules using the MIT/BIH arrhythmia database." *IEEE Transactions on Biomedical Engineering*, 33(12): 1157-1165, 1986.
[71] Sedghamiz, H., "Matlab Implementation of Pan Tompkins ECG QRS detector." *Research Gate*, March 2014, On-Line: https://www.researchgate.net/publication/313673153_Matlab_Implementation_of_Pan_Tompkins_ECG_QRS_detect
[72] Zong, W., Moody, G. B., and Jiang, D., "A robust open-source algorithm to detect onset and duration of QRS complexes." *Computers in Cardiology*, pp. 737-740, 2003.
[73] Pahlm, O., and Sörnmo, L., "Software QRS detection in ambulatory monitoring—a review". *Medical and Biological Engineering and Computing*, 22(4): 289-297, 1984.
[74] Kohler, B. U., Hennig, C., and Orglmeister, R., "The principles of software QRS detection." *IEEE Engineering in Medicine and Biology Magazine*, 21(1): 42-57, 2002.
[75] Beraza, I., and Romero, I., "Comparative study of algorithms for ECG segmentation." *Biomedical Signal Processing and Control*, 34: 166-173, 2017.
[76] Antía Testa Pérez, "Comparación de algoritmos de detección de complejos QRS", *B.Sc. Thesis*, Universidad Politécnica de Madrid, July 2016.
[77] Oran Lasri, "Comparative study of ECG segmentation algorithms", *B.Sc. Thesis*, Shamoon College of Engineering, 2017.
[78] Ana Garci-Nuño Verbo, "Desarrollo de un método fiduciario robusto para biometría basada en ECG", *B.Sc. Thesis*, Universidad Politécnica de Madrid, Dec. 2019.
[79] "Cardiovascular diseases (CVDs)", World Health Organization (WHO), On-Line: https://www.who.int/en/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds)
[80] Watrous, R. L., "Computer-aided auscultation of the heart: From anatomy and physiology to diagnostic decision support." *International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 140-143, 2006.
[81] Uğuz, H., "A biomedical system based on artificial neural network and principal component analysis for diagnosis of the heart valve diseases." *Journal of Medical Systems*, 36(1): 61-72, 2012.
[82] Schmidt, S. E., Holst-Hansen, C., Graff, C., Toft, E., and Struijk, J. J., "Segmentation of heart sound recordings by a duration-dependent hidden Markov model." *Physiological Measurement*, 31(4): 513-529, 2010.

[83] Ari, S., Hembram, K., and Saha, G., "Detection of cardiac abnormality from PCG signal using LMS based least square SVM classifier." *Expert Systems with Applications*, 37(12): 8019-8026, 2010.
[84] Raza, A., Mehmood, A., Ullah, S., Ahmad, M., Choi, G. S., and On, B. W., "Heartbeat sound signal classification using deep learning." *Sensors*, 19(21): 1-15, 2019.
[85] Krishnan, P. T., Balasubramanian, P., and Umapathy, S., "Automated heart sound classification system from unsegmented phonocardiogram (PCG) using deep neural network." *Physical and Engineering Sciences in Medicine*, 43: 505-515, 2020.
[86] Demir, F., Şengür, A., Bajaj, V., and Polat, K., "Towards the classification of heart sounds based on convolutional deep neural network." *Health Information Science and Systems*, 7(1): 1-9, 2019.
[87] Clifford, G. D., *et al.*, "Recent advances in heart sound analysis." *Physiological Measurement*, 38(8): E10-E25, 2017.
[88] John P. Finley, *Teaching heart auscultation to health professionals*, Canadian Pediatric Cardiology Association (CPCA), 2015. Companion Website (with data): https://teachingheartauscultation.com/HEART-SOUNDS-MP3-DOWNLOADS
[89] "Heart sound and murmur library", University of Michigan, On-Line: http://www.med.umich.edu/lrc/psb_open/html/repo/primer_heartsound/primer_heartsound.html
[90] Coviello, J. S., *Auscultation skills: breath & heart sounds*. Lippincott Williams & Wilkins, 2013.
[91] Liu, C., *et al.*, "An open access database for the evaluation of heart sound algorithms." *Physiological Measurement*, 37(12): 2181-2213.
[92] "Physionet/CinC Challenge 2016: Training Sets", On-Line: https://archive.physionet.org/pn3/challenge/2016/
[93] Sarkady, A. A., Clark, R. R., and Williams, R., "Computer analysis techniques for phonocardiogram diagnosis." *Computers and Biomedical Research*, 9(4): 349-363, 1976.
[94] Zeiler, A., Faltermeier, R., Keck, I. R., Tomé, A. M., Puntonet, C. G., and Lang, E. W., "Empirical mode decomposition-an introduction." *IEEE International Joint Conference on Neural Networks (IJCNN)*, pp. 1-8, 2010.
[95] Bajelani, K., Navidbakhsh, M., Behnam, H., Doyle, J. D., and Hassani, K., "Detection and identification of first and second heart sounds using empirical mode decomposition." *Proceedings of the Institution of Mechanical Engineers, Part H: Journal of Engineering in Medicine*, 227(9): 976-987, 2013.
[96] Gavrovska, A., Zajić, G., Bogdanović, V., Reljin, I., and Reljin, B., "Identification of S1 and S2 heart sound patterns based on fractal theory and shape context." *Complexity*, 2017: 1-9, 2017.
[97] Ignacio Medina Rivas, "Detección de patologías cardíacas por medio del análisis de fonocardiogramas", *B.Sc. Thesis*, Universidad Politécnica de Madrid, Oct. 2020.
[98] Javier Rubio Tardío, "Desarrollo de una aplicación en Matlab para el procesado de señales de fonocardiograma", *B.Sc. Thesis*, Universidad Politécnica de Madrid, Oct. 2020.

ECG-Based Biometric Recognition

David Meltzer[1[0000-0003-3074-9023]], David Luengo[1[0000-0001-7407-3630]] and Tom Trigano[2[0000-0002-9912-1499]]

[1] Universidad Politécnica de Madrid, 28031 Madrid, Spain
[2] Shamoon College of Engineering, Ashdod, Israel

Abstract. Biometric recognition refers to the identification of a subject through some characteristic physical or behavioral feature. Some of these traits (like fingerprints) have been used for a long time to identify individuals, while others (like facial or iris recognition) are becoming increasingly more popular. The electrocardiogram (ECG) is a biomedical signal that is always present when the subject is alive and which has been considered in biometrics during the last 20 years. In this book chapter we perform an introduction to biometrics, the ECG and biometric recognition systems first. This is followed then by a brief description of the two main blocks from a signal processing/machine learning point of view: the feature extractor and the pattern matcher/decision rule.

Keywords: Biometric recognition, Electrocardiogram (ECG), Feature extraction, Pattern matching.

1 Introduction

1.1 Biometrics

Recognizing the identity of individuals is an intrinsic characteristic that humans use without even emerging to consciousness. Identifying individuals is a basic ability needed for social interaction that takes as inputs external information, such as the face image, the speaker's voice or the gait.

The study of human characteristics has always been of interest in different knowledge areas with many different applications. During the latter 19th century a growing interest emerged, but it was not until the 20th century that biometrics was first used. The term was initially related to the study of massive biological phenomena and the derivation of knowledge by means of statistical methods in the field of evolutionary biology [1]. The last quarter of the century saw how the term biometrics changed its meaning until it got its contemporary significance: "the quantification of unique distinguishing of individuals for the purpose of individual personal identification" [2].

Biometrics is founded on a generally accepted principle of immutability, affirming that some specific traits produced by a human can be quantified and then compared with new obtained samples of the same traits and individual, experiencing only minor

changes between them. It is generally accepted to use the term *pattern* when referring to a quantified trait sample. Recognizing individuals by means of pattern-matching problem solving is commonly referred as *biometric recognition*.

As awareness increased about the potential applications of biometrics, the interest in the subject grew dramatically during the last two decades; altogether with increased availability of sensors and data processing capacity. A simple evidence of the previous assertion is presented in **Fig. 1**, where a dramatic increase in the number of citations associated with the term "biometric identification", found in the Web of Science Core Collection since 1995, is shown.

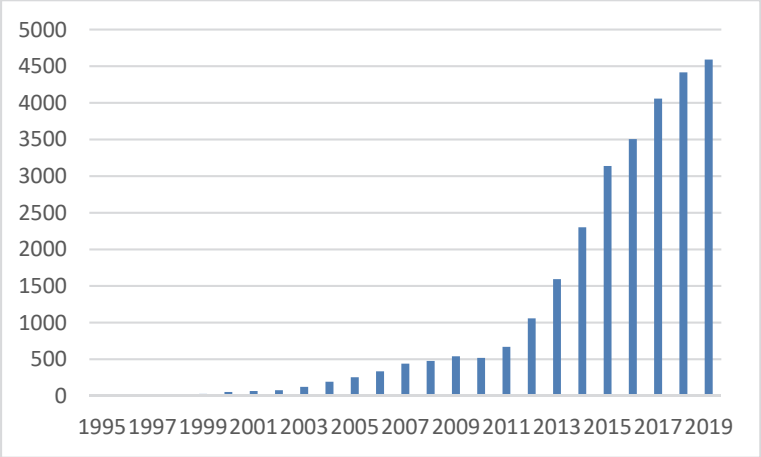


Fig. 1. Number of citations associated to the term "biometric identification" found on the Web of Science Core Collection since 1995.

Biometric recognition, considered as the recognition of individuals, particularizes in two alternatives. On the one hand, recognition is referred as *identification* when the problem to solve consists in determining the individuals' identity that originated a set of patterns by considering all possible identities as feasible. On the other hand, when identification considers only a claimed identity as possible, recognition is referred as *identity verification*.

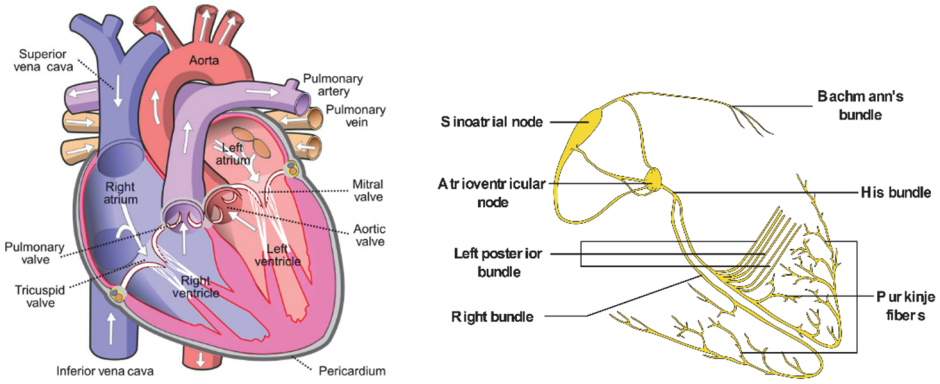
Three basic methods have been described for individual identification: demonstrating the individuals' possession of some physical (or logical) token, providing a proof that the individual possesses a unique knowledge or by means of individuals' intrinsic, biometric personal traits (biological, physical and/or behavioral). The referred mechanisms are commonly called ownership, knowledge and inherence respectively. Providing proof of the possession of a token, whether physical or logical, is referred as **ownership-based identification**. A physical key needed to open a padlock or an authentication key stored inside a SIM card that enables a subscriber to associate with a network are both simple examples of ownership-based identification methods. **Knowledge-**

based identification, on its side, can be experienced when a numeric combination is used to open a padlock or when a digital signature is made public after being generated from a private key in a cryptosystem. Both previously presented identification mechanisms rely on the capacity of the individual to manipulate extrinsic tokens. **Biometrics-based identification** differs from the previously shown mechanisms in the fact that it does not rely on extrinsic elements for identification, but only depends on the inherent capacity of the individual to generate unique trait samples that can be the subject of quantification and matching. Both ownership and knowledge-based mechanisms have been subject of criticism [3], as physical and logical tokens can be subject of disclosure or stealing that can result in impersonation and illicit use. On its side, biometrics-based identification is not subject of the mentioned vulnerabilities, so no special risk mitigation systems are required. The fact that no special protection measures are needed on the individuals' side when considering adopting biometrics for identification makes this option specially interesting. Whereas biometrics provide solid advantages in specific identification scenarios, criticism for its use still exists when taking into account different evaluation criteria for identification schemes [2] [4].

1.2 The electrocardiogram

The electrocardiogram (ECG) consists on capturing the electrical dynamics of the heart. The dynamics registration is usually performed by means of non-invasive electrodes that are placed in direct contact to the skin on the body's surface. The captured variations in potential between electrodes are recorded, where those variations are created by cardiac cells forcing other cells to contract. Even though in the late 19th century technical development enabled for the registration of the low level of potentials generated by the heart, it was not until the first quarter of the 20th century that the ECG registration techniques dramatically developed and widespread adoption began. A number of clinical applications of the ECG have been identified since then, and health has remained its most relevant area of application [5].

The main objective of the heart is pumping blood throughout the body to deliver oxygen to its cells. As depicted in **Fig. 1**(a), it is divided into two sides, right and left. Each side pumps blood belonging to a different circulatory system and both work rhythmically synchronized. Each side is divided into two chambers: the atrium, where blood enters, and the ventricle, that forces blood to circulate through the body.



(a) (b)

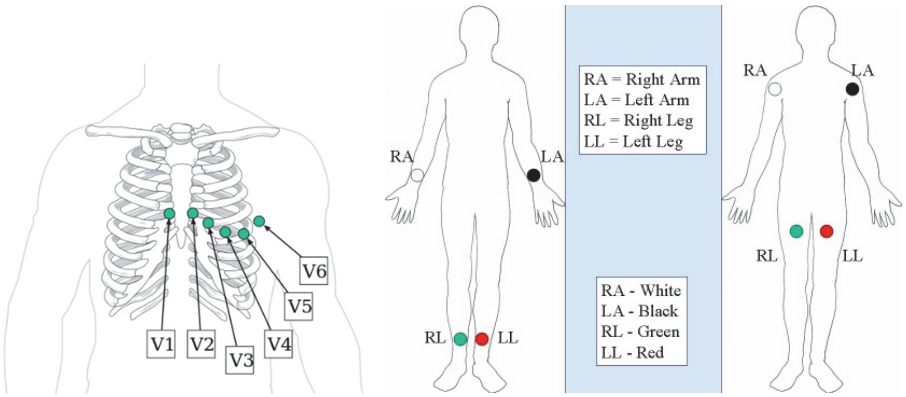
Fig. 2. a) Schematic diagram of the heart. *"File:Diagram of the human heart (cropped).svg" by Wapcaplet is licensed under CC BY-SA 3.0*. b) Schematic of the heart conduction system. *"File:ConductionsystemoftheheartwithouttheHeart-en.svg" by Madhero88 and Angelito7 is licensed under CC BY-SA 3.0.*

The blood is forced to circulate in a determined direction by means of valves: the atrioventricular valves, that set the boundaries between the atria and the ventricles, and the pulmonary and mitral valves, placed between the ventricles and the pulmonary artery and the aorta, respectively.

The walls of the heart are composed of muscle cells, the myocardium, enclosed into an outer protective layer, the pericardium. The myocardium is capable of contraction, reducing its enclosed volume and resulting in blood being pumped into the pulmonary artery and aorta. The myocardium's contraction is driven by electrical impulses that propagate by means of specialized myocardium cells that conform a network called the conduction system, depicted in **Fig. 2(b)**. Physiological myocardium relaxation follows the contraction, and the process starts again, forming what is called a cardiac cycle. In normal conditions, the cells in the sinoatrial node start the cycle by spontaneously generating an electrical impulse (called electrical wavefront) that propagates into the atrioventricular node and forces contraction of the atria. After a delay, the signal propagates through the His bundle and into the Purkinje fibers, forcing the ventricles to contract subsequently. This is known as the activation phase, which is described as depolarization in electrical terms. The recovery phase follows activation and it is characterized by a fast repolarization of the cardiac cells.

The most common ECG registration configuration, known as the 12-lead ECG, uses 10 electrodes that are placed in standardized body positions on the skin's surface, as shown in **Fig. 3**. The precordial leads are a group of six electrodes, identified as V_1 to V_6, placed on the chest. Being close to the heart, they provide detailed information on its electrical activity, taking as reference the Wilson central terminal, corresponding to

the average tension values recorded by the right and left arm and the left leg. The bipolar limb leads are used to record the cardiac potentials on the frontal plane.



(a) (b)

Fig. 3. Electrodes placement in the 12-leads ECG recording. a) Precordial leads. b) Bipolar limb leads. *Source Wikimedia Commons*.

The signals received by the electrodes result in 12 ECG signals called waves. A common ECG wave, depicted in **Fig. 4(a)**, shows the activity of atrial depolarization as the P wave, while ventricular depolarization is reflected by the QRS complex. The T wave represents ventricular repolarization. The amplitudes of waves are measured with respect to the so-called isoelectric line reference, considered just before the QRS complex. The ECG recording equipment must provide a high gain, altogether with a large dynamic range, because the ECG wave can present maxima and minima in the microvolts range, but it can also attain amplitudes higher than one Volt. Additionally, the isoelectric line, also called baseline, can show a wandering effect of up to one Volt. **Fig. 4(b)** shows an actual ECG wave sample corresponding to the precordial lead V4 of a subject from the Physikalisch-Technische Bundesanstalt (PTB) database [6] available on PhysioNet [7].

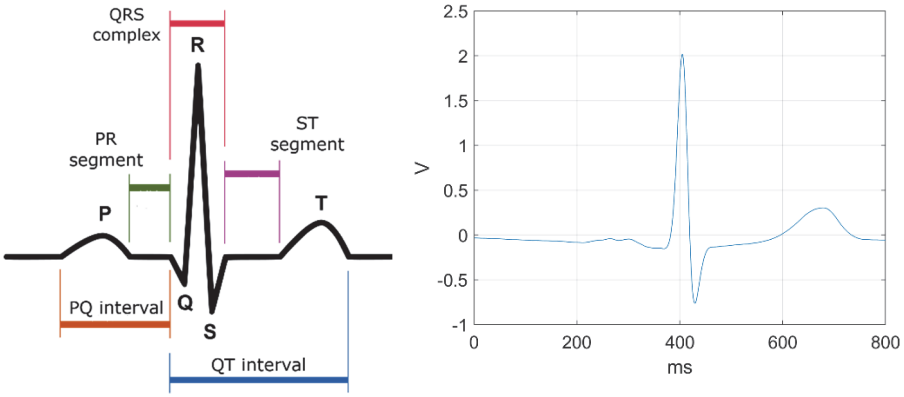


Fig. 4. a) Simplified representation of an ECG wave. *Source Wikipedia Commons, Created by Agateller (Anthony Atkielski)*. b) Actual ECG wave without wandering effect.

2 Biometric systems

A biometric system is designed for the resolution of the personal identification of individuals, or the validation of their claimed identities, by considering the acquisition of one or more quantified physiological and/or behavioral traits [8]. Once this information is considered, an output is generated that consists in a probable identity, or an assessment on the claimed identity, for identification and authentication, respectively [9].

As previously stated, biometrics rely on solving a pattern matching problem for identification or authentication. The patterns (alternatively referred as templates) have to be obtained from the set of users of the biometric system in order for the system to take them into consideration. Obtaining patterns from the users of the biometric system and storing them into a pattern storage system for further use is commonly referred to as the **enrollment** process. A basic enrollment process is depicted in **Fig. 5**.

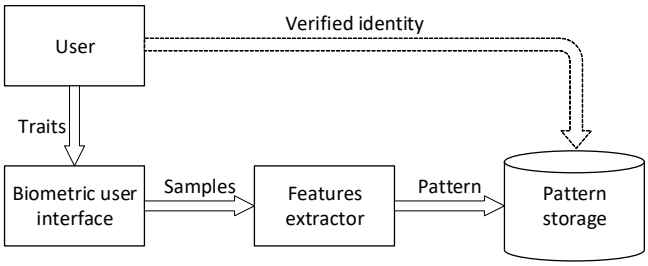


Fig. 5. Biometric enrollment process.

The future user of the biometric system lets the biometric user interface acquire one or more specific personal traits that are quantified into samples. The obtained samples are then used as the input to the features extractor, that extracts unique features from them, called patterns. Finally, the pattern (or patterns) obtained corresponding to the future user are stored, altogether with the verified identity of the user, that, from this moment, can be considered a user of the biometric system.

When the pattern storage contains the intended set of users' patterns, the biometric system can be used for identification or identity verification. **Fig. 6** shows a simple biometric system for identification.

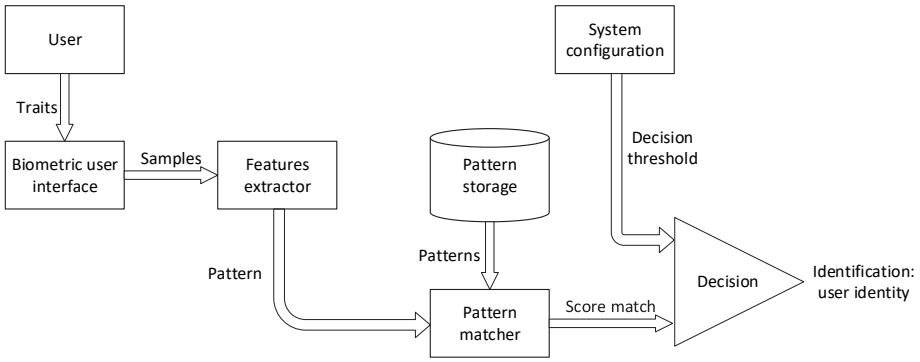


Fig. 6. Biometric identification of an enrolled user.

The **identification** process starts analogously as the enrollment: the user of the biometric system lets the biometric user interface acquire one or more specific personal traits that are quantified into samples; those samples are subsequently the subject of the features extraction process in order to obtain a pattern. At this point, the pattern is matched with the existing patterns in the pattern storage system, producing a number of pattern scores that are used by the decision stage, altogether with a configurable decision threshold, to output the most likely identity.

The **authentication** process is similar to identification, as can be seen in **Fig. 7**. The personal traits are processed as in the case for identification, but in this case the user claims for an identity. The claimed identity permits the system to identify the pattern (or set of patterns) that corresponds to it in the pattern storage system and the matching is performed exclusively against it, resulting in a single matching score. Altogether with the decision threshold, the score match is used to output a positive or negative identity authentication.

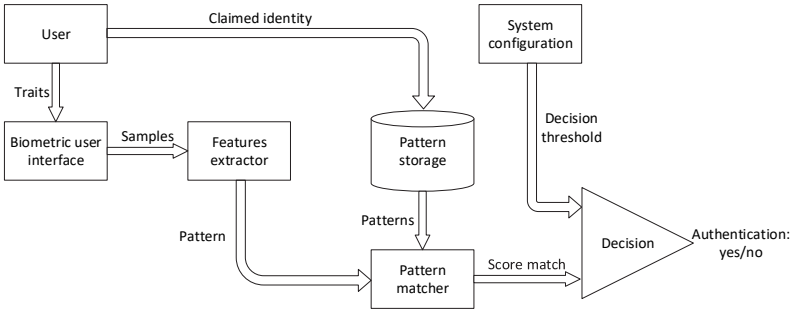


Fig. 7. Biometric authentication of an enrolled user.

In the following two sections we briefly review the two key blocks in any biometric recognition system: the features extractor, and the pattern matcher and decision rule.

3 Feature extraction

From a machine learning point of view, feature extraction refers to the process where the most relevant characteristics are obtained from the data. Ideally we would like to be able to extract a minimal set of independent features that include all the relevant information (for the following pattern matching and decision blocks) which is contained in the data. In biometric recognition, feature extraction approaches can be classified in three large groups:

- **Fiducial methods:** These methods are based on the location of several prominent and physiologically relevant points in the biomedical signals of interest. After locating these fiducial points, features are typically computed on the basis of their (relative) locations and amplitudes.
- **Non-Fiducial methods:** These methods do not require the location of any relevant samples in the signal of interest. They rely on performing some transformation of the data (possibly after dividing the signals in arbitrary segments of fixed length) and computing the features in the transformed space.
- **Semi-Fiducial methods:** A very small set of easy to compute fiducial points (e.g., 1 or 2) are extracted in order to perform the signal segmentation and then the features are computed by applying some transformation as in non-fiducial methods.

In the sequel, we briefly review these three approaches.

In ECG-based biometric recognition, **fiducial methods** locate first the R peak associated to each heart beat (the most prominent waveform, as shown in **Fig. 4**), and then perform waveform delineation (see Chapter 10) to determine the location and duration

of all the other relevant waveforms (P-QRS-T) in each cardiac cycle. Once this information has been obtained, the most widely used features correspond to the duration of the different waveforms and clinically relevant intervals. However, features related to the amplitude of the different peaks and other aspects of the P-QRS-T waves can also be used. Indeed, the first ECG-based biometric approaches used a mixed set of 30 features which were a combination of temporal features (9), amplitude features (13) and others (QRS wave area, T wave morphology, ST segment slope, etc.) [10, 11]. A similar approach was followed by other authors in the following years: [12] used 4 temporal and 3 amplitude features, while [13] used 13 temporal features, 4 amplitude features and 3 angular features. However, in 2005 Israel *et al.* proposed a set of 15 temporal features that have been extensively used since then [14], and which are shown in **Fig. 8**. A few years later, Wang *et al.* proposed a set of 6 amplitude features that measure the difference in amplitude among the following points [15]: PL', PQ, RQ, RS, TS and TT'. Although many other features based on fiducial points have been proposed, altogether the 15 temporal features proposed in [14] and the 6 amplitude features from [15], have become a *de facto* standard feature set in many fiducial ECG-based biometric approaches. **Fig. 9** shows these 21 features.

A related and very relevant topic for fiducial ECG-based biometric approaches is **dimensionality reduction**. The aforementioned temporal and amplitude features are based on physiological criteria, and they are heavily correlated. Consequently, using them directly as the input of the pattern matching/decision stages usually leads to a suboptimal performance and an unnecessarily high computational cost. Dimensionality reduction approaches typically select either a subset of the most relevant features or a reduced number of linear combinations of features to optimize the performance and reduce the computational/storage burden. The first approach is also known as feature selection and is used, for example, in [10, 11], where features with a high correlation with other features are removed, thus reducing the number of features used from 30 to 12. A second example of this kind of strategy is the one applied in [14], where a stepwise canonical correlation procedure, based on Wilk's lambda, is followed. Regarding the second approach, the most common procedure to obtain a reduced number of linear combinations of features is principal component analysis (PCA). This method is applied, for example, in [16].

No.	Feature	Description
1	RQ	Elapsed time from the Q to the R waveform
2	RS	Elapsed time from the R to the S waveform
3	RP	Elapsed time from the mid-point of the P waveform to the R waveform
4	RL	Elapsed time from the beginning of the P waveform to the R waveform
5	RP'	Elapsed time from the end of the P waveform to the R waveform
6	RT	Elapsed time from the Q waveform to the mid-point of the T waveform
7	RS'	Elapsed time from the R waveform to the beginning of the T waveform
8	RT'	Elapsed time from the R waveform to the end of the T waveform
9	P width	Time width of the P waveform
10	T width	Time width of the T waveform
11	ST	Elapsed time from the S waveform to the mid-point of the T waveform
12	PQ	Elapsed time from the P to the Q waveform
13	PT	Elapsed time from the P to the T waveform
14	LQ	Elapsed time from the beginning of the P waveform to the Q waveform
15	ST'	Elapsed time from the S waveform to the end of the T waveform

Fig. 8. Temporal features proposed in [14] and frequently used in fiducial methods.

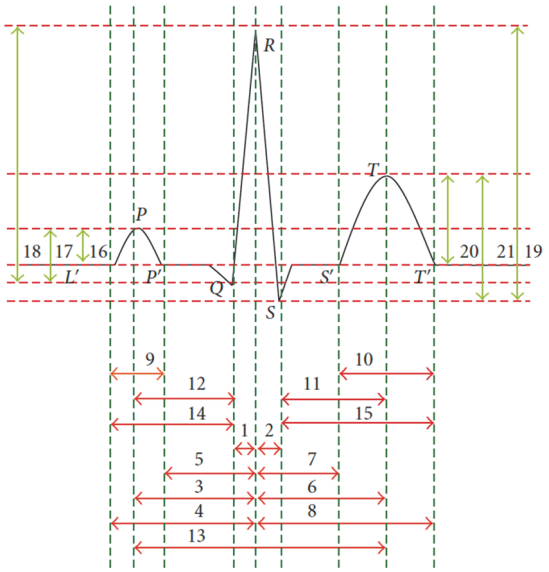


Fig. 9. Temporal and amplitude features frequently used in fiducial methods 15. Features 1-15 are the temporal features shown in Fig. 8 14, whereas features 16-21 are amplitude features 15.

Non-fiducial methods for ECG-based biometric recognition were initially proposed by Plataniotis et al. 17. Their proposed approach consists in taking segments of fixed length of the ECG (typically 10 seconds), performing their normalized autocorrelation (AC), which is unique for each subject, and then applying the discrete cosine transform (DCT) in order to obtain just a few non-zero coefficients (i.e., relevant features). This approach was further investigated in 18, and a comparison with the fiducial approach described above was performed in 15. Non-fiducial methods can be more robust than fiducial approaches and require less computational cost, since they do not depend on the location of fiducial points, which is computationally demanding and difficult for noisy signals. However, further research is required to show their scalability to larger data bases, for example.

Finally, **semi-fiducial** ECG biometric approaches can be devised by using the location of the R peak (which can be easily and quickly located) for waveform segmentation, and then applying the AC/DCT approach described above or any other non-fiducial method.

Note that a fourth alternative has recently emerged. Several researches have proposed the application of deep learning approaches to ECG-based biometrics during the last few years. In this class of methods, the separation between the feature extraction and the pattern matching/decision stages disappears. Typically, the deep neural network (DNN) receives the raw data (formatted in some appropriate way) as the input, performs some internal dimensionality reduction and feature extraction automatically in some complicated and non-linear way which is difficult to interpret, and provides the matching score at the output. A brief account of several proposed DNN-based approaches for ECG biometric recognition is provided at the end of the following section.

4 Pattern matching and decision

As it was described in the previous section, patterns can be obtained both taking into account specific time-domain features or by considering frequency-related ones. Regardless of the considered features, patterns are presented as vectors that are used by the pattern matcher and decision subsystems in order to output a probable user identity or an assertion of authentication. The two referred subsystems have been described separately for the sake of clarity, but in some cases their actual implementations do not pose strict boundaries between them, as already mentioned in the case of DNNs.

Pattern matching represents the capability to compare patterns and obtain a quantification of their similarity, usually referred to as a score match. Strictly related to this concept lays the idea of pattern recognition, which involves the problem of searching regularities in data that can be used, for instance, to categorize it. A simple approach to

the problem consists in creating mainly heuristics-based algorithms that rely on the ability to take all relevant features into consideration, but most of the times this method is simply not feasible. A more generic approach can be taken by using a machine learning paradigm 12. A sufficiently large set of n patterns $\{x_1, \dots, x_n\}$, called *training set*, is initially used to learn the parameters of an adaptive model. The training set is usually fed with patterns obtained during the biometric enrollment process, in this mode each pattern x_i is directly related to a known user of the system. Multiple patterns from a single user can be used in the training set, and all of them are considered to belong to the same *class* (also called *category*). Therefore, in a training set with n patterns, k discrete classes exist $\{C_1, \dots, C_k\}$, one for each single user of the biometric system.

The class to which a single pattern belongs can be represented by means of a target vector $\boldsymbol{t}$. The training process (also commonly known as learning phase) consists in establishing the parameters of an adaptive model taking as inputs the training set and the set of target vectors of every single training pattern in the training set. Training with the aforementioned input data is also referred to as *supervised learning*. The outcome of the training process is a set of parameters for the adaptive model that can be expressed as a function $y(x)$ such that, given an input pattern x, a corresponding target vector y is obtained with the same format as the target vectors used during training. The recognition process naturally uses different patterns other than the ones used in training and target vectors are obtained that identify concrete classes. This fundamental capability of the trained models is known as *generalization*.

As previously stated, the matching and decision stages cannot always be independently treated when it comes to solve decision problems. Indeed, three different approaches based on multivariate statistical methods have been identified for solving them 12:

1. **Generative models:** These models use the posterior class probabilities that a pattern x belongs to a particular class C_i, $p(C_i|x)$, calculated using the Bayes' theorem,
$$p(C_i|x) = \frac{p(x|C_i)p(C_i)}{p(x)},$$
after inferring the class-conditional densities $p(x|C_i)$ and $p(C_i)$ from the training set data. After determining the posterior class probabilities, they use decision methods to finally decide what class x belongs to.
2. **Discriminative models:** They use direct inference of the posterior class probabilities, $p(C_i|x)$, and then use decision methods to determine what class x belongs to.
3. **Discriminant functions:** These methods aim to define a function $f(x)$ that maps each pattern x to a class label. This approach is characterized by joining the matching and decision processes into one single stage.

Different approaches have been followed in actual ECG-based recognition problems to date. Relevant works are presented here attending to the classification techniques they use.

Linear Discriminant Analysis (LDA) for classification 12 was taken as the classification technique in the works of Shen *et al.* 20 and Israel *et al.* 14, while **Quadratic Discriminant Analysis (QDA)**, on its side, was selected as classifier by Sarkar *et al.* 21.

A number of non-conventional different methods have been tried too. Wang *et al.* 15, as well as Agrafioti *et al.* 22, used **Nearest Neighbor (NN)** classification with the Euclidean distance for similarity quantification. **Suppor Vector Machines (SVMs)** 23 have been used as classifiers in the works of Choi *et al.* 24, Chu *et al.* 25 and Liu *et al.* 26.

Meltzer *et al.* 27 compare the most common conventional and non-conventional classifiers (LDA, Quadratic DA, kNN 28 and multiclass SVM 2930) applied to selected sets of ECG fiducial features in terms of their resulting confusion matrix derivations.

Special interest exists lately in the use of deep learning techniques for ECG recognition, which represents an interesting alternative to both widely used multivariate statistical and other non-conventional techniques: Hong *et al.* 31, Labati *et al.* 31 and Zhao *et al.* 33 are some examples of recent developments that use the aforementioned technique.

5 Discussion

In this book chapter we have briefly reviewed some of the main aspects related to ECG-based biometric recognition. Our main goal has been describing the potential of the ECG signal as a biometric trait, defining the structure of the typical systems required for enrollment, identification and authentication, and providing an overview of some of the key blocks in any biometric system. Several reviews of ECG-based biometric systems that go beyond the scope of this book chapter have been published during the last few years 343536. However, there are still many open issues that need to be addressed:

- The availability of public databases with a large number of individuals. So far the most commonly used are the following ones: MITDB, MIT-BIH, ECG-ID, CEBS-DB, Long-Term ST Database, MIT-BIH Noise Stress Test Database, MIT-BIH Supraventricular Arrhythmia Database, PTBDB, and CYBHi. However, none of them is large enough to test the scalability of current ECG-based biometric systems to large populations of individuals.
- The robustness of current ECG-based biometric systems with respect to noise, healthy and unhealthy patients (i.e., patients with some kind of pathology), intra-class ECG effects (e.g., sporadic pathologies, aging or changes in physical condition), etc.

- The practical implementation of ECG-based biometric systems: privacy issues (which is particularly relevant in this case, as the ECG can expose pathologies from the individuals making use of this system), low-cost devices for ambulatory measurement of the ECG signal with a proper quality level, combination of ECG-based biometry with other types of biometric recognition (e.g., fingerprints) for increased robustness, etc.

A basic introduction to the topics of ECG-based biometrics and biometric systems can be provided in any 3rd or 4th year biomedical signal processing course, or as part of a larger biometric recognition course, at the B.Sc. level regardless of the background of the students. However, a more detailed analysis of the feature selection, pattern matching and decision rules applied requires knowledge of statistical signal processing and machine learning concepts and is more appropriate for the M.Sc. level. ECG-based biometric recognition is also a good topic for B.Sc./M.Sc. thesis 3738.

References

1. Galton, F.: Biometry. *Biometrika*, 1(1), pp. 7-10 (1901).
2. Li, Stan Z., and Anil Jain. *Encyclopedia of biometrics*. Springer, 2015.
3. Bonneau, J.: "The quest to replace passwords: A framework for comparative evaluation of web authentication schemes." *IEEE Symposium on Security and Privacy*, pp. 553-567, 2012.
4. Zimmermann, V.: "The quest to replace passwords revisited–rating authentication schemes." *Twelfth International Symposium on Human Aspects of Information Security & Assurance (HAISA 2018)*, pp. 38, 2018.
5. Sörnmo, Leif, and Pablo Laguna. *Bioelectrical signal processing in cardiac and neurological applications*. Academic Press, 2005.
6. Bousseljot R, Kreiseler D, Schnabel, A. "Nutzung der EKG-Signaldatenbank CARDIODAT der PTB über das Internet." *Biomedizinische Technik*, Band 40, Ergänzungsband 1 (1995) S 317
7. Goldberger, A., et al. "PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals." *Circulation*, 101(23), pp. e215–e220, 2000.
8. Jain, Anil K., Arun A. Ross, and Karthik Nandakumar, *Introduction to biometrics*. Springer Science & Business Media, 2011.
9. J. R. Vacca, "Biometrics," *Computer and Information Security Handbook*, Morgan Kaufman (2nd ed.), 2012.
10. Biel, L., Pettersson, O., Philipson, L., and Wide, P. "ECG analysis: a new approach in human identification." *16th IEEE Instrumentation and Measurement Technology Conference*, Vol. 1, pp. 557-561, May 1999.
11. Biel, L., Pettersson, O., Philipson, L., and Wide, P. "ECG analysis: a new approach in human identification." *IEEE Transactions on Instrumentation and Measurement*, 50(3): 808-812, 2001.
12. Shen, T. W., Tompkins, W. J., and Hu, Y. H. "One-lead ECG for identity verification." *Second Joint 24th Annual Conference and the Annual Fall Meeting of the Biomedical Engineering Society*, Vol. 1, pp. 62-63), Oct. 2002.

13. Singh, Y. N., and Gupta, P. “ECG to individual identification.” *IEEE Second International Conference on Biometrics: Theory, Applications and Systems*, pp. 1-8, Sep. 2008.
14. Israel, Steven A., et al. “ECG to identify individuals.” *Pattern recognition* 38(1): 133-142, 2005.
15. Wang, Y., Agrafioti, F., Hatzinakos, D., and Plataniotis, K. N. “Analysis of human electrocardiogram for biometric recognition.” *EURASIP Journal on Advances in Signal Processing*, 2008(1): 1-11, 2008.
16. Irvine, J. M., Israel, S. A., Scruggs, W. T., and Worek, W. J. “eigenPulse: Robust human identification from cardiovascular function.” *Pattern Recognition*, 41(11): 3427-3435, 2008.
17. Plataniotis, K. N., Hatzinakos, D., and Lee, J. K. “ECG biometric recognition without fiducial detection.” *2006 Biometrics symposium: Special session on research at the biometric consortium conference*, pp. 1-6, Aug. 2006.
18. Agrafioti, F., and Hatzinakos, D. “ECG based recognition using second order statistics.” *6th IEEE Annual Communication Networks and Services Research Conference (CNSR 2008)*, pp. 82-87, May 2008.
19. Bishop, Christopher M. *Pattern recognition and machine learning*. Springer, 2006.
20. Shen, Tsu-Wang David, Willis J. Tompkins, and Yu Hen Hu. “Implementation of a one-lead ECG human identification system on a normal population.” *Journal of Engineering and Computer Innovations*, 2(1): 12-21, 2010.
21. Sarkar, Abhijit, A. Lynn Abbott, and Zachary Doerzaph. "ECG biometric authentication using a dynamical model." *IEEE 7th International Conference on Biometrics Theory, Applications and Systems (BTAS)*, 2015.
22. Agrafioti, Foteini, and Dimitrios Hatzinakos. “ECG biometric analysis in cardiac irregularity conditions.” *Signal, Image and Video Processing*, 3(4): 329, 2009.
23. Cortes, Corinna, and Vladimir Vapnik. “Support-vector networks.” *Machine Learning* 20(3): 273-297, 1995.
24. Choi, Hyun-Soo, Byunghan Lee, and Sungroh Yoon. “Biometric authentication using noisy electrocardiograms acquired by mobile sensors.” *IEEE Access*, 4: 1266-1273, 2016.
25. Chu, Yifan, Haibin Shen, and Kejie Huang. “ECG authentication method based on parallel multi-scale one-dimensional residual network with center and margin loss.” *IEEE Access*, 7: 51598-51607, 2019.
26. Liu, Jikui, *et al.* “A multiscale autoregressive model-based electrocardiogram identification method.” *IEEE Access*, 6: 18251-18263, 2018.
27. Meltzer, David, and David Luengo. “Fiducial ECG-Based Biometry: Comparison of Classifiers and Dimensionality Reduction Methods.” *42nd IEEE International Conference on Telecommunications and Signal Processing (TSP)*, 2019.
28. Cover, Thomas, and Peter Hart. “Nearest neighbor pattern classification.” *IEEE Transactions on Information Theory*, 13(1): 21-27, 1967.
29. Duan, Kai-Bo, and S. Sathiya Keerthi. “Which is the best multiclass SVM method? An empirical study.” *International Workshop on Multiple Classifier Systems*, 2005.
30. Crammer, Koby, and Yoram Singer. “On the algorithmic implementation of multiclass kernel-based vector machines.” *Journal of Machine Learning Research*, pp. 265-292, 2001.
31. Hong, Pei-Lun, *et al.* “ECG biometric recognition: template-free approaches based on deep learning.” *41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2019.

32. Labati, Ruggero Donida, *et al.* "Deep-ECG: convolutional neural networks for ECG biometric recognition." *Pattern Recognition Letters*, 126: 78-85, 2019.
33. Zhao, Zhidong, *et al.* "ECG authentication system design incorporating a convolutional neural network and generalized S-Transformation." *Computers in Biology and Medicine* 102: 168-179, 2018.
34. Odinaka, I., Lai, P. H., Kaplan, A. D., O'Sullivan, J. A., Sirevaag, E. J., and Rohrbaugh, J. W. "ECG biometric recognition: A comparative analysis." *IEEE Transactions on Information Forensics and Security*, 7(6): 1812-1824, 2012.
35. Fratini, A., Sansone, M., Bifulco, P., and Cesarelli, M. "Individual identification via electrocardiogram analysis." *Biomedical Engineering Online*, 14(1): 78, 2015.
36. Pinto, J. R., Cardoso, J. S., and Lourenço, A. "Evolution, current challenges, and future possibilities in ECG biometrics." *IEEE Access*, 6: 34746-34776, 2018.
37. Gerardo Aguirre Vivar, "Desarrollo de métodos no fiduciales para sistemas biométricos basados en ECG", *B.Sc. Thesis*, Universidad Politécnica de Madrid, July 2019.
38. Ana Garci-Nuño Verbo, "Desarrollo de un método fiduciario robusto para biometría basada en ECG", *B.Sc. Thesis*, Universidad Politécnica de Madrid, Dec. 2019.

Modern Technologies for Biomedical Systems Prototyping

Anzhelika Parkhomenko[1] [0000-0002-6008-1610], Olga Gladkova[1] [0000-0002-6834-2854], Artem Tulenkov[1] [0000-0003-4863-4144], Yaroslav Zalyubovskiy[1] [0000-0002-6847-8778], Andriy Parkhomenko[1] [0000-0001-8265-0530] and Oleksandr Tarasov[2][0000-0002-0493-1529]

[1]National University Zaporizhzhia Polytechnic, 64, Zhukovskogo str., Zaporizhzhia, 69063, Ukraine
[2] Donbass State Engineering Academy, 72, Akademichna str., Kramatorsk, 84313, Ukraine
parhom@zntu.edu.ua

Abstract. This chapter focuses on the basics of prototyping technologies usage for design of cyber-prosthesis of human upper limb and wearable device for remote monitoring of patient physical condition. The results of the development and investigation of virtual and physical prototyping of schemes and structures of biomedical systems are presented.

Keywords: Biomedical Systems, Cyber Prostheses, Wearable Devices, Prototyping Technologies, Virtual Prototype, Digital Twins, 3D Printing, Remote Monitoring, Indicators of Human Physical State, Wi-Fi Module, Cloud Service.

1 Introduction

The development of information technologies in the field of medicine provides new opportunities for the creation of modern biomedical systems, in particular, cyber prostheses and wearable devices for monitoring indicators of the physical condition of patients [1-2]. Nevertheless, improvement of their functions, form, weight, dimensions, appearance and control algorithms remains a challenge [3].

It is known that prototyping has become an integral stage of the modern computer-aided design process, especially in the context of reducing the time and financing of production with the simultaneous requirement to ensure the quality of the finished product. Today prototyping is the name of several related technologies that can be effectively used in the design of biomedical systems [4-5].

In particular, virtual prototyping can be distinguished, in which a visual model of the digital twin of a designed object is created and studied, and physical prototyping, in which a full-scale model is created and studied.

The application of modern prototyping technologies helps to save time and money on development since it allows to identify unwanted errors and optimize the parameters and characteristics of biomedical products at the early stages of project implementation. Therefore, the study of these technologies is an urgent educational task.

2 Development of the Prototype of Cyber Prosthesis

2.1 Criterias of cyber prostheses design

The authors of work [6], have tested a big amount of cyber prostheses and identified three groups of criteria for their development. They are Durability / Cycle of use, Physical properties and Actuation or executive properties (Fig. 1). Developers have to find a compromise between these groups of criteria, so the design process for cyber prostheses is quite long, and their cost is too high for most people.

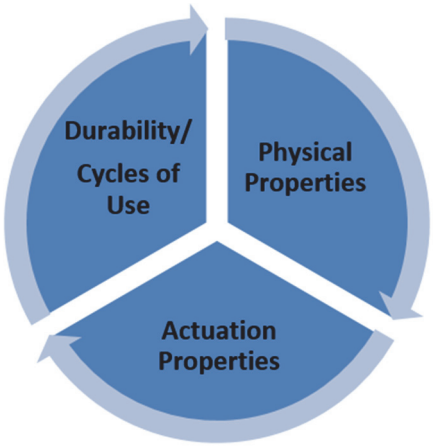


Fig. 1. The groups of criterias for cyberprostheses development.

2.2 Technologies and tools for prototype development

The diagram of the activities of the developers of the cyber prosthesis of the human upper limb using virtual and physical prototyping is presented in Fig. 2. The design process includes the creation of mechanical and electronic parts of the prototype and the development of a software.

The stages of the development of the mechanical part of the cyber prosthesis prototype are:

- Digitization with a scanner (for example, Gotcha) of the patient's healthy limb to obtain a general view of the 3D model of the hand and its dimensions (Fig. 3).
- Work in the CAD-environment on the creation of separate mechanical parts of the cyber prosthesis and the general assembly unit, taking into account the kinematic properties of the future product.
- Printing of individual mechanical parts using a 3D printer (for example, LeapFrog Creatr HS) and assembling a physical prototype [3].

The stages of the development of the electronic part of the cyber prosthesis prototype are:

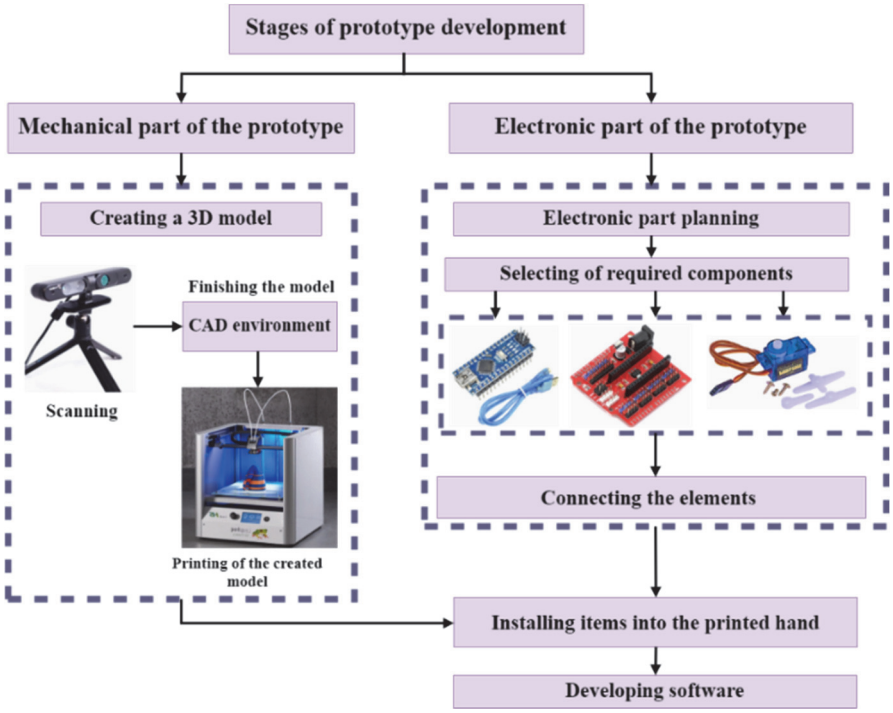


Fig. 2. The diagram of designers of cyber prosthesis prototype activities.

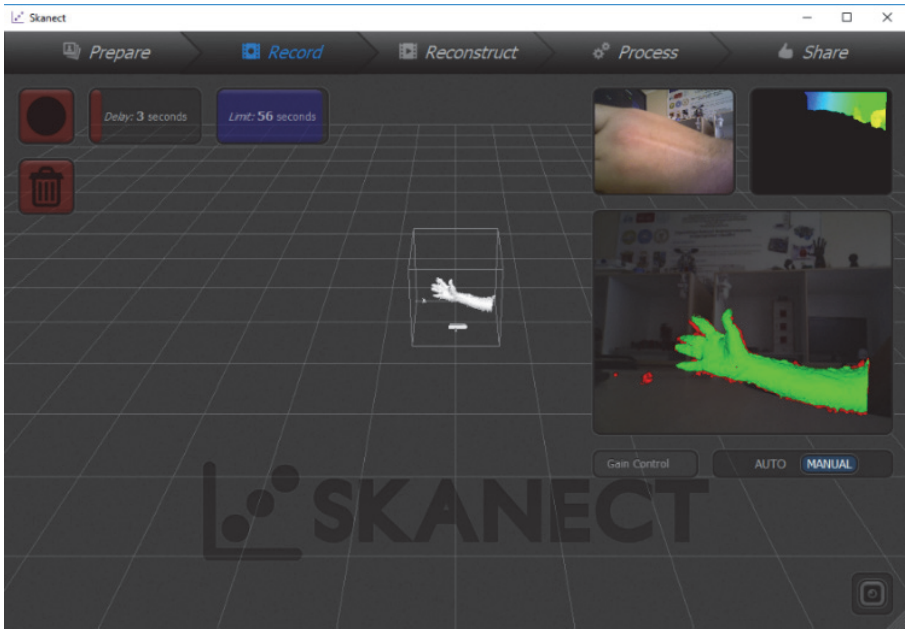


Fig. 3. Scanning process of human upper limb.

- Analysis and selection of electronic components.
- Connection of electronic components and their installation into the finished mechanical part of the prototype.

The development of the software for the cyber-prosthesis prototype functioning may begin immediately after the selection of electronic components as it related to certain development environments (for example, Arduino IDE), as well as software libraries associated with specific components (for example, Servo). The choice of electronic components is also important for the mechanical part of a cyber prosthesis design as it determines the size of the places for their further installation.

The main stages and steps of information technology for creating a mechanical part of a cyber prosthesis prototype are shown in Fig. 4.

At the first stage, based on the principles of hierarchy and decomposition, a structural diagram of the mechanical part of the cyber prosthesis is developed (Fig. 5). It contains separate parts (fingers, palm, base of the thumb) and an assembly subunit (the palm + base of the thumb).

PTC Creo Parametric software can be recommended for virtual prototypes of cyber prostheses construction creation and investigation as it is one of the eight best CAD environments in the world according to [7]. The system has powerful functionality and provides an intelligent design and parametric modeling environment.

Tabular parameterization can be used in the computer-aided design of a cyber prosthesis to quickly build models of similar parts (fingers) based on one parametric description. For its realization, a table with sizes and simplified equations for model parameters calculation should be implemented in the Creo environment.

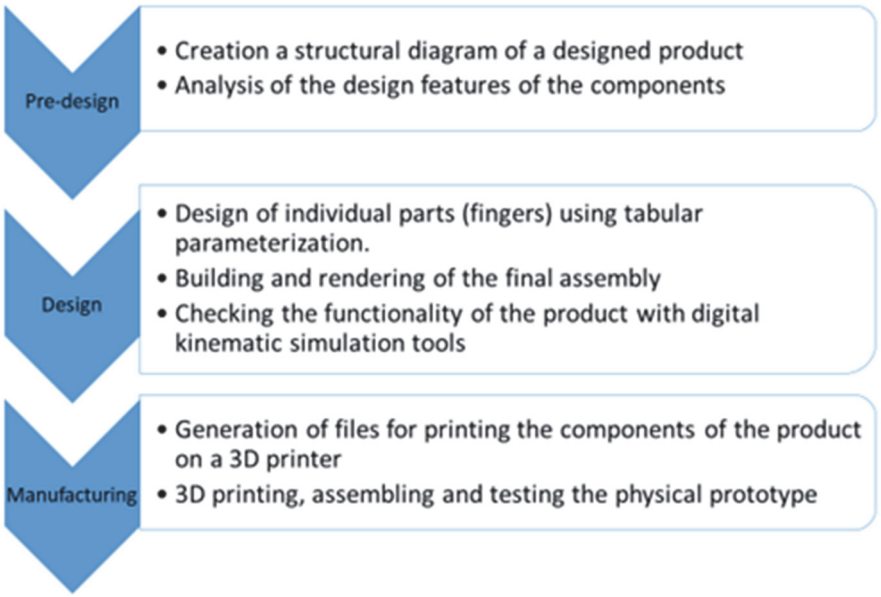


Fig. 4. Information technology for creating a mechanical part of a cyber prosthesis prototype.

After the creation of the files with 3D models of all five fingers and a palm in Creo Parametric, it is possible to develop the 3D model of the entire assembly (Fig. 6). First of all, it is convenient to apply all fingers for 3D printing at once instead of each finger separately (Fig. 7), as well as to evaluate how the artificial hand will look in the final.

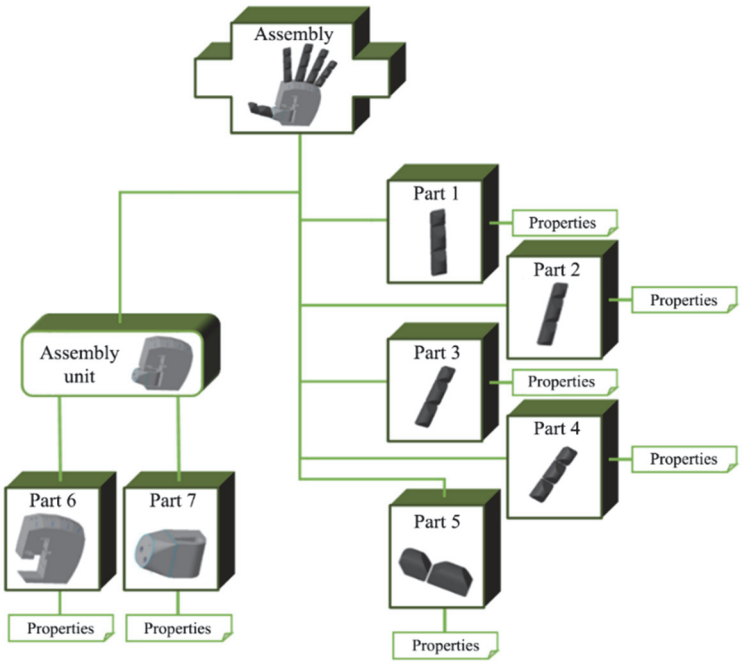


Fig. 5. Structure diagram of the mechanical part of the cyber prosthesis prototype.

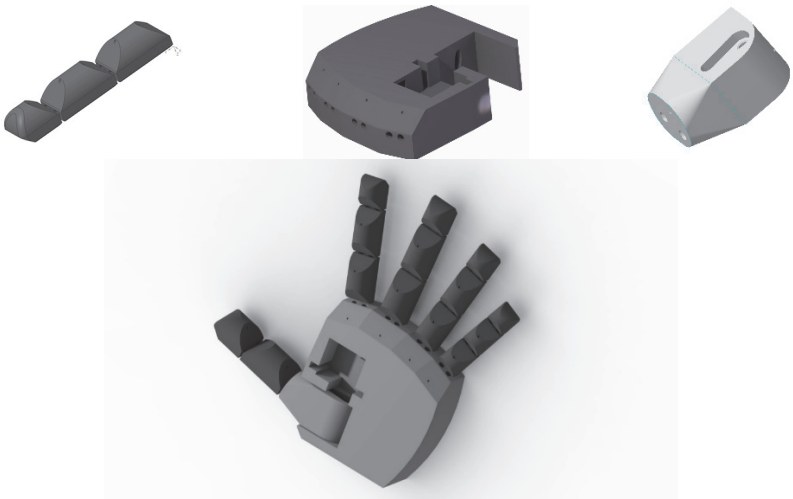


Fig. 6. Parts and assembly of cyber prosthesis virtual prototype in PTC Creo.

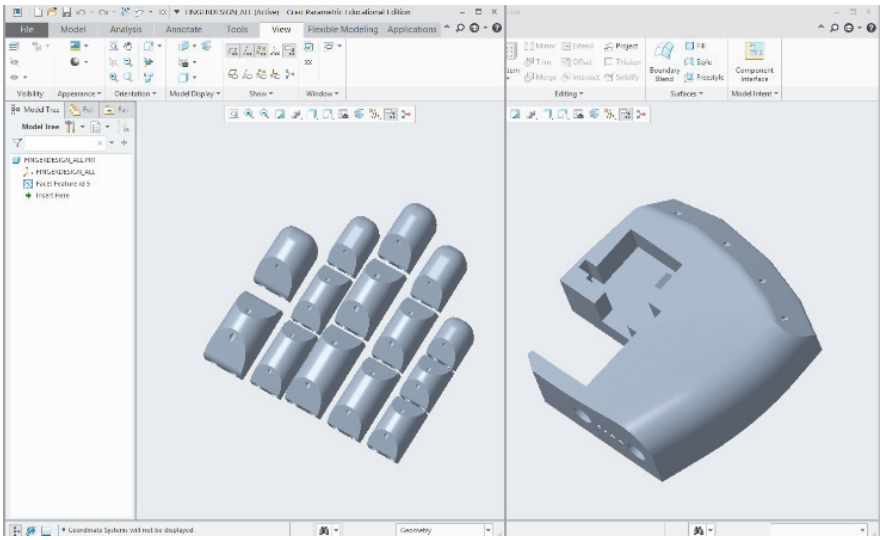


Fig. 7. Development of cyber prosthesis virtual model in PTC Creo.

2.3 Manufacturing and testing the physical prototype

After the development of the 3D models, STL files can be generated in the PTC Creo environment, that are the input data for the Simplify3D software. This software performs 3D model slicing for the 3D printing process, as well as setting the printing parameters (Fig. 8). The mechanical parts of the designed physical prototype (fingers, palm, base of the thumb) can be made using FDM (Fused Deposition Modeling) additive technology. As a material for printing on the printer LeapFrog Creatr HS, ABS plastic is used, which is resistant to adverse operating conditions, is not toxic and has a low cost.

Arduino Nano board, Arduino Nano Shield V3.0 expansion board and Tower Pro 9g SG90 servo motors can be selected as components of the electronic part of the prototype. The electronic components are connected and positioned on the mechanical basis of the prototype (Fig. 9).

The software developed in the Arduino IDE realizes the control of five fingers flexion and extension. Library Servo was used and special software components were developed for the functioning of the control program.

Functional testing is carried out according to different scenarios to check the performance of the cyber prosthesis prototype, as well as comparison with a digital twin is performed to check the conformity of shape and size (Fig. 10).

Therefore, the technology of complex virtual and physical prototyping using parametric modeling allows to quickly design and investigate a prototype of a future biomedical product to search errors and optimize design solutions.

In the future, it is possible to add sensors to control the force of pressure of parts of the prosthesis on an object during gripping, as well as to expand the list of achievable grips to improve control algorithms.

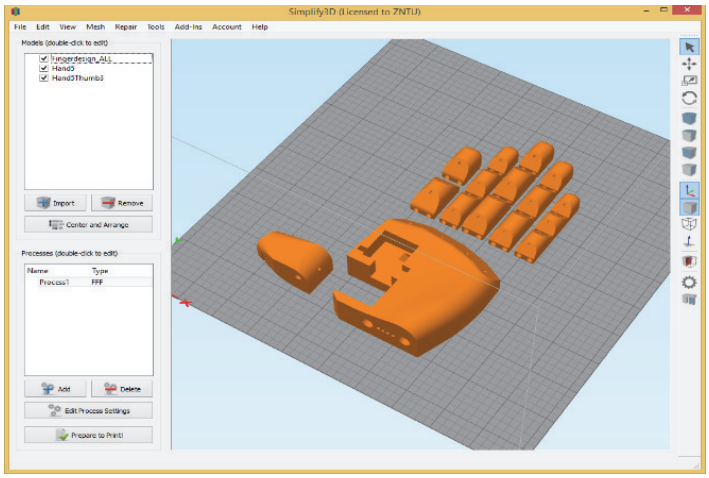


Fig. 8. Preparation of virtual prototype in Simplify3D for next 3D printing.

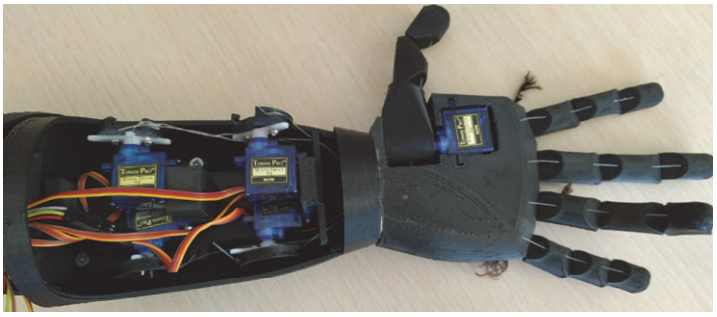


Fig. 9. Physical prototype of the cyber prosthesis.

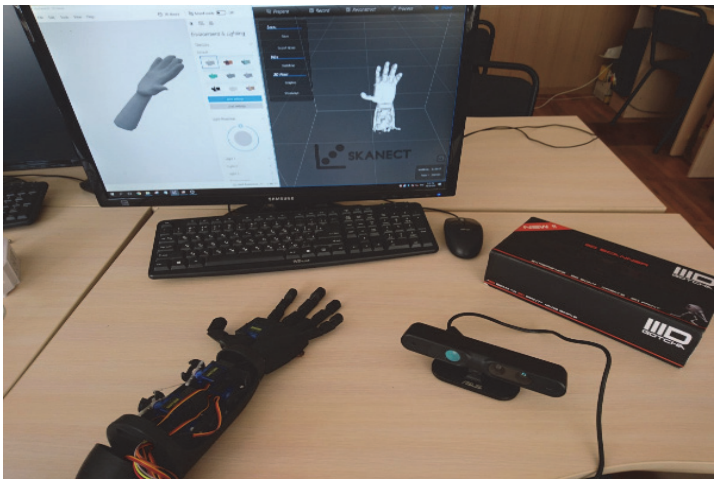


Fig. 10. Testing of the prototype of cyber prosthesis.

3 Development of the Prototype of Paramedical Bracelet

3.1 Wearable Devices Usage

Today, not only professional medical equipment but also popular wearable devices and corresponding mobile applications are actively used as means of indicators of the physical condition of people (temperature, heart rate, blood oxygen saturation, blood pressure, etc.) monitoring [8-9]. Bill Gates called the application of wearable devices for monitoring cardiac activity one of the most important technologies in 2019 because their usage allows to identify timely a critical state of a person and help to avoid consequences [10]. The existing wearable devices can be classified based on their function, appearance, proximity to the human body, and other parameters. Fig. 11 shows the classification of wearable devices according with investigations [2].

Fig. 11. Classification of wearable devices.

Wearable devices with various sets of sensors are already used in medicine, sports and security. They allow to notify about dangerous situations and to control physiological parameters and symptoms. This technology has changed the approach to health care, as it made it possible to monitor the patient state without hospitalization continuously. Through the activity tracker, medical professionals can access information about the patient's body temperature, heart rate, brain activity, muscle movement, and other data. Most of the existing wearable devices are aimed at multifunctional usage, including communication. Many of them have disadvantages, which can be associated with high cost, redundant or insufficient functionality, low accuracy, as well as with certain features of application. Most of them are focused on individual usage and self-diagnosis. The monitoring data is displayed and stored locally, so it can`t be used to transfer it to the information network of a medical institution for quick response and emergency assistance or prescribing treatment. The existing wearable devices use applications created by the manufacturer of the respective device and the possibilities of their customization are not provided.

Therefore, the development of a prototype of an inexpensive wearable device for remote monitoring of indicators of a person's physical condition, which can be integrated into the information system of a medical institution, is an urgent practical task.

3.2 Approache to Monitoring System Realization

It is known, that Wi-Fi technology refers to WLAN (Wireless Local Area Networks, IEEE 802.11) which are used for connection of personal computers and devices wirelessly (Fig. 12). The main advantages of WLAN are the prevalence of technologies, a significant coverage area and high data transfer rates [11].

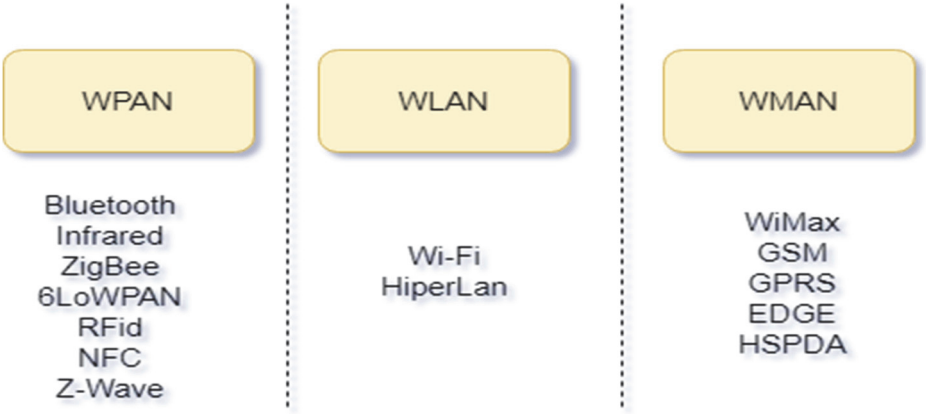
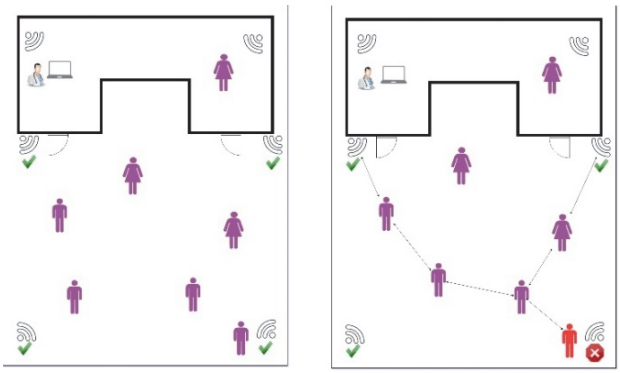


Fig. 12. Classification of wireless technologies.

Wi-Fi technology, due to its characteristics and existing implementation possibilities, can be successfully used in medical institutions to create an information system infrastructure.

Unlike a common Wi-Fi system, which consists of a router and several repeaters, the Mesh system provides efficient data transmission over a modular connection without losing speed. Also, the Mesh system provides better stability of work due to the creation of a "seamless" architecture because when using a router and repeater, you still need to disconnect and establish a new one with a new device. The advantages of the Mesh system, in particular the large coverage radius and seamless roaming of the network, determine the relevance of its application.

For trouble-free functioning of the monitoring system, all devices must always be online to ensure reliable data transfer to the server. Therefore, various solutions for automatic connection of access points to the Internet were analyzed, which can be simply implemented based on available resources. The first option is based on the usage of multiple access points throughout the hospital. The second option is based on the implementation of the Mesh network. The third option is a hybrid organization of the information system for collecting and transmitting medical data. It is more reliable due to the usage of an additional connection option in case the patient leaves the range of the Wi-Fi access point, or the access point does not work (Fig.13). As soon as the patient disconnects from the Wi-Fi hotspot, the monitor sends a request to neighboring devices to switch to Mesh mode to transfer data through the nearest neighbor connected to the Wi-Fi hotspot [8].



a b

Fig. 13. Variants of medical information system organization based on multiple routers (a) and Mesh network (b).

The implementation of a hybrid option based on a seamless network and the usage of perspective algorithms for network self-organizing and finding the best ways to transfer data provides some advantages. They are coverage of the entire area of activity of customers, as well as a stable signal for data transmission, which prevents customers from disconnecting when moving out of a zone [8].

The UML diagram of the remote monitoring system functioning includes the server and the wearable device (Fig. 14). On the server-side, the registration of the patient is performed, as well as the monitoring data is stored and visualized. In addition to the hospital server itself, the cloud service can be used as a server. On the side of the wearable device, its connection to the network, reading and processing of data, visualization and sending data to the server are performed.

3.3 Development of Software and Hardware for Paramedical Bracelet

A convenient form of a wearable device for monitoring the parameters of a patient's physical condition is a bracelet (Fig. 15) [8]. It provides measurement of the pulse and blood oxygen saturation, determination of the change in the patient's position, graphical and audible notification and transfer data to the server.

A wearable device consists of hardware and software that ensures the functioning of the system. Hardware integrates mechanical and electronic parts. The mechanical part provides placement and protection of the electronic part components, as well as fastening to the user's hand.

The electronic part consists of the following components:

- Wi-Fi module WeMos D1 mini on the ESP8266 chip that provides calculations, storage, and execution of the program, implementation of Wi-Fi connection, and receiving program updates via a wireless connection.
- Battery and charger (Li-Po 3.7V 2000mAg 103450, micro USB based on TP4056 controller) that provides autonomous power supply and recharging.

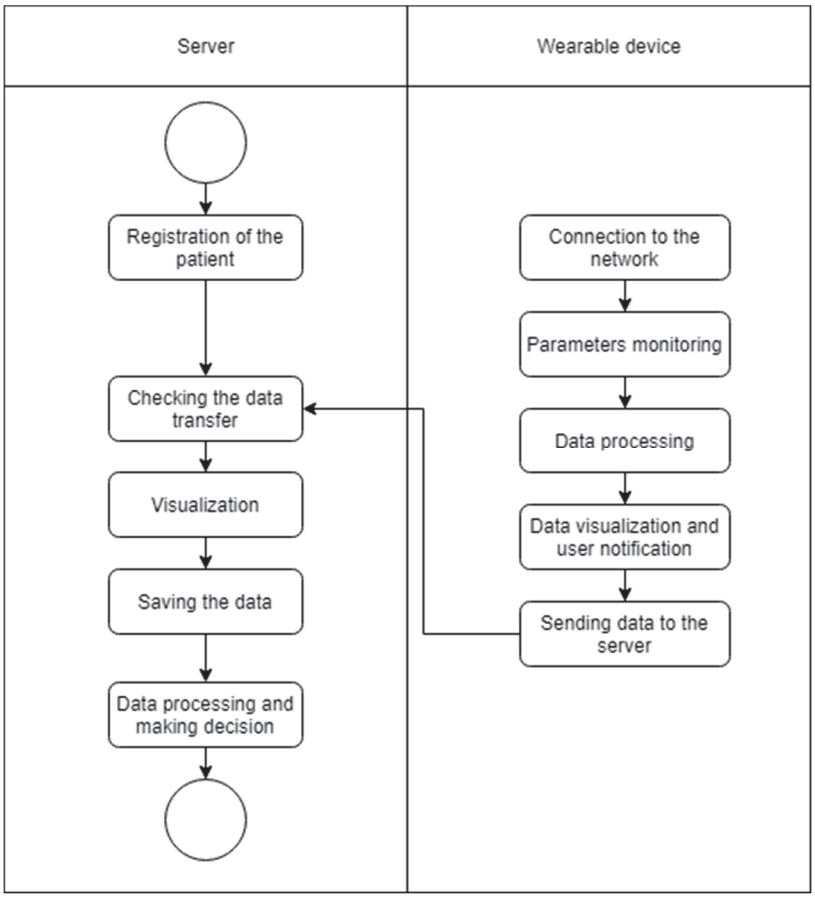


Fig. 14. UML diagram of the remote monitoring system functioning.

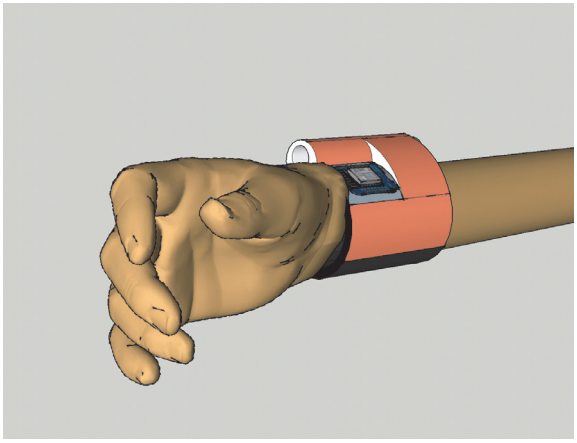


Fig. 15. Concept of paramedical bracelet.

- Integrated heart rate and pulse oximetry sensor (MAX 30100) to measure pulse values and oxygen content in the user's blood.
- Integrated accelerometer and gyroscope sensor (GY-521) to determine the position of the device in space to use this data for user's movements monitoring.
- Display (OLED screen 0.91" l2C 128x32) for data vizualization.
- Passive buzzer (KY-006) for audible alerts.

It should be noted that there are still many possibilities to modify the wearable device. For example, it is possible to replace the indicated components with the following: ESP32 module with OLED display and 18650 compartment; accelerometer and gyroscope MPU-6050 6DOF module; integrated sensor MAX30102; Liitokala NCR18650B Protected 3400mAh battery with protection. Even with the replacement of components, the total cost will be around 25 euros.

The connection diagram of the main components created in the fritzing environment is shown in Fig. 16. A feature of the circuit is the connection of sensors via the l2C protocol. It is an asymmetric serial bus for communication between integrated circuits that uses two bi-directional communication lines (SDA and SCL), which are used to connect low-speed peripheral components to processors and microcontrollers, thus reducing the number of wires. Another feature is the use of different sensors with different addresses, which makes it possible not to use an additional wire for setting the module (sensor) address.

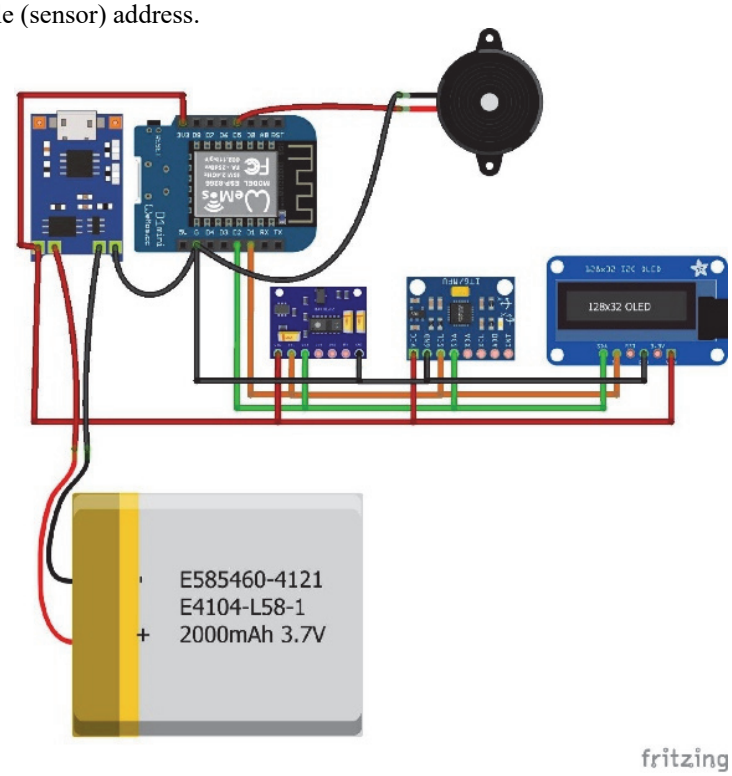


Fig. 16. Virtual prototype of the sheme in Fritzing.

3D models of all parts (body, cover, fasteners), components layout, and general assembly were developed in the CAD system (Fig. 17). Project shortcomings were identified and the construction was optimized based on the investigation of the virtual prototype. Based on the 3D models, files in the STL format were obtained, that are used for printing on a 3D printer. The physical prototype of the mechanical part of the designed wearable device was produced as a result of options setting in Simplify 3D software and further printing on the 3D printer Leapfrog Creatr HS. ABS plastic has been used as the printing material. The case and the cover with some electronic components arranged inside are presented in Fig. 18.

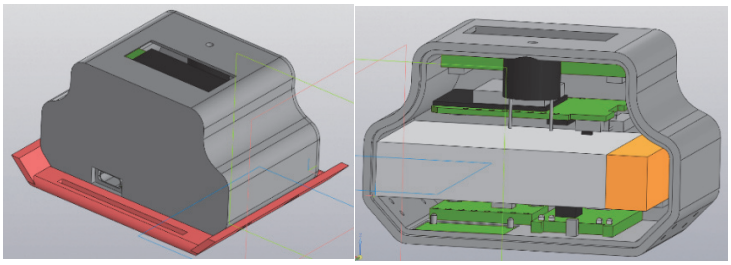


Fig. 17. Virtual prototype of bracelet parts.

Fig. 18. Physical prototype of body and cover.

The Arduino IDE was used for software development. The program code is loaded into the ESP8266 Wi-Fi module microcontroller as firmware. The algorithm of the program combines several steps:

- Initialization of variables with the addresses of the sensors connected to the I2C data bus.
- Initialization of the OLED screen to display information.
- Declaration of variables for displaying parameters on the screen.
- Initialization and start of I2C bus at a speed of 400kHz.
- Start of the display usage.
- Reading data from sensors, their processing and visualization.
- User notification about alerts.
- Data transfer to the server.

In each cycle, there is a continuous reading of data from the sensors. All data are saved in separate variables. When all data are identified, a JSON request is formed to the server to save and process the data. In the future, the possibility of remote updating can be implemented (to update the code and download additional functionality) by downloading new versions of the program code via a Wi-Fi network.

Modern cloud services can be used for solving the task of information storage and visualization. As such a service, free service Freeboard.io can be recommended, which is built on a secure, highly productive enterprise-class cloud system. It allows to create interactive dashboards and real-time visualizations using an intuitive interface conveniently and simply (Fig. 19). Using a Wi-Fi module based on ESP8266 and HTTP requests the information is transmitted to the cloud service, visualized, and stored. In the future, it is possible to implement an intelligent analysis of the received monitoring data to support the decision-making of medical personnel in emergencies or when prescribing treatment for a patient.

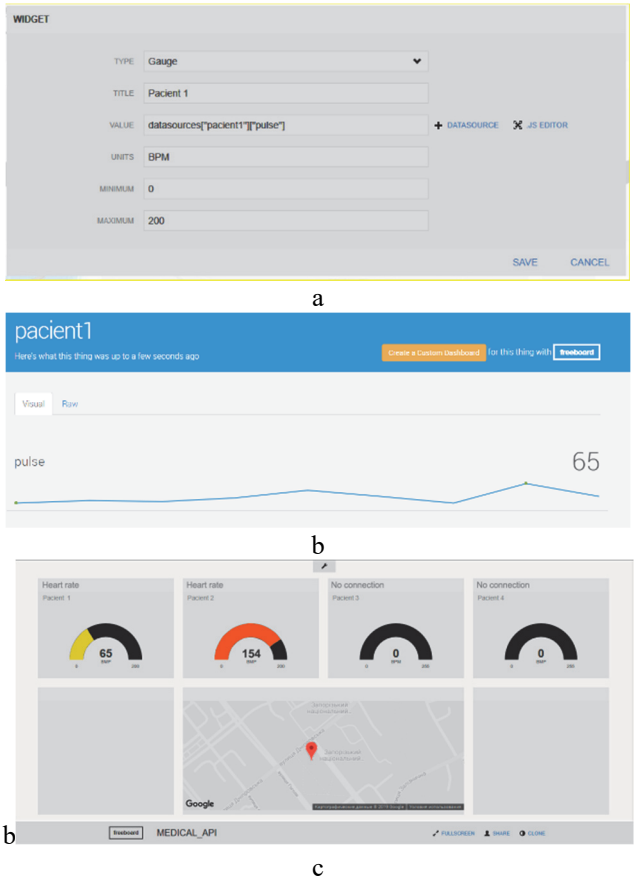


Fig. 19. Dashboards for patient registration (a), parameters visualizations (b) and monitoring (c) in Freeboard.io.

4 Conclusion

In this chapter, the features of the computer-aided design of biomedical systems (cyber prostheses and wearable devices for monitoring the indicators of the physical condition of patients) were investigated.

It is shown that the application of prototyping technologies allows to improve the efficiency of the development process by reducing the design time, the number of errors and provides opportunities to improve the functional characteristics, shape, weight, and dimensions of the designed biomedical products.

References

1. Dorrier, J.: Bionic athletes with exoskeletons, robotic limbs, and brain-control devices to compete in cybathlon, https://singularityhub.com/2014/04/23/bionic-athletes-with-exoskeletons-robotic-limbs-and-brain-control-devices-to-compete-in-2016-cybathlon, last accessed 2020/12/10.
2. Mardonova, M., Choi Y.: Review of wearable device technology and its applications to the mining industry. Energies 11(3), 14 p. (2018).
3. Parkhomenko, A., Gladkova, O., Zalyubovskiy, Y.: Investigation and realization of prototyping technologies for robotic-prostheses computer aided design. In: Proceedings of XV International Conference on The Experience of Designing and Application of CAD Systems in Microelectronics (CADSM 2019), pp. 7/5-7/8. IEEE, Los Alamitos, USA (2019).
4. Liou, W.F.: Rapid prototyping and engineering applications: a toolbox for prototype development. CRC Press Taylor & Francis Group Boca Raton, 592 p. (2007).
5. McEwen, A., Cassimally, H.: Designing the Internet of things. Wiley, 338 p. (2014).
6. Belter, T.J., Segil, L.J., Aaron, M.D., Weir, F.R.: Mechanical design and performance specifications of anthropomorphic prosthetic hands. Journal of Rehabilitation Research & Development, 599-618 (2013).
7. Gaget, L.: 2018 Top 8 of the best parametric modeling software in 2020, https://www.sculpteo.com/blog/2018/03/07/top-8-of-the-best-parametric-modeling-software/, last accessed 2020/12/10.
8. Parkhomenko, A., Presaizen, Y., Gladkova, O., Tulenkov, A., Kalinina M.: Remote heart rate monitoring of the hospital patients. In: Proceedings of the 2019 10th IEEE International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS 2019) pp. 991-996. IEEE, Los Alamitos, USA (2019).
9. Parkhomenko, A., Volynska, A., Zalyubovskiy, Y., Parkhomenko, A., Kalinina M.: Method of monitoring of young athletes' physical state indicators based on wearable devices usage. In: CEUR Workshop Proceedings, 2608, pp. 436-449. CEUR, M. Jeusfeld c/o Redaktion Sun SITE, Informatik V, RWTH Aachen, Germany (2020).
10. How we'll invent the future, by Bill Gates. MIT technology Review, February, 2019, https://www.technologyreview.com/lists/technologies/2019/, last accessed 2020/12/10.
11. Tulenkov, A., Parkhomenko, A., Sokolyanskii, A., Stepanenko, A., Zalyubovskiy, Y.: The features of wireless technologies application for Smart House systems. In: Proceedings of the 4th IEEE International Symposium on Wireless Systems within the IEEE International Conferences on Intelligent Data Acquisition and Advanced Computing Systems, pp.1-6. IEEE, Los Alamitos, USA (2018).

System of three-dimensional human face images formation for plastic and reconstructive medicine

Sergii Pavlov[1] [0000-0002-0051-5560], Oleg Avrunin[2][0000-0002-6312-687X], Oleksandr Hrushko[3][0000-0001-5551-375X], Yana Nosova[4][0000-0003-4310-5833], Natalia Shushlyapina[5][0000-0001-8958-3270]
[1, 3] Vinnytsia National Technical University, 21021, Vinnytsia, Khmelnytske shose 95, Ukraine
[2]Kharkiv National University of Radio Electronics, 61166, Nauki Ave. 14, Kharkiv, Ukraine
[4,5]Kharkiv National Medical University, 61022, Nauki Ave. 4, Kharkiv, Ukraine
[1]psv@vntu.edu.ua
[2]oleh.avrunin@nure.ua
[3]grushko1alex@gmail.com
[4]yana.nosova@nure.ua
[5]schusha75@ukr.net

Abstract. Improving the methods and means of computer planning of functional and aesthetic interventions on the human face based on realistic 3D visualization methods is an urgent task of modern biomedical engineering. Despite a sufficient amount of work in the field of medical imaging and image processing, a more accurate and realistic presentation of data is needed for imaging systems aimed at modifying the initial data to predict the results of surgery. Software for 3D visualization in computer modeling of functional and aesthetic surgical interventions should ensure the solution of the following set of tasks: initial computer processing of tomographic data and 3D surface scanning data; selection of a visualization window and tone mapping; segmentation of anatomical structures on tomographic images; determination of the contours of segmented objects on tomographic images; construction and visualization of polygonal models of the anatomical structures of the investigated area.

Keywords: Surface Scanning, Tomographic Images, Visualization.

1 Introduction

An increase in the efficiency of preoperative planning of plastic interventions can be achieved through the use of 3D visualization methods and modeling of the deformation of structures subject to plastic correction with maximum realism. Initial data for such visualization can be obtained as a result of tomographic studies of internal structures and surface 3D scanning. In accordance with this, software systems should be developed that allow for realistic visualization and correction of the obtained 3D images, taking into account the biophysical properties of the operated area [1,2].

The wide application of this approach requires the use of universal (non-specialized) computer systems. This achieves the flexibility of improving and expanding the capabilities of software for planning operations in rhino surgery [3,4,5].

Creation of three-dimensional computer models of the human head precedes the stage of modelling changes in various anatomical parts of the head in order to obtain an optimal model corresponding to the wishes of the patient and the capabilities of the surgeon [6,7,8]. Such modeling requires the development of segmentation algorithms, complex or separate visualization and virtual deformation of anatomical structures.

2 Analysis of literary sources

Improving the methods and means of computer planning of functional and aesthetic interventions on the human face based on realistic 3D visualization methods is an urgent task of modern biomedical engineering.

The purpose of considering the features of the developed models and methods is to create a concept for the presentation, processing and analysis of tomographic data for computer planning systems in functional-aesthetic rhinosurgery.

Analysis of literature sources showed that, despite a sufficient number of works in the field of medical imaging [1,8] and image processing [9,10,11], a more accurate and realistic presentation of data is needed for imaging systems aimed at modifying the initial data to predict the results of surgery [12,13,14].

Software for 3D visualization in computer modeling of functional and aesthetic surgical interventions should provide a solution to the following set of tasks:

- initial computer processing of tomographic data and 3D surface scanning data;
- selection of the visualization window and carrying out tone mapping;
- segmentation of anatomical structures on tomographic images;
- determination of the contours of segmented objects on tomographic images;
- construction and visualization of polygonal models of the anatomical structures of the investigated area [15,16,17].

The main source of input data for the operation of the software for planning functional-aesthetic surgeries is a set of images of tomographic slices (DataSet) [1], represented as a series of files in DICOM (Digital Imaging and Communications in Medicine) format. The DICOM format is a standard for the generation, storage and transmission of radiological images and other medical information. This format, in addition to the direct results of an introscopic examination, contains data about the patient, the conditions of the study, data about the medical institution, the operators of medical equipment, the technical characteristics of the introscopic devices, the values of the physical fields underlying the study, etc. The standard enables digital interconnection between different diagnostic and therapeutic equipment from different manufacturers.

In a DICOM file, information about the study performed is stored as a set of information elements - special information structures that can store data about the patient, the study being carried out, and diagnostic equipment. The order of the elements of the tomographic dataset is such that as the element address increases, first the voxel number

in the line increases, then the line number, and then the number of the slice in the axial projection [18,19,20].

3 Preliminary data preparation

Computed tomography images are gray-scale. Each slice element determines the average value of X-ray absorption in the corresponding area of the object under study during tomography. Such a 2D dataset can also be represented as anarray of brightnessor discrete intensity function, where x, y, and z are integer variables defining voxel coordinates. The domain of the function V is:

$$\begin{cases} x \in [0..w-1], \\ y \in [0..h-1], \\ z \in [0..d-1]; \end{cases} \tag{1}$$

where w, h, d are the sizes of the dataset in voxels.

The function V takes non-negative integer values in the range $[0 \ldots V_{MAX}]$. The value V_{MAX} depends on the bit depth N of the dataset and is determined by the formula

$$V_{MAX} = 2^N - 1 \tag{2}$$

In turn, the number of bits allocated for encoding the brightness of one element of the tomographic raster N usually takes values from 8 to 16, depending on the tomograph and the study protocol. Thus, one voxel can occupy 1 or 2 bytes in computer memory, which significantly complicates the effective processing of tomographic data and requires writing software implementations for both voxel bit widths and using generalized programming principles [21,22,23].

If the number of bits per one voxel of the initial tomographic data is not a multiple of 8, then it is advisable to round this value to the nearest higher multiple of 8, since This approach will greatly simplify the addressing of the brightness values of the voxels of the dataset and will make the software processing of tomographic data faster.

Storing tomographic slices in PC memory can require significant amounts of RAM. The content of all voxels in one memory block is preferable from the point of view of convenience and speed of programmatic access to dataset elements. However, this approach with sequential storage of voxels is not always possible due to the fact that the program may fail when allocating memory necessary to store the entire dataset in a single block. Therefore, when developing software for processing tomographic slices, it is advisable to implement the data loading algorithm in such a way that, in the event of a failure on a new memory block requested from the operating system, memory allocation is performed for each slice separately [24,25,26].

The most used strategies for storing tomographic data in RAM are slice-by-slice storage of elements and storage of elements in cubic blocks, as shown in Fig. 1.

The slice-by-cut arrangement of elements in memory involves storing raster lines sequentially one after another. The pixels in a row are stored linearly in ascending order of the x coordinate in the row. The element next to the last pixel in a line in memory is the first element of the next line. The advantage of this way of storing voxels is the simplicity of addressing an arbitrary element of the dataset.

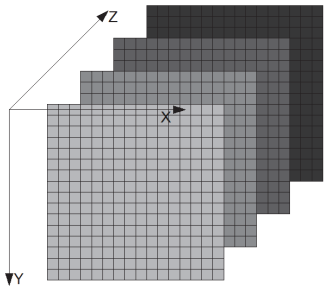


Fig. 1. Cut-by-slice presentation of tomographic study data (*w*=*h*=16, number of slices - 4).

The address of the voxel in RAM, depending on its coordinates (*x*, *y*, *z*), is easily calculated in accordance with the following expression:

$$Pnt_{AxVox}(x,y,z) = Pnt_{Slice}(z) + (y \cdot w + x) \cdot C(b) \tag{3}$$

where $Pnt_{Slice}(z)$ is the function that determines the memory address of the beginning of the *z*-th axial slice;
w - the width of the dataset in voxels;
b is the number of bits used to encode one voxel;
C (b) is a function that returns the number of bytes required to store *b* bits, i.e. rounding *b* to *8* upwards, and

$$C(b) = \left(\left\lfloor \frac{(b-1)}{8} \right\rfloor + 1 \right) \tag{4}$$

When storing a dataset as a set of cubic blocks (see Fig. 2), all voxels are combined into smaller volumes with dimensions (D_X,D_Y,D_Z). Reading voxels located inside a block is also performed by progressive scanning, but within the block, not a slice. Access to the elements of the dataset with a block structure is performed in accordance with the expression

$$Pnt_{Vox}(x,y,z) = Pnt_{Slice}\left(\left\lfloor \frac{z}{D_Z} \right\rfloor \right) + \left(\left\lfloor \frac{y}{D_Y} \right\rfloor \cdot \left\lceil \frac{w}{D_X} \right\rceil + \left\lfloor \frac{x}{D_X} \right\rfloor \right) \cdot D_X \cdot D_Y \cdot D_Z \cdot C(b) + \begin{pmatrix} mod(z,D_Z) \cdot D_X \cdot D_Y + \\ + mod(y,D_Y) \cdot D_X + \\ + mod(x,D_X) \end{pmatrix} \cdot C(b) \tag{5}$$

where *mod*(*i*,*j*) is a function that returns the remainder of dividing *i* by *j*.

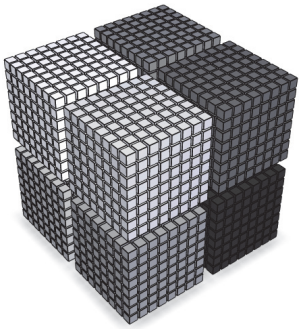


Fig. 2. Block storage of tomographic dataset voxels ($D_X = D_Y = D_Z = 8$).

As can be seen from expressions (2) and (4), addressing an arbitrary voxel of a dataset when using a block strategy for storing bulk data requires more computations than with other storage methods, but for some tasks (for example, tracing or casting rays), this approach shows higher productivity. The general increase in the performance of algorithms that actively use the operation of fetching voxel values when dividing a dataset into blocks occurs primarily due to more efficient use of the caching mechanism by the central processor of cells stored in RAM [27,28,29].

Note also that it is possible to locally reduce the bit depth of some entire dataset blocks. So you can reduce the number of bits, which encodes the brightness of the voxels of this block, to N if the following conditions are met:

$$\begin{cases} V_{MAX} - V_{MIN} \le 2^N - 1, \\ N < M. \end{cases} \tag{6}$$

This approach also makes it possible to somewhat reduce the average access time to the elements of a three-dimensional array of tomographic data and, as a rule, significantly reduce the amount of memory required for storing the tomographic dataset as a whole. A three-dimensional array of luminosities using the described approach to storing elements is called a sparse dataset [30,31].

Storing tomographic data in the form of a sparse dataset and its corresponding software processing are in some cases the only possible effective way to work with datasets characterized by large sizes or in conditions of a small available amount of RAM.

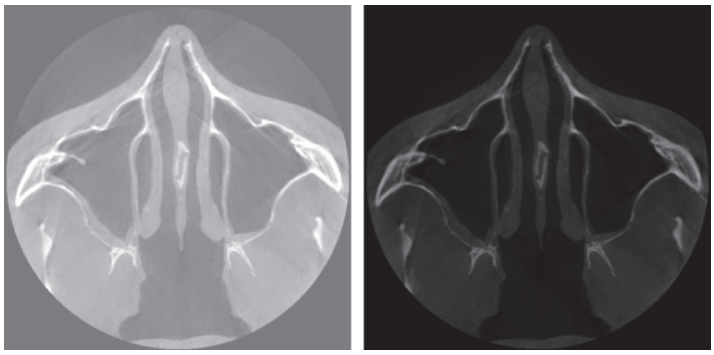
4 Tone Mapping

The next step is tone mapping, which allows more efficient use of the dynamic range of tomographic imaging and promotes better perception of the obtained images by the operator, allows better differentiation in the image of tissues of various kinds that differ in density, and, consequently, in brightness in the output images.

In the case of linear transformation of brightness ranges during tone mapping, the brightness of an image element after I_D compression depends on the brightness of the corresponding pixel I_S in the original image according to the expression

$$I_D = W_{DL} + \frac{(I_S - W_{SL}) \cdot (W_{DR} - W_{DL})}{W_{SR} - W_{SL}} \quad (7)$$

where $[W_{DL} ... W_{DR}]$ is the brightness range (brightness window) used to display the image to which the tonal range has been applied;
$[W_{SL} ... W_{SR}]$ - the tonal range of the original image.
The results of the method according to formula (7) (see Fig. 3).



a b

Fig. 3. Tomographic section of the patient's head: a) before tone mapping ($W_{SL} = 30920$, $W_{SR} = 40375$); b) after tone mapping ($W_{DL} = 35760$, $W_{DR} = 44760$).

5 Segmentation of Anatomical Structures

In the automated analysis of tomographic data, one of the most important operations preceding the main stages of processing the original images is the task of segmentation - determining which image pixels carry information about the amount of X-ray absorption by the object under study, and which do not belong to it. It is advisable to perform this operation by recursive filling of connected pixels that have passed binarization with a lower threshold, which will be discussed in detail below.

During the segmentation and classification of tomographic slices, an additional two-dimensional array A is created for each tomographic image with the same dimensions as the raster being processed. Before the start of segmentation, all elements of this array are initialized to zero. The result of the classification of any voxel in the dataset is entering one of the nonzero values T_1, T_2 or T_3 into the cell of array A. These values are determined by the type of region to which this voxel belongs, and as a result, this voxel becomes an element of one of the following segmented sets D_1, D_2, D_3, for which

$$(x, y, z) \in D_i, \quad \text{if } A(x, y, z) = T_i \quad (i = 1, 2, 3). \quad (8)$$

When binarization with a lower threshold, if the brightness of a given voxel $V(x, y, z)$ is less than or equal to a certain threshold value (V_1), then the value of the element in

array A remains unchanged. If the brightness is greater than V_1, then $A(x,y,z)=T_1$. Thus, for binarization with a lower threshold, the following expression is valid:

$$A(x,y.z)=\begin{cases}T_1, & V(x,y,z)>V_1\\ A(x,y,z), & V(x,y,z)\le V_1\end{cases} \quad (9)$$

After isolating the background pixels of the slices, the stage of isolating and classifying the tissues of the human head on the images of tomographic slices begins using the operation of filling connected pixels that satisfy the following binarization condition with an upper threshold:

$$A(x,y.z)=\begin{cases}T_2, & V(x,y,z)<V_2\\ A(x,y,z), & V(x,y,z)\ge V_2\end{cases} \quad (10)$$

where V_2 is some threshold value.

Another type of threshold classification that is used at the stage of segmentation and classification of tissues imaged on tomographic sections is binarization with double limitation in accordance with the expression:

$$A(x,y,z)=\begin{cases}A(x,y,z), & V(x,y,z)\le V_1\\ T_3, & V_2>V(x,y,z)>V_1\\ A(x,y,z), & V(x,y,z)\ge V_2\end{cases} \quad (11)$$

As a result of this operation, those elements of the raster are selected, the brightness of which lies in the range. This operation can be useful, for example, when recognizing soft tissues on tomographic sections.

The first stage of the fill operation is to determine the initial (seed) pixels, starting from which the operation of growing the segmented area will take place. In the case of automated background selection on a slice image, such points are usually selected pixels located at the edges and corners of the slice, provided that their brightness is not higher than the threshold value.

If it is necessary to fill the pixels of the biological object under study, the operator chooses pixels that exactly belong to the biological object as starting points. Segmentation of the airway anatomical structures, such as the upper airways, usually requires manual setting of the initial filling points in an interactive mode, as well as the definition of special limiters that do not allow many segmented pixels to expand through specified areas of the image (some nasal passages, anastomosis) that satisfy the threshold condition. classification.

The specified starting points of the fill are added to a special data structure that has a stack organization. This structure is necessary, since during each iteration of the filling cycle, for each processed element there can be up to 4 neighboring elements in the case of filling a four-connected area, or up to 8 neighbors in the case of processing an eight-connected environment, which are to be processed in subsequent iterations of the

algorithm. By its principle of operation, a stack is similar to a queue with the only difference that the first element added to the stack will be popped last, and the last added element, on the contrary, will be the first. This difference between a queue and a stack makes it easier and more efficient to use the latter for storing image elements waiting to be processed by this segmentation algorithm with the same final result.

In case of multi-slice segmentation of a human head in a tomographic dataset, in addition to 4 or 8 pixels of this slice, which are adjacent to the considered pixel, it is also necessary to check the values of voxels having the same two-dimensional coordinates, but located on adjacent slices. This approach is used when pouring on sections of head tissue and, as a rule, makes it possible to select all the space occupied by the patient's head tissue at one time.

6 Voxel Rendering and Surface Rendering

A simpler implementation of the fill function is also possible, recursively calling itself when checking for connected pixels. If the traversal of connected areas is implemented using a regular recursive call of the subroutine for checking neighboring pixels, then with each call, the address from which this function was called, its arguments, local variables, and the current values of general-purpose registers are saved on the stack. This leads to an avalanche-like increase in the amount of used memory of the system stack and, ultimately, can lead to its overflow and, consequently, to the abnormal termination of the program. Therefore, the use of the filling algorithm using recursive calls is fraught with overflow of the system stack at a large depth of recursion nesting and is not suitable for stable and fast segmentation of large and medium volumes of tomographic data.

It is possible to speed up the segmentation procedure for large tomographic volumes even more by using the principle of row-wise filling. The main difference between this method and the filling of connected pixels is that during one iteration of the cycle, a series of pixels are filled sequentially to the left and right of the element processed during this iteration of the fill cycle. This approach is much more efficient by processing more pixels per loop iteration, reducing the number of repetitions, and also by reducing the amount of stack used. Below is table 1 containing data reflecting the pouring processes, the stages of which are shown in Fig. 4 and 5.

The characteristics given in Table 1 prove the significant superiority of the row-by-row fill method over the four-connected pixel fill method for such tasks of processing tomographic images as highlighting the background in a tomographic image, as well as segmenting the tissues of the head and upper respiratory tract. Line-by-line filling is the most optimal not only for reasons of the general performance of the algorithm, but also because of the significantly less use of additional memory. In addition, the line-by-line pixel traversal also improves the speed of such segmentation by taking into account the organization of the CPU's cache memory. The stage of identifying the contours of anatomical structures on tomographic images is necessary for a clear morphological separation of models of these structures [32,33,34].

Table 1. Comparison of the characteristics that determine the effectiveness of the threshold segmentation algorithms when filling the background of the test slicen.

	Segmented area,%	Number of fill iterations performed	The amount of used stack memory, Bytes	Algorithm execution time, μs
Quadruple fill	10	8414	33528	1370
	40	33731	133812	3230
	70	63778	215176	5174
	100	128108	159408	9507
Line-by-line fill	10	41	12	214
	40	510	30	930
	70	1045	36	1445
	100	1225	12	1848

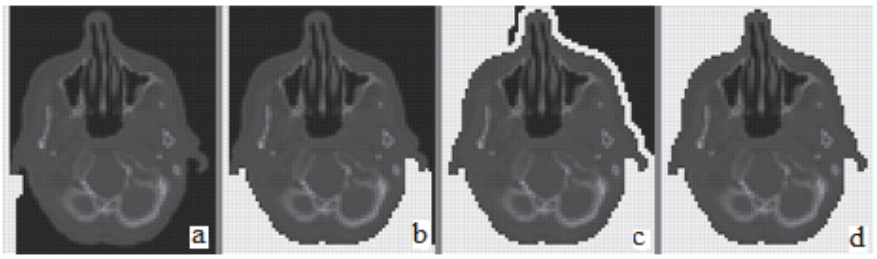


Fig. 4. Stages of selecting the background of a tomographic slice by the method of filling a four-connected area: a) 10% of the background is filled; b) 40%; c) 70%; d) 100%.

Fig. 5. Stages of separation of tissues of the human head on a tomographic section: a) 10% of the pixels of the head are filled in; b) 40%; c) 70%; d) 100%.

Algorithms of this kind make it possible to clearly distinguish morphologically on the sections of biological structures (objects) after their preliminary classification according to the necessary attribute and to form a set of segments describing a closed

broken line of the contour. The information obtained at this stage about the contours of objects subsequently forms the basis of the wire 3D model.

The marking of segmented areas is necessary for the subsequent selection of the contours of the selected areas and preparation for building wire models.

At this stage, each element of the auxiliary buffer is filled according to the type of area to which it belongs. Now, in this buffer, it is necessary to select separate areas and number the segmented voxels, so that later it will be possible to determine to which area each segmented element of the tomographic data array belongs.

Further, upon completion of the stage of marking the segmented parts of the image, the pixels located on the inner border of the segmented areas are selected. Let the matrix F have the following form:

$$F = \begin{pmatrix} 1 & 1 & 1 \\ 1 & -8 & 1 \\ 1 & 1 & 1 \end{pmatrix} \tag{12}$$

A pixel is considered to be on the border of the segmented area if, during the check in the auxiliary array of its eight-connected environment by the sliding window F, the result is nonzero. This condition is true if at least one of the eight pixels surrounding the given one does not belong to the processed area.

In order not to start another additional array for storing the results of the operation of selecting the boundaries of the segmented regions, the edge pixels are marked in the previously used auxiliary array by setting a certain bit to one state. The setting of a bit can be done either by performing the "logical or" operation with the number C, or by performing the operation:

$$A'(x,y,z) = A(x,y,z) + C \tag{13}$$

moreover $A(x,y,z) < C$.

To set this bit not to affect the bits of the auxiliary array, which are responsible for encoding the segmented area number, the value of the number C must be calculated in accordance with the following formula:

$$C = 2^{\lceil \log_2 a \rceil} \tag{14}$$

where a is the number of segmented areas in the tomographic image.

In order to build a contour model in the future, it is necessary to find a set of line segments that describe the boundaries of the selected areas. This stage begins with the creation of an array of chain codes, into which a bit mask is written, corresponding to the presence of border pixels in the eight-connected region of a certain raster element (see Fig. 6).

An array of chained codes is another auxiliary two-dimensional array, the number of rows and columns in which is the same as in the original tomographic raster. To store information about the presence or absence of a contour pixel in a certain position of an eight-connected environment, 1 bit of information is enough. Based on this, the size of

the element of the chain codes buffer from the point of view of minimizing the use of the RAM resource should be chosen equal to 1 byte.

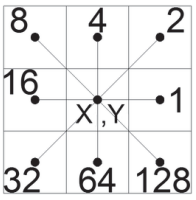


Fig. 6. The value of the bit mask of the element of chain codes, depending on the location of the associated edge pixels.

It is possible to determine the value of the chain code for a certain element of the tomographic image, based on its eight-connected environment, in accordance with the following expression:

$$B(x,y,z)=\sum_{i=0}^{7}2^{\left\lfloor \frac{ang(x_{i-x},y_i-y)\cdot 4}{\pi} \right\rfloor}\cdot\left\lfloor \frac{A(x_i,y_i,z)}{C} \right\rfloor \tag{15}$$

where i is the number of the pixel (x_i, y_i, z) in the array of the eight-connected environment of the considered pixel (x, y, z) (see Fig. 6);

$ang(dx,dy)$ - function, which calculates the angle, plotted in the counterclockwise direction, between the segment given by the points (x, y), (x_i, y_i) and the horizontal ray emitted from the point (x, y) in the positive direction along the X axis;

$$ang(dx,dy)=\begin{cases} arctg\dfrac{dy}{dx}, dx>0, dy>0 \\ arctg\dfrac{dy}{dx}+\pi, dx<0 \\ arctg\dfrac{dy}{dx}+2\pi, dx>0, dy<0 \end{cases} \tag{16}$$

where $dx=x_i-x$; $dy=y_i-y$.

The set of connected segments that form a polyline of a contour is obtained as a result of sequential traversal of nonzero elements of the array of chain codes. If a certain pixel belongs to the contour (see Fig. 7, a) and has no more than two contour neighbors in an eight-connected area, then the centers of these pixels are connected by segments, otherwise the traversal occurs in a certain angular direction, for example, clockwise (see Fig. 7, c) in order to avoid the appearance of transverse lines that do not define the border of the segmented area (see Fig. 7, b). The resulting contour can be simplified by combining into one segment a sequential series of connected segments (see Fig. 7, d), having the same direction.

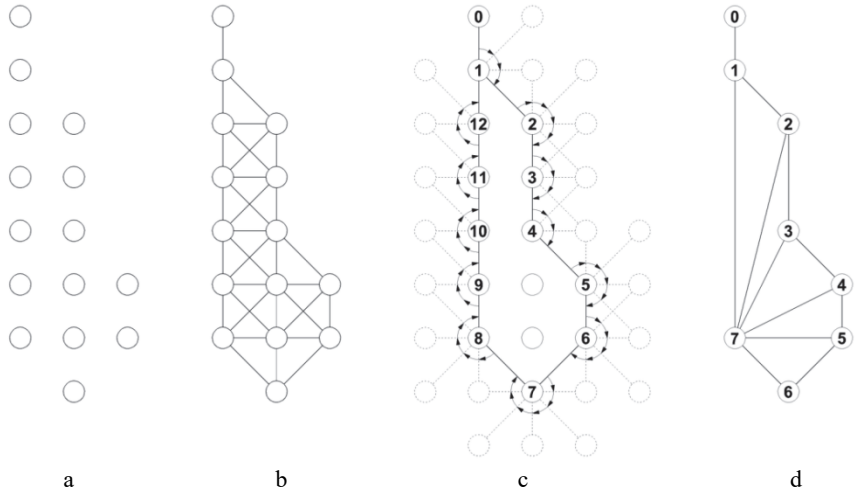


Fig. 7. Construction of a polyline based on the data of contour pixelsa) original contour pixels; b) incorrect connection of the contour nodes; c) angular traversal of contour pixels; d) the contour after its simplification and triangulation

It is considered that one segment of the contour given by the points (x_1, y_1) and (x_2, y_2) is a continuation of another segment with coordinates (x_2, y_2), (x_3, y_3) if the condition specified in the expression

$$\frac{x_2 - x_1}{y_2 - y_1} = \frac{x_3 - x_2}{y_3 - y_2} \tag{17}$$

In this way, as a rule, it is possible to achieve a significant reduction in the geometric complexity of the contour of the segmented area, which reduces the amount of memory required for its storage, and accelerates its subsequent processing.

Having received a closed outline of the selected area, you can calculate such its characteristics as the perimeter and area. In the context of planning surgical interventions on the upper respiratory tract and the human face, this can be useful, for example, to assess the geometry of the anatomical structures of the face or to calculate the functional parameters of the patient's nasal breathing [15,16]. An illustration of the results of the algorithm for segmentation of the contours of the airways of the upper respiratory tract is shown in Fig. 8.

The advantage of the algorithm created in the work is, first of all, the lower complexity of the contours of the selected areas in the images, which simplifies and accelerates the further processing of these images, in particular, their specified contours.

By presenting a set of contours of different slices of one tomographic dataset in three-dimensional space, it is possible to form a spatial wire model (see Fig. 9, a).

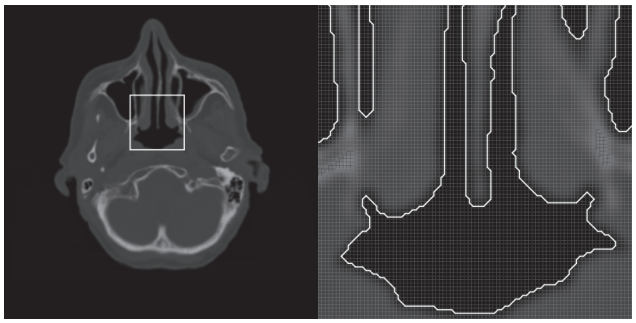
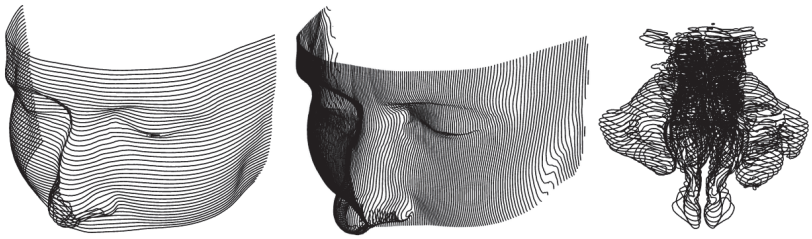


Fig. 8. Selection of the contours of the upper respiratory tract on the tomographic image: a) the processed area of the original image; b) the result of constructing a contour for a given area of the image using the proposed method

Operations such as segmentation of anatomical structures and the selection of their contours often need to be performed not only in the axial, but also in other projections (see Fig. 9, b, Fig. 9, c). In this case, the operation of multiplanar reconstruction is performed, in which the initial set of slices (usually axial) is recalculated into the same set of slices, but corresponding to a different projection (see Fig. 10). In particular, this is necessary when calculating the geometric characteristics of the air-conducting paths of the nose based on the results of processing a series of frontal tomographic sections [13,14].



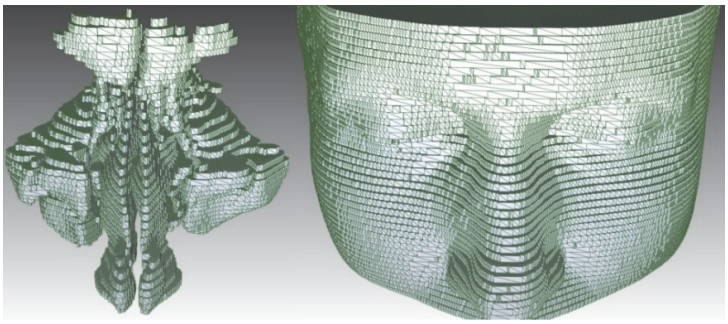
a b c

Fig. 9. Spatial contour models obtained as a result of cut-through processing of tomographic data: a) a model of the face, built from axial slices; b) a model of the face, built on sagittal sections; c) model of the upper respiratory tract.

One of the widely used methods of visualization of the studied biological objects based on the results of computed tomography is the display of voxel polygonal models. The voxel model is a three-dimensional array of the simplest volumes of the same size. When rendering a voxel model of tomographic data, each segmented element of the tomographic volume is displayed as a separate small cube (rectangular parallelepiped if the lengths of the voxel sides are not equal to one), each face of which consists of 2 triangles. Voxels that are not related to the biological object are ignored and are not displayed in 3D space. Thus, it turns out that in order to visualize one voxel in three-dimensional space, it is necessary to draw 12 triangles built on 8 vertices, which makes

a rather serious computational load on the block of the used graphics library, which is responsible for vertex transformations and lighting, when rendering large volumes of tomographic data. As an example, voxel models of the upper respiratory tract (see Fig. 11, a) and the human face (see Fig. 11, b) are given.

Fig. 10. Presentation of computed tomography data, segmentation results and contouring in three main projections.



a b

Fig. 11. Voxel models: a) model of the upper respiratory tract; b) model of the human facial region.

To render a voxel model in real time with a frame rate acceptable for visual perception, as a rule, it is necessary to perform preliminary optimization aimed at simplifying the geometry of the model, reducing the number of simple geometric shapes that describe the model (primitives). This simplification of geometry implies the elimination of individual invisible faces and entire voxels from the model, as well as combining several faces of adjacent voxels into one face (see Fig. 12).

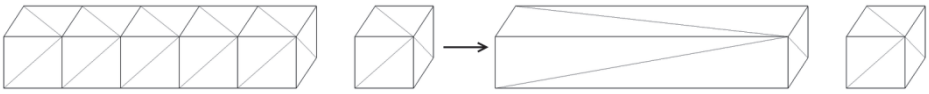


Fig. 12. Simplification of the geometric component of the voxel model by combining the faces of adjacent voxels.

The voxel representation model is quite informative for the purposes of medical diagnostics and planning of operations. It most closely matches the input tomographic data and allows you to directly calculate the geometric characteristics of the anatomical objects of the facial region and the upper respiratory tract of a person, visually assess the spatial relationships between the air cavities, nasal passages, fistulas and surrounding structures, and display the localization of the location of pathologically altered areas.

The voxel method of data visualization is quite visual, but it has a number of disadvantages: the mutually perpendicular voxel edges are not convenient for perception, and the standard lighting calculation methods applied to the model do not add depth and volume to the resulting image. The application of the "Marching Squares" algorithm [17] allows obtaining a three-dimensional surface consisting of triangles, limiting the parts of the tomographic volume that differ in brightness or segmentation mask (see Fig. 12, Fig. 13). Visualization of anatomical structures at the subvoxel level is also relevant for the study of parietal air flows in the nasal cavity [18-20].

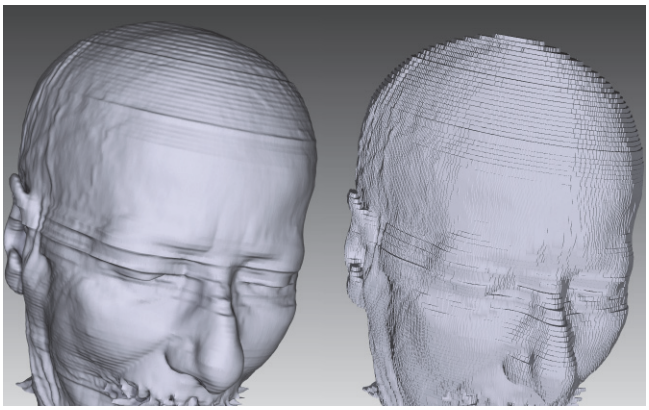


Fig. 13. Result of applying the "Marching Cubes" algorithm to a tomographic dataset 512x512x106 voxels with and without interpolation of the position of the triangle vertices.

7 Conclusions

Voxel models of tomographic data presentation are quite informative for the purposes of medical diagnostics and planning of operations. It is advisable for realistic visualization to triangulate the original voxel volume using the Marching Squares method and obtain three-dimensional surfaces that limit anatomical structures that differ in brightness or other features during segmentation.

Using modern high-speed computing systems with multi-core processors, it is possible, on the basis of the developed algorithms and software modules, to solve the problem of creating computer 3D models of the anatomical structures of the human head in real time. The created computer models are the basis for the subsequent solution of such problems: modification of 3D models in order to determine the optimal shape for the

patient of certain anatomical structures subject to plastic surgery; assessment of the functioning of the upper respiratory tract and computer planning of surgical correction of intranasal structures, if necessary. The prospect of the work is the development of methods of volumetric deformation of models of the anatomical structures of the nasal cavity for the possibility of their virtual correction, taking into account the biophysical properties of the investigated area and the realism of 3D visualization.

References

1. Avrunin, O. G., Alkhorayef, M., Saied, H. F. I., &Tymkovych, M. Y. The surgical navigation system with optical position determination technology and sources of errors. Journal of Medical Imaging and Health Informatics, 5(4), 689-696 (2015).
2. Moscatiello, F., Jover, J. H., Ballester, M. A. G., Hernández, E. C., Piombino, P., &Califano, L.. Preoperative digital three-dimensional planning for rhinoplasty. Aesthetic plastic surgery, 34(2), 232-238(2010).
3. Park, S. S.. Fundamental principles in aesthetic rhinoplasty. Clinical and Experimental Otorhinolaryngology, 4(2), 55(2011).
4. Adamson, P., Smith, O., & Cole, P.. The effect of cosmetic rhinoplasty on nasal patency. The Laryngoscope, 100(4), 357-359(1990).
5. Kherani, S., Javer, A. R., Woodham, J. D., & Stevens, H. E.. Choosing a computer-assisted surgical system for sinus surgery. Journal of otolaryngology, 32(3) (2003).
6. Innis, W., Byrne, P., &Tufano, R. P.. Image-guided osteoplastic frontal sinusotomy. American journal of rhinology, 19(5), 430-434(2005).
7. Strauss, G., Koulechov, K., Richter, R., Dietz, A., Trantakis, C., &Lüth, T.. Navigated control in functional endoscopic sinus surgery. The International Journal of Medical Robotics and Computer Assisted Surgery, 1(3), 31-41 (2005).
8. Cole, P., &HollierJr, L. Use of Intraoperative CT Scanning in Endoscopic Sinus Surgery: A Preliminary Report. Journal of Craniofacial Surgery, 20(1), 265(2009).
9. Tingelhoff, K., Moral, A. I., Kunkel, M. E., Rilk, M., Wagner, I., Eichhorn, K. W., ...&Bootz, F. Comparison between manual and semi-automatic segmentation of nasal cavity and paranasal sinuses from CT images. In 2007 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (pp. 5505-5508). IEEE. (2007).
10. Pirner, S., Tingelhoff, K., Wagner, I., Westphal, R., Rilk, M., Wahl, F. M., ...&Eichhorn, K. W. CT-based manual segmentation and evaluation of paranasal sinuses. European archives of oto-rhino-laryngology, 266(4), 507-518. (2009).
11. Tymkovych, M. Y., Avrunin, O. G., Paliy, V. G., Filzow, M., Gryshkov, O., Glasmacher, B., ...&Kozbekova, A.. Automated method for structural segmentation of nasal airways based on cone beam computed tomography. In Photonics Applications in Astronomy, Communications, Industry, and High Energy Physics Experiments 2017 (Vol. 10445, p. 104453F). International Society for Optics and Photonics. (2017, August)
12. Avrunin, O. G., Tymkovych, M. Y., Moskovko, S. P., Romanyuk, S. O., Kotyra, A., &Smailova, S. Using a priori data for segmentation anatomical structures of the brain. PrzeglądElektrotechniczny, 3, 102-105. (2017).
13. Al_Omari, A. K., Saied, H. F. I., &Avrunin, O. G.. Analysis of Changes of the Hydraulic Diameter and Determination of the Air Flow Modes in the Nasal Cavity. In Image Processing and Communications Challenges 3 (pp. 303-310). Springer, Berlin, Heidelberg. (2011).

14. Saied, H. F. I., Al_Omari, A. K., &Avrunin, O. G. An Attempt of the Determination of Aerodynamic Characteristics of Nasal Airways. In Image Processing and Communications Challenges 3 (pp. 311-322). Springer, Berlin, Heidelberg. (2011).
15. Avrunin, O. G., Nosova, Y. V., Paliy, V. G., Shushlyapina, N. O., Kalimoldayev, M., Komada, P., &Sagymbekova, A. Study of the air flow mode in the nasal cavity during a forced breath. In Photonics Applications in Astronomy, Communications, Industry, and High Energy Physics Experiments 2017 (Vol. 10445, p. 104453H). International Society for Optics and Photonics. (2017).
16. Knigavko, Y. V.. Calculation of venturi nozzles diameter for nasal breathing evaluation device. International Journal of Mechanical Engineering, 2, 21-28. (2013)
17. Newman, T. S., & Yi, H.. A survey of the marching cubes algorithm. Computers & Graphics, 30(5), 854-879. (2006)
18. Nosova, Y., Avrunin, O., Tymkovych, M., Khudaieva, S., Yurevych, N., &Glasmacher, B.. Determination of nasal breathing disorders according to computer tomography. In 2020 IEEE 40th International Conference on Electronics and Nanotechnology (ELNANO) (pp. 516-519). IEEE(2020).
19. Avrunin, O. G., Nosova, Y. V., Shushliapina, N. O., Kulish, S. M., Krekoten, E. G., Maciejewski, M., &Rakhmetullina, S.. Method for determination of laminar boundary layer of airflow in the upper respiratory tract. In Photonics Applications in Astronomy, Communications, Industry, and High-Energy Physics Experiments 2019 (Vol. 11176, p. 1117631). International Society for Optics and Photonics. (2019).
20. Nosova, Y. V., Faruk, K. I., &Avrunin, O. G. A tool for researching respiratory and olfaction disorders. Telecommunications and Radio Engineering, 77(15), 1389-1395.(2018).
21. Vyatkin Sergey, Romanyuk Alexander, Romanyuk Oksana, Nechyporuk Mykola, Troyanovskaya Tatiana and Tsikhanovska Olena. Photorealistic object reconstruction using perturbation functions and features of passive stereo projection. CONFERENCE ACIT' (2020).
22. Bezsmertnyi Y.O., Shevchuk V.I., Grushko O.V., Tymchyk S.V., Bezsmertna H.V.? Dzierżak R, et al. Information model for the evaluation of the efficiency of osteoplasty performing in case of amputations on below knee. Proc. SPIE 10808, Photonics Applications in Astronomy, Communications, Industry and High-Energy Physics Experiments 2018, 108083H (2018).
23. Shevchuk V.I., Bezsmertnyi Y.O., Kyrychenko V.I. M'iazova plastyka pid chas amputatsii i reamputatsii homilky [Muscular plastic during amputation and abdominal shaft. Orthopedics, traumatology and prosthetics]. Ortopedija, travmatologija i protezirovanie. 4 (2010), pp. 13-18 9 (2010). (in Ukrainian).
24. Shevchuk V.I., Bezsmertnyi Y.O., Maiko V.M.. Kistkova plastyka pry amputatsiiakh ta reamputatsiiakh nyzhnikh kintsivok [Bone plastics with amputations and reamputs of the lower extremities]. Ortopedija, travmatologija i protezirovanie. 1 (2011), pp 47-55 (2011). (in Ukrainian).
25. Yurii O. Bezsmertnyi, Viktor I. Shevchuk, Sergii V. Pavlov, "Prognosis of efficacy of medical and social rehabilitation in disabled individuals with respiratory diseases", Proc. SPIE 11176, Photonics Applications in Astronomy, Communications, Industry, and High-Energy Physics Experiments 2019, 1117633 (6 November 2019); https://doi.org/10.1117/12.2537340.
26. Yurii O. Bezsmertnyi, Sergii V. Pavlov, and etc. "Information model for forecasting of violation reparative osteogenesis of long bonds", Proc. SPIE 11176, Photonics Applications in Astronomy, Communications, Industry, and High-Energy Physics Experiments 2019, 111762A (6 November 2019); https://doi.org/10.1117/12.2536250.

27. Sergey I. Vyatkin, Olexander N. Romanyuk, Sergii V. Pavlov, and etc. "Transformation of polygonal description of objects into functional specification based on three-dimensional patches of free forms", Proc. SPIE 11176, Photonics Applications in Astronomy, Communications, Industry, and High-Energy Physics Experiments 2019, 1117622 (6 November 2019); https://doi.org/10.1117/12.2537043.
28. Sergey I. Vyatkin, Olexander N. Romanyuk, Sergii V. Pavlov, and etc. Offsetting and blending with perturbation functions // Proc. SPIE 11045, Optical Fibers and Their Applications 2018, 110450W, 2019; doi: 10.1117/12.2522353; https://doi.org/10.1117/12.2522353.
29. Sergey I. Vyatkin, Olexander N. Romanyuk, Sergii V. Pavlov, and etc. A GPU-based multivolume rendering for medicine // Proc. SPIE 11045, Optical Fibers and Their Applications 2018, 1104513, 2019); doi: 10.1117/12.2522408.
30. Sergey I. Vyatkin, Olexander N. Romanyuk, Sergii V. Pavlov, and etc. Offsetting and blending with perturbation functions // Proc. SPIE 10808, Photonics Applications in Astronomy, Communications, Industry, and High-Energy Physics Experiments 2018, 108082Y, doi: 10.1117/12.2501694.
31. Sergey I. Vyatkin, Sergii A. Romanyuk, Sergii V. Pavlov, and etc. Using lights in a volume-oriented rendering // Proc. SPIE 10445, Photonics Applications in Astronomy, Communications, Industry, and High Energy Physics Experiments 2017, 104450U; doi: 10.1117/12.2280982..
32. Oleg G. Avrunin; Maksym Y. Tymkovych; Sergii V. Pavlov; Sergii V. Timchik; Piotr Kisała, et al. Classification of CT-brain slices based on local histograms, *Proc. SPIE* 9816, Optical Fibers and Their Applications 2015, 98161J (December 18, 2015); doi:10.1117/12.2229040.
33. Olexander N. Romanyuk; Sergii V. Pavlov; Olexander V. Melnyk; Sergii O. Romanyuk; Andrzej Smolarz, et al. Method of anti-aliasing with the use of the new pixel model, *Proc. SPIE* 9816, Optical Fibers and Their Applications 2015, 981617 (December 18, 2015); doi:10.1117/12.2229013.
34. S. O. Romanyuk; S. V. Pavlov; O. V. Melnyk. New method to control color intensity for antialiasing. Control and Communications (SIBCON), 2015 International Siberian Conference. - 21-23 May 2015. - DOI**:** 10.1109/SIBCON.2015.7147194.

Multi-criteria decision-making system for design and implementation on the market rehabilitation toys

Kinga Korniejenko[0000-0002-8265-3982] and Katarzyna Harpak

Faculty of Materials Engineering and Physics, Cracow University of Technology, Cracow, Poland
kinga.korniejenko@pk.edu.pl

Abstract. Currently, devices, toys used in the practice of children's rehabilitation centres usually have simple forms and shapes, a monolithic form and are devoid of mechanical or electronic control systems. In practice, the use of interactive tools is rare and associated mainly with expensive and advanced solutions, such as rehabilitation robots and devices using virtual reality. Rehabilitation centres mainly use simpler and cheaper devices. The design of this type of aid is usually simple according to technology, but quite complex to ensure full functionality for the rehabilitation purpose. A lot of aspects should be taken into consideration. The aim of this chapter is to analyze the possibility of using multi-criteria methods for support in the designing and implementation on the market rehabilitation toys. Multi-criteria methods used to support the decision-making process, are a response to the complexity of contemporary issues in conditions of uncertainty, incomplete data and changing environment. The research methods used in the article are: critical analysis of literature sources and the case study of the use of the Analytic Hierarchy Process (AHP) for implementation on the market rehabilitation toys.

Keywords: Analytic Hierarchy Process, Rehabilitation Toy, Rehabilitation Equipment.

1 Introduction

The life of each individual affected by disability is more or less disturbed, but disability in children has a specific character [1,2]. It disturbs their proper development and affects the activity characteristic of the development period (playing, learning) [2,3]. Moreover, the child's disability significantly disturbs the functioning of the whole family [4]. Children with motor organ dysfunction have to deal with specific problems related to this kind of disability. Apart from typical life tasks, they must undertake special activities resulting from the specificity of their dysfunction, such as: pain relief, long-term immobilization of certain parts of the body, absolute compliance with medical recommendations [4,5]. They have to endure separation from their parents, caused by hospitalization or the need to stay in rehabilitation centers.

Disability, regardless of its type, has an impact on the functioning of the individual, while motor disability, also referred to as motor organ dysfunction, causes significant psychosocial consequences [4,5]. The influence of movement limitations on a child's development is particularly negative, because movement is one of the child's basic needs. Blocking these need to move is very disadvantageous because it exposes a child with motor organ dysfunction to unpleasant frustration experiences and violates his self-esteem [5,6]. Various forms of rehabilitation aimed at restoring or achieving independence of the patient are an inseparable part of the treatment of motor organ dysfunction in children. Congenital disability is a factor determining the nature of the child's relationship with the environment from the moment of birth, thus influencing its functioning [6,7].

The motor development of a small child is closely related to his mental development. By manipulating objects, the child learns about their structure and function. By affecting objects, the child gets new impressions, thus enriching his knowledge about the world [6,7]. The children with a congenital lack or deformity of limbs have a significantly reduced ability to manipulate objects, but it is not impossible. Such a child can be helped by the parent by, for example, moving the object over the child's face, torso, legs. With appropriate interaction, the development of cognitive processes in children with limb injuries can proceed properly.

However, congenital damage to the musculoskeletal system rarely implies unfavourable consequences in terms of mental development, more often they are associated with social development. Damage to motor organs eliminates the child from many motor games and thus the inflow of social stimuli is limited, and the child is deprived of the possibility of learning social rules and interpersonal behaviour within the group [7,8]. It is also important that isolation from the social group (rejection by the group, fear of joining it) makes it difficult to develop the need for emotional contact.

Very important in supporting the treatment of children with limb injuries, apart from specialist medical care, is an appropriate rehabilitation program with the use of equipment adjusted to the needs and age of the patient [3,9]. Children's motor rehabilitation is based mainly on exercises conducted under the supervision of a specialist. Various therapeutic aids are often used, which are to be functional, effective, but also to mobilize and encourage the child to actively participate in exercises. This effect is usually obtained thanks to the selection of appropriate colours, textures, and giving therapeutic aids or their elements the form of toys. These aids are used, among others (depending on the type of device): to maintain the correct body posture, maintain balance, improve spatial orientation, improve small and high motor skills, superficial and deep sensation, improve eye-hand coordination. Rehabilitation toys play an important role in sensory integration therapy. Different types of information received by different types of senses and reaching the central nervous system after integration generate an answer that builds the necessary foundations for the psychomotor development of a child [3,10].

The essence of sensor integration therapy is to provide the body with a controlled amount of sensory stimuli, especially proprioceptive, tactile and vestibular stimuli. It should be related to various motor stimulation, for better processing of the received stimuli and their proper adaptation to them. As a result, it leads to an improvement in equivalent reactions, a better integration of reflexes, an improvement in fine and gross

motor skills and the normalization of postural tension. During therapy, the child does not acquire specific skills, but by improving sensory integration, it stimulates the nervous processes that form the basis of these skills [1,10].

The use of interactive toys is aimed at, among others, increasing the attractiveness and effectiveness of rehabilitation activities with children. In the literature on the subject, a lot of space is currently devoted to devices using virtual reality and rehabilitation robots [2,11]. A rehabilitation robot is considered to be virtually any reprogrammable, flexible platform that enables the physical interaction of the robot with the patient and manipulation of the patient's body parts by the robot for therapeutic purposes (called robot mediated therapy). Robot-assisted rehabilitation may be particularly effective in ensuring greater patient motivation through the use of biofeedback, computer games, the aforementioned virtual reality systems and other forms of feedback [2,11].

Currently, devices, toys used in the practice of children's rehabilitation centres usually have simple forms and shapes, a monolithic form and are devoid of mechanical or electronic control systems. In practice, the use of interactive tools is rare and associated mainly with expensive and advanced solutions, such as rehabilitation robots and devices using virtual reality. Rehabilitation centres mainly use simpler and cheaper devices such as: balls, swings, balancers, platforms, harnesses, skateboards, tunnels, textural sensory aids. The design of this type of aid is usually simple according to technology, but quite complex to ensure full functionality for the rehabilitation purpose. A lot of aspects should be taken into consideration, including the principle of ensuring the smoothest possible surfaces (or soft to the touch), oval edges, as few details as possible and protruding elements, so as to ensure the safety of the equipment by children.

The aim of this chapter is to analyze the possibility of using multi-criteria methods for support in the designing and implementation on the market rehabilitation toys. The chapter is divided into fourth part. The first section is a short introduction into the topic. The second section characterizes used research methods, including the description of the Analytic Hierarchy Process (AHP). The third section shows the achieved results and discuss their importance for the designing of the rehabilitation toys. The last section summarizes the most important findings.

2 Research method

2.1 The Analytic Hierarchy Process (AHP)

The Analytic Hierarchy Process (AHP) was used as a core empirical research method for determination of the most important features for rehabilitation toys. This method provides a comprehensive and rational framework for structuring a decision problem and helps decision makers in taking final decisions.

This method was created and developed by Thomas L. Saaty in the 1970s (University of Pittsburgh) [12,13]. This method is used to structure, evaluate and synthesize the considered problems. The method uses the principles of the multi-attribute utility theory and allows for the decomposition of a complex decision problem, leading to the ordering of a finite set of decision alternatives. An important attribute of this

method is the possibility of taking into account both countable and non-quantifiable factors in the analysis of the problem [12,14].

The AHP method is a response to the growing complexity of the contemporary world and economies, where decisions are made under conditions of increasing uncertainty. This is mainly due to the excess of information coming from the changing external environment. One could argue that the more information is available, the better from the point of view of the decision-making process, but in practice it turns out that the excess of information is as unfavourable as its shortage [12]. The decision maker is not able to process all the data available to him, therefore he is forced to choose the most important factors that have a decisive influence on the decision problem presented to him. Such a selective and superficial selection of information may lead to the omission of relevant data from the point of view of the problem under consideration. It should also be noted that uncertainty in decision making and the associated risks also increase with the complexity of the problem and the degree to which it is new. The method has been used to solve many decision problems, incl. in the following areas: economics and management; in the fields of marketing, finance, transportation, resource allocation, planning and forecasting [15,16], politics; in negotiations, conflict resolution, arms control, war games [12], social problems; in the field of education, medicine, law, public sector, sport [17,18], as well as technology, including market selection and technology transfer [13,19].

The AHP method, due to its hierarchical structure and the possibility of taking into account many qualitative and quantitative variables (criteria) in the decision-making analysis, can undoubtedly find its application in the study of the optimal features of an innovative rehabilitation toy for children introduced to the market. Traditional methods of determining the desired features and functions of the designed toy (SWOT, linear programming, econometric methods), although often used by decision makers, do not always correspond to real market realities. The AHP method, due to the holistic approach to the analysis and evaluation of alternatives and decision variables, and the possibility of taking into account the opinions of various experts, may be a rational tool supporting the decision-making process in the field of designing and building an innovative rehabilitation toy for children.

The AHP method comprises three basic steps [19, 20]:

1. Decomposition of the problem and construction of a hierarchical decision model.
2. Comparative evaluation of the strategy.
3. Prioritizing the pairwise comparison matrix.

The first step is linked with defining the problem. After that, it is necessary to decompose it and define the internal and external factors that make up the issue under study. The problem is divided into components, which include: main goal, main criteria (sub-goals), detailed criteria (attributes) and possible decision alternatives (optionally). This division makes it possible to distinguish the structure and to notice the multi-level nature of the problem, and consequently to build a hierarchical structure. The hierarchical scheme in the AHP method consists of the number of levels, which depends on the complexity of the problem [12, 13].

The next stage is comparative evaluation of the strategies. It leads to a synthetic assessment and partial assessments for the tested alternatives: individual elements of the hierarchical tree (criteria and sub-criteria) listed in the first stage are compared in pairs in terms of benefits, costs, opportunities and risks. Based on the presented relations, the expert expresses his preferences regarding all pairs of decision-making elements of the model. Ratings are determined most often on the basis of expert judgments, less often on the basis of existing measurements or statistics needed to make a decision. The AHP method can be used in making decisions by an individual decision-maker as well as in group experts. Group decision making can be done by reaching a consensus in a group of decision makers or by completing questionnaires separately by each member of the decision group and then analyzing the consistency of the responses.

Comparisons are most often made on a verbal scale (Table 1) and constitute the basis of the AHP method. Thanks to it, the priorities of individual elements are established. Numbers are assigned to the individual degrees of preferences. T. Saaty proposed a scale of natural numbers from 1 to 9 and the reciprocal of these numbers in the opposite relation. The higher the number, the greater the intensity of preferences, and assigning a number equal to 1 means that the significance of both model elements is equivalent. These numbers are assessments of decision makers [21].

Table 1. The fundamental scale of comparisons by T. Saaty [21].

Intensity	Definition
1	Equally preferred
3 (1/3)	One element is slightly more relevant than another
5 (1/5)	One element is strongly more relevant than another
7 (1/7)	One element is very strongly more relevant than another
9 (1/9)	One element is extremely more relevant than another
2 (1/2), 4 (1/4), 6 (1/6), 8 (1/8)	For compromise comparisons between the above-mentioned values

It is assumed that the evaluations expressed by the decision makers are inaccurate, and the exact evaluations are the ratios of the respective components of the searched ordering vector. Limiting the scale to 9 natural numbers results from the conclusions of the psychologist Georg Miller, who analyzed the results of research on the scope of direct memory and concluded that regardless of the type of material, people reproduce from 5 to 9 elements, i.e. 7 ± 2. This shows that decision makers cannot efficiently compare more items at the same time because they lose consistency. The number of incorrect opinions and conclusions among decision makers is increasing. The scale proposed in the method allows for the inclusion of the experience and knowledge of the decision-maker and allows for indicating how many times a given element outweighs others in relation to a given criterion.

The final stage is prioritizing the pairwise comparison matrix by establishing the eigenvector of the comparison matrix. For each comparison matrix, its eigenvector has a local priority. These priorities can also be compared with each other by using global

priorities. The eigenvector comparison of the matrix can also be used to arrange the elements according to their relative importance for the overall problem being solved (prioritizing). The final ranking is made by calculating for each variant the value of the aggregating utility function, which is the sum of the products of absolute weights of the variant on the way from the variant through the criteria to the goal. The absolute weights of each matrix are computed by finding the eigenvector of the given matrix. The last element of the method is the selection of the best solutions - decision alternatives.

A slight inconsistency in the responses is permissible, which proves that the expert providing the answer can change his mind (which is most often the result of gaining new information and changing environmental conditions). However, the discrepancy should not be too great. Most often it is stated that the value of the permissible error should not exceed 10%. In some situations, a greater inconsistency of a given matrix is acceptable, even at the level of 15%. The occurrence of an error higher than 15% means that the answers are inconsistent and in this case the decision-maker should redefine his preferences, which in practice means repeating all or some comparisons [19].

2.2 Presentation of the problem in the form of a multi-criteria mathematical model of the AHP

In order to determine the most optimal model of the rehabilitation toy for children, a hierarchical AHP model was built in terms of "benefits" and "costs" resulting from the introduction of an innovative rehabilitation toy for children on the market. Based on literature research and expert opinions, the main criteria and detailed sub-criteria, reflecting the most important elements determining the introduction of the toy to the market, were proposed. The surveyed experts, using the 9-point Saaty scale, determined the dominance of individual decision-making elements of the AHP "benefits" and "costs" models. Two multi-criteria AHP mathematical models - "benefits" and "costs" were built. Each model consists of three hierarchical levels. At level I of the hierarchy, the main goal was placed, defined as "Introducing an innovative rehabilitation toy for children to the market". At level II, there are the main criteria: safety, functionality, price, design, while at level III, there are sub-criteria in relation to the basic criteria.

The development of a hierarchical model in terms of "benefits" and "costs" of introducing an innovative rehabilitation toy for children to the market was carried out using the AHP method on the basis of the following scheme:

- Problem definition - verification of internal and external factors influencing the benefits / costs of introducing an innovative rehabilitation toy for children to the market.
- Identification of the main goal - "introducing an innovative rehabilitation toy for children to the market".
- Presentation of the problem structure in the form of a hierarchical tree of benefits (Fig. 1.) and costs (Fig. 2.). In the developed structure, criteria (safety, functionality, price, design) influencing the main goal and detailed sub-criteria determining the main criteria were defined.
- Determining the verbal opinions of the respondents, obtained on the basis of their knowledge, experience, knowledge of the rehabilitation toys market, professional competences by conducting a direct and individual interview with the questionnaire.

- Defining the preferences (dominance) of the main criteria by the respondents, by making a pairwise comparison of their importance in relation to the main goal.
- Determination by the respondents of the preferences (domination) of the sub-criteria, by making pairwise comparisons of their importance in meeting the main criteria.
- Quantification of the verbal opinions of the respondents based on the 9-step fundamental scale of Saaty's preferences.
- Calculating the priorities of the elements of the "hierarchical tree" of benefits / costs in relation to their impact on the main goal on the basis of normalization of the vectors of own comparisons.
- Calculation of CR non-compliance rates for each respondent.
- Synthesis of the results obtained, necessary to indicate the most beneficial / costly elements of introducing an innovative rehabilitation toy for children to the market.

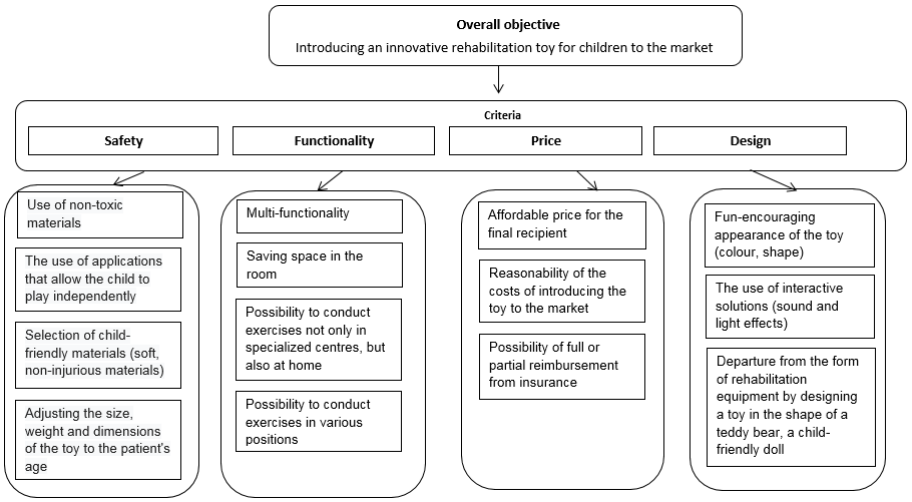


Fig. 1. AHP's decision-making hierarchy in terms of "benefits"

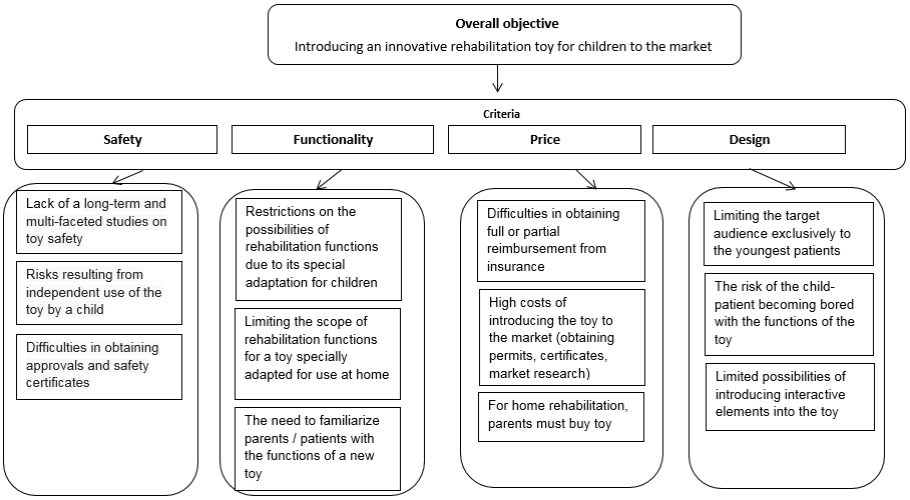


Fig. 2. AHP's decision-making hierarchy in terms of “costs"

2.3 Characteristic group of experts

The research was aimed at verifying the decision-making elements determining the launch of an innovative rehabilitation toy for a child on the market. Based on literature research and an interview with a questionnaire with experts dealing with the treatment and rehabilitation of the youngest patients with both physical and mental disabilities, the most important criteria that should be met by the designed rehabilitation toy were determined. The decision-making elements of the model proposed by the author include both quantitative variables - relatively easy to present in numerical form, as well as qualitative variables expressing the subjective views of the surveyed experts.

In order to verify the elements of the AHP hierarchical model, a study was carried out on experts in the medical sector who had knowledge and experience in the field of rehabilitation of children with upper limb dysfunctions. The group of experts included both physiotherapists (9 people) with experience in the rehabilitation of children, as well as specialist doctors (4 people) dealing with the treatment of children with motor disabilities, in total 13 experts - industry specialists.

When analyzing the age structure of the surveyed experts, people aged 26-35 (8 people) dominated, the remaining age groups were represented as follows: two people aged 36-45, two 46-55 and one person under 26 age.

The data was collected using a structured questionnaire. Apart from filling in the questionnaire, they included the experts proposing criteria or changing the name for individual criteria and sub-criteria of the questionnaire. All data and information collected during the study were analysed qualitatively and quantitatively. The expert computer program Super Decisions was used to analyse the results.

The study based on the AHP method on a group of experts was carried out according to the following procedure:

- on the basis of the structure of the decision problem in the form of a hierarchical tree (goal, criteria, sub-criteria, construction of strategic models, definition of the elements of the decision structure), specified in the literature, a structured interview was conducted with an expert, allowing to determine what modifications, according to him, should occur in the structure of the problem (reference to criteria and sub-criteria in the method),
- conducting a specific study on a group of experts using a questionnaire,
- calculating preferences for each of the surveyed entities separately,
- pairwise comparison of models with regard to sub-criteria,
- interpretation of the results.

3 Results

Based on the comparisons of experts, information was obtained on the structure of the importance of the main criteria of the AHP “benefits” model. When analysing the priority values of the global criteria of the main "benefits" model of introducing an innovative rehabilitation toy for children to the market, it can be clearly observed that the most important criterion in the opinion of experts is "safety". Based on the comparisons of experts, this decision-making element was given priority (0.4950), which is almost half in the structure of all main criteria of the AHP "benefits" model. The high value of the "safety" criterion in the general structure of "benefits" draws the attention of experts that when introducing an innovative rehabilitation toy for children to the market, the main aspect taken into account during the design process should be ensuring the safe use of the product, especially for children using it and people assisting in the rehabilitation of the youngest. Safety of use should be ensured by the selection of appropriate, certified materials, as well as by adjusting the form and size of the toy so that it does not pose a threat to users. The second most important criterion is the “functionality” criterion, which received the priority weight (0.3187). The respondents indicated that the "functionality" of the toy is not the most important when introducing it to the market, although it is an important element when choosing a device for rehabilitation. The two remaining decision-making elements - the criteria "price" and "design" received similar, relatively low priority weights, respectively (0.1147) and (0.0789), which shows that in the opinion of experts, these criteria are the least important during the process designing and introducing an innovative device to the market. This means that when choosing a rehabilitation toy, assessing it from the point of view of benefits, attention is paid to the aspects of safety and functionality, and to a lesser extent to its price and appearance.

In the area of "safety" of the hierarchical AHP model of "benefits", the element which, in the opinion of experts, obtained the highest local values is the sub-criterion "selection of child-friendly materials (soft materials, non-traumatic)" (0.3006). This element is, according to the respondents, the most important benefit of the safety area resulting from the introduction of an innovative rehabilitation toy to the market, and those who design rehabilitation equipment for children should pay special attention to the use of appropriate technologies and materials to ensure this element is provided. As

regards importance, the sub-criteria, assessed by the respondents with almost equal values of local priorities, are: "adjusting the size, weight and dimensions of the toy to the patient's age" and "using applications that allow the child to play independently". These elements obtained respectively (0.2732) and (0.2722). The experts agreed that the least important in the hierarchy of benefits of the security aspect is "the use of non-toxic materials", the importance of which, in the respondents' opinion, amounts to (0.1540), which is slightly over 15% in the overall share of the security benefits.

The next most important criterion in the respondents' assessment is the functionality of the AHP "benefits" model. When analysing the importance structure of the functional area of the hierarchical "benefits" model, it is noted that three out of the four specifying sub-criteria are dominant. The most important decision-making element of the analysed model is "the possibility of conducting exercises not only in specialized centres, but also at home" (0.3540). Experts pointed out that, in particular in the case of the rehabilitation of the youngest patients, very important in choosing rehabilitation equipment is the possibility of conducting activities at home, in a safe and familiar environment for children, not associated with hospital conditions. Another, almost equally important sub-criterion of the APH model of benefits in the area of functionality is "the possibility of exercising in various positions" (0.2867), which allows the use of equipment taking into account the rehabilitation of various parts of the body. The third in the structure of the benefits of the functional area is the "multifunctionality" element, which has obtained a similarity to the sub-criterion described above (0.2734). The element that, in the opinion of the respondents, was assessed as the least important in the model hierarchy is "saving space in the room" (0.0859).

When analysing the values of the sub-criteria in terms of the price area, the factor “possibility of total or partial reimbursement from insurance” is clearly dominant in the opinion of the surveyed experts. The local priority for the analysed sub-criterion constitutes almost 60% of the overall structure of benefits for this area, which proves the very high importance of this element in the implementation of the main criterion in question (0.5961). The possibility of qualifying equipment to the list of reimbursed medical devices, and thus obtaining a partial or total reimbursement from health insurance, is a very important determinant that may affect decisions about the purchase and wider use of a toy, both by individual recipients and specialized centres dealing with rehabilitation of children. Another sub-criterion which obtained relatively lower values of the local priority is "affordable price for the final recipient" (0.2402). Experts, relying on the statement that the innovative rehabilitation toy is to be used, among others exercises at home, assessed that the undoubted advantage of its introduction will be a competitive, affordable price - possible to cover by individual recipients. The lowest value in the discussed structure was obtained by the element "cost-effectiveness of introducing the toy to the market" (0.1637).

In the area of design of the hierarchical model of "benefits", the dominance of values of two decision-making elements is noticeable. The most advantageous in terms of design of the innovative rehabilitation toy for children introduced to the market is "the use of interactive solutions (sound effects)" (0.4107). During the preliminary research and direct interviews with manufacturers of medical equipment, information was ob-

tained that there is a lack of rehabilitation devices on the Polish market dedicated specifically to children, due to the unique, interesting design or the use of interactive solutions encouraging the child to work and play. The examined experts pointed out the need to introduce innovations in the field of designed devices so that they would be more friendly and encouraging for the rehabilitated child. Thus, attention is drawn to the possibility of filling an existing market gap in the field of rehabilitation equipment specifically aimed at young patients. Next, in the structure of importance there is the element “lack of the typical form of rehabilitation equipment by designing a toy in the shape of a teddy bear, a child-friendly doll” (0.3655). According to experts' opinions, the lowest rated sub-criterion of the AHP model "benefits" model area is "the appearance of the toy encouraging to play (colour, shape)”. Based on experts' assessments, this element accounts for less than 23% of the overall benefit structure of the model design area.

In addition to the previously discussed local values for individual sub-criteria of the AHP benefits model, the global values of these decision elements are also presented. They define the importance of all sub-criteria of the AHP "benefits" model in achieving the adopted goal of "introducing an innovative rehabilitation toy for children to the market". The presented list includes all the previously discussed detailed sub-criteria of the hierarchical AHP model (Fig. 3).

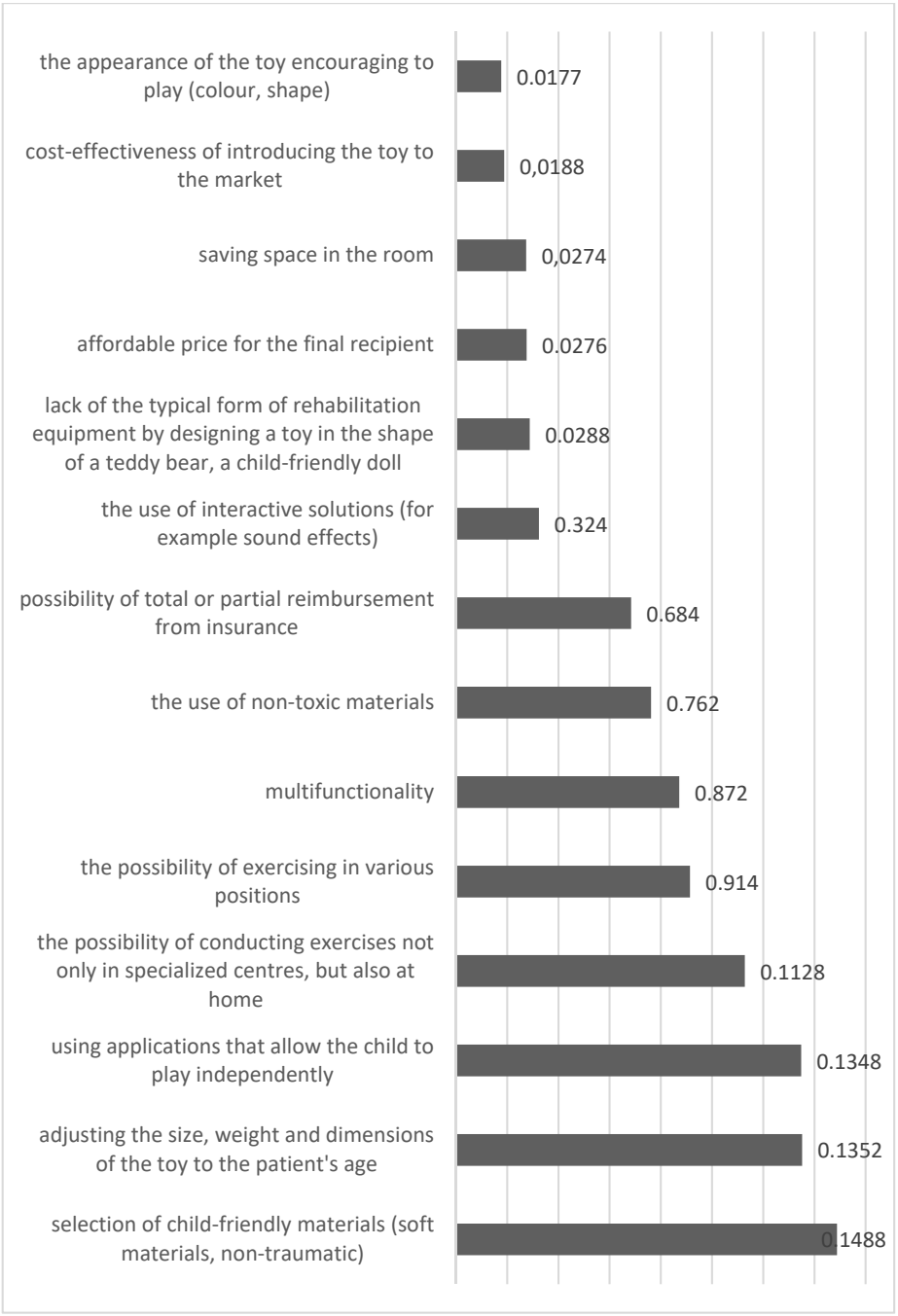


Fig. 3. Global priority values for the AHP "benefits" sub-criteria

In a study with a questionnaire, by comparing the elements of the model in pairs, the experts also assessed the importance of the main criteria and sub-criteria of the AHP model in relation to the "costs" resulting from the "introduction of an innovative rehabilitation toy for children to the market". The values of local and global priorities, presented in the further part of the study, were obtained through a diagram analogous to the model of "benefits", comparing the importance of decision elements in terms of "costs".

As in the case of the AHP "benefit" model analysis presented in the previous section, the analysis of the hierarchical "cost" model of "introducing an innovative rehabilitation toy for children to the market" began with determining the importance structure of the main criteria. When analysing the structure of importance of the main criteria of the AHP "cost" model, there are clear convergences with the priority hierarchy of individual criteria of the "benefits" model. Similarly to the hierarchical structure of "benefits", the highest priority value was obtained by the criterion of "security". This criterion accounts for almost 50% of all safety area criteria. The high importance of safety in the analysis of the "costs" of introducing an innovative rehabilitation toy on the market, in the opinion of experts, indicates that this area is associated not only with significant benefits, but also possible risks, and thus costs, both for recipients and device manufacturers. The second most important criterion in the respondents' assessment is "functionality", which compared to the most important criterion "safety" (0.4416), received a much lower priority value (0.2877). As in the "benefits" model, the lowest values were obtained by the criteria "price" (0.1503) and design (0.1204).

According to experts, the most important cost in terms of "safety" are "difficulties in obtaining approvals and safety certificates". This element obtained the dominant local priority value (0.4076). Another important item in the structure of costs related to the area of "safety" is "threats resulting from the independent use of a toy by a child" (0.3313). The least expensive element of the hierarchical structure of the model, according to the respondents' assessments, is “the lack of long-term and multifaceted research on toy safety” (0.2610). The relatively low value of this sub-criterion in the overall cost structure may be justified, inter alia, by the fact that the toys will be made of materials and technologies that have previously been thoroughly researched and tested.

Successively, respondents participating in the study assessed the value of the priorities of local sub-criteria in the criterion of "functionality" of the AHP "cost" model. The most important in the assessment of the "costs" of the functional area of the AHP model is the sub-criterion "limitation of the rehabilitation functions due to its special adaptation for children" (0.5164). This element reflects the economic costs of introducing the toy to the market related to the special adaptation of its functions for use by children, which excludes its use by a wider target group - e.g. Adults. The other two decision-making elements of the model "limiting the scope of rehabilitation functions for a toy specially adapted for use at home" and "the need to familiarize rehabilitators / parents / patients with the functions of a new toy" obtained significantly lower priority values - respectively 0.3207 and 0.1629.

In the assessment of the cost of the price area of the hierarchical model, according to the respondents, the highest cost is concerns about difficulties in “obtaining total or

partial reimbursement from insurance" (0.5483). The procedure of including a rehabilitation toy on the list of medical products partially or fully reimbursed by health insurance is complicated and requires meeting many requirements and obtaining the necessary certificates. Possible difficulties are related to the formal requirements as well as the time needed to meet the guidelines set out in the Act on Medical Devices. Another item in the structure of importance of the AHP model price sub-criterion "costs" is, in the opinion of medical sector experts, "for the purpose of home rehabilitation the need to purchase a toy by parents" (0.3219). Experts emphasized that in order for a toy to fill a significant market gap in the field of rehabilitation equipment for children, which can be used at home, it must be characterized by an affordable price, which can be covered by parents even in the absence of co-financing from external sources. The least important price sub-criterion in the opinion of the respondents is "high costs of introducing the toy to the market (obtaining permits, certificates, market research)" (0.1298).

The main criterion "design" of the AHP "cost" model obtained the lowest priority value (0.1204) in the respondents' assessment. When analysing the values of the sub-criteria for the AHP "cost" model design area, one can notice similarities in the priority values assigned by experts to all three decision-making elements. The difference in values between the analysed sub-criteria is several percentage points, which means that these decision elements have similar importance in the structure of the main criterion. The sub-criterion "risk of child-patient getting bored with the functions of the toy" obtained the value (0.3969), thus constituting the most important element of the presented hierarchy. In the order of importance, the cost consisted of "limited possibilities of introducing interactive elements into the toy" (0.3060). The least significant sub-criterion of the hierarchy is "limiting the target audience to the youngest patients only" (0.2971).

Moreover, the analysis of the significance of decision-making elements on a local basis for individual main criteria of the AHP "cost" model, calculations were also made of the values of global sub-criteria. The structure of the importance of individual sub-criteria in achieving the main goal, defined as "introducing an innovative rehabilitation toy for children to the market" was presented. Figure 4 shows the structure of the importance of global priorities for all sub-criteria of the AHP "cost" model.

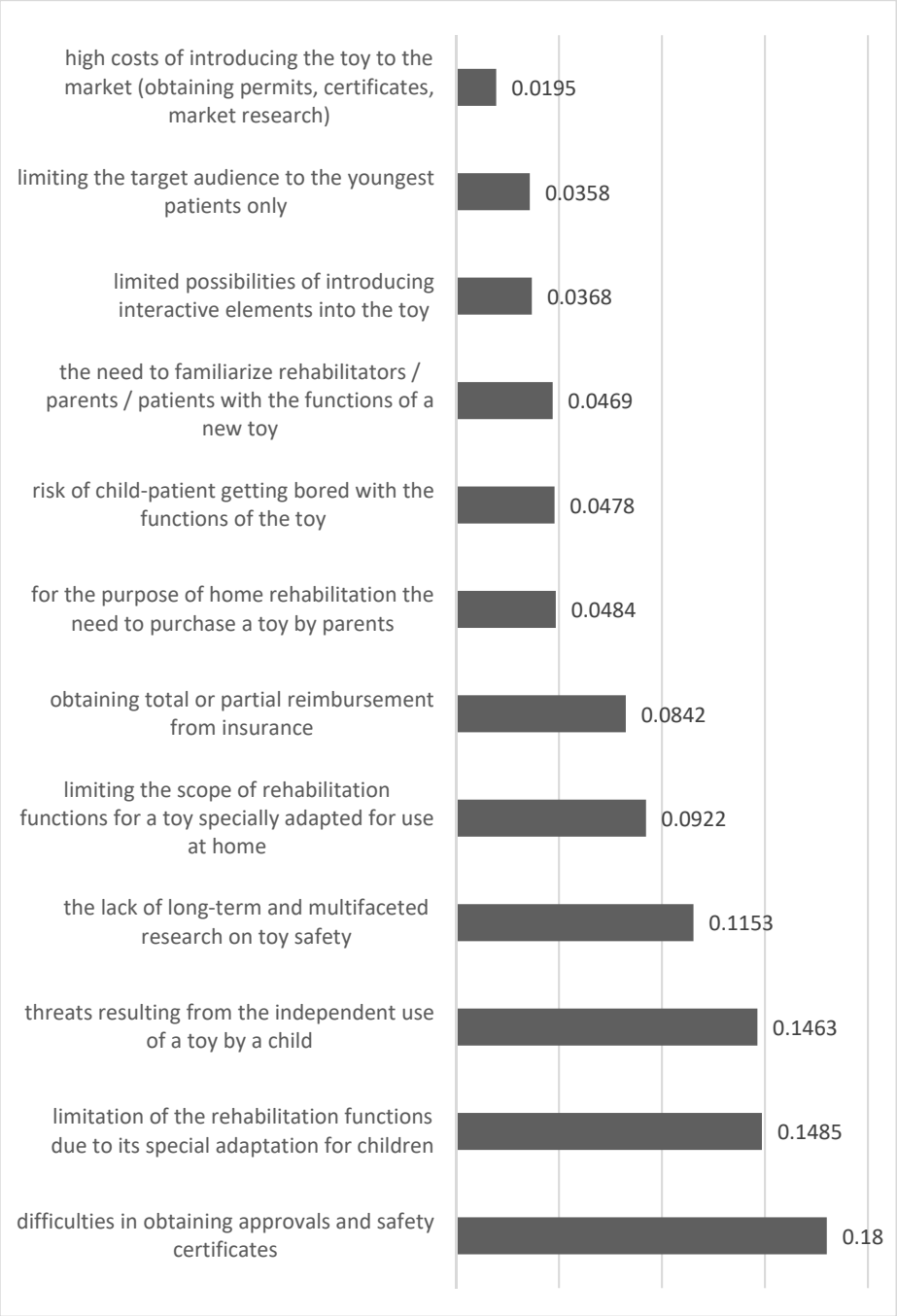


Fig. 4. Global priority values for the AHP "costs" sub-criteria

4 Conclusions

There is a significant market gap in medical equipment specifically dedicated to small patients. On the rehabilitation equipment market you can find numerous medical devices and aids for the rehabilitation of the upper limbs (rehabilitation tapes and weights, hand trainers, manual charts, rehabilitation balls, etc.), although these devices are not specially adapted for children. In the case of rehabilitation of children and adolescents, the basic problem is their involvement in exercising. The reason for this is the attractiveness of the activities they perform. Children tend to give up activities quickly which they find boring and unattractive. For this reason, in order to encourage children to perform rehabilitation activities, the equipment, in addition to appropriate technical parameters (adaptation to the child's height, weight, strength), should, with its design, shape, colors and interactive functions, encourage the child to further exercise. In this respect, few devices meet these characteristics.

Rehabilitation equipment for children available on the market in many cases, due to its high price or complicated form of use, cannot be used at home, but only in hospitals or rehabilitation centers. Undoubtedly, this affects the quality and duration of treatment, as well as the inconvenience for the patient's families related to frequent visits to specialized facilities offering appropriate rehabilitation equipment. There is a lack of multifunctional devices that would be specially adapted to the needs of the youngest patients with upper limb dysfunctions, and would allow for effective rehabilitation at home, without the need for constant care by medical personnel, at a price that could be covered by the patient's family.

The presentation of the problem in the form of a multi-criteria mathematical model of the AHP, both by analyzing the elements affecting the "benefits" and "costs" resulting from the introduction of an innovative rehabilitation toy for children onto the market, is an attempt at a complementary approach to the issue. When analyzing the market and interviewing the medical staff involved in the rehabilitation of children with all kinds of handicaps of the upper limbs, two basic conclusions were drawn. The results of AHP analysis could be applied into practice to design the rehabilitation toy.

References

1. dos Santos Nunes E.P., da Conceição Júnior V.A., Giraldelli Santos L.V., Pereira M.F.L., de Faria Borges L.C.L. (2017) Inclusive Toys for Rehabilitation of Children with Disability: A Systematic Review. In: Antona M., Stephanidis C. (eds) Universal Access in Human–Computer Interaction. Design and Development Approaches and Methods. UAHCI 2017. Lecture Notes in Computer Science, vol 10277. Springer, Cham.
2. Valadão C.T., Alves S.F.R., Goulart C.M., Bastos-Filho T.F. (2017) Robot Toys for Children with Disabilities. In: Tang J., Hung P. (eds) Computing in Smart Toys. International Series on Computer Entertainment and Media Technology. Springer, Cham.
3. Lindsay, S., Edwards, A.: A systematic review of disability awareness interventions for children and youth. Disability and Rehabilitation 35(8), 623–646 (2013).

4. Venkatesan, S., Yashodhara, G.Y.: Parent Opinions and Attitudes on Toys for Children with or Without Developmental Disabilities. International Journal of Indian Psychology 4 (4), 6-20 (2017).
5. Martín, A., Pulido, J.C., González, J.C., García-Olaya, Á., Suárez C.: A Framework for User Adaptation and Profiling for Social Robotics in Rehabilitation. Sensors 20, 4792 (2020).
6. Fraser, A., Doan, D., Lundy, M., Bevill, G., Aceros, J.: Pediatric safety: review of the susceptibility of children with disabilities to injuries involving movement related events. Injury Epidemiology 6 (12), 1-9 (2019).
7. Guzik-Kopytko, A., Wodarski, P.. Kowol, M. Wielgosz, A.: Use of toys in the process of rehabilitation of children with cerebral palsy (in Polish). Aktualne Problemy Biomechaniki 7/2013, 49–52 (2013).
8. Boyd, R.N, Morris, M.E., Graham, H,K.: Management of upper limb disfunction in children with cerebral palsy: a systematic review. European Journal of Neurology 8, 150–166 (2001).
9. Proença, J.P., Quaresma, C., Vieira, P.: Serious games for upper limb rehabilitation: a systematic review. Disability and Rehabilitation: Assistive Technology 13 (1), 95–100 (2018).
10. Rihar, A., Mihelj, M., Pašič, J., Sgandurra, G., Cecchi, F., Cioni, G., Dario, P., Munih, M.: Infant posture and movement analysis using a sensor-supported gym with toys. Medical and Biological Engineering and Computing 57 (2), 427–439 (2019).
11. Saleh, M.A., Hashim, H., Mohamed, N.N., Almisreb, A.A., Durakovic, B.: Robots and autistic children: A review. Periodicals of Engineering and Natural Sciences 8 (3), 12477–1262 (2020).
12. Adamus, W., Gręda A.: Multiple Criteria Decision Suport in Organizational and Management Chosen Problems Solving (in Polish), Badania Operacyjne i Decyzje 2, 1-11 (2005).
13. Fiore, P., Sicignano, E., Donnarumma, G.: An AHP-Based Methodology for the Evaluation and Choice of Integrated Interventions on Historic Buildings. Sustainability 12, 5795 (2020).
14. Jacquet-Lagreze E., Siskos J.: Assessing a Set of Additive Utility Functions for Multicriteria Decision Making, the UTA Method. European Journal of Operational Research 10(2), 151–164 (1982).
15. Peng C., Jing C. (2021) Evaluation of Tourism Landscape Ecological Environment Based on AHP and Fuzzy Mathematics. In: Sugumaran V., Xu Z., Zhou H. (eds) Application of Intelligent Systems in Multi-modal Information Analytics. MMIA 2020. Advances in Intelligent Systems and Computing, vol 1234. Springer, Cham.
16. Pawlewicz, K., Senetra, D., Gwiaździńska-Goraj, M., Krupickaitė, D.: Differences in the Environmental, Social and Economic Development of Polish–Lithuanian Trans-Border Regions, Social Indicators Research 147, 1015–1038 (2020).
17. Haddad M. et al. (2021) Use of the Analytical Hierarchy Process to Determine the Steering Direction for a Powered Wheelchair. In: Arai K., Kapoor S., Bhatia R. (eds) Intelligent Systems and Applications. IntelliSys 2020. Advances in Intelligent Systems and Computing, vol 1252. Springer, Cham.
18. Buyuk A.M., Temur G.T. (2021) A Framework for Selection of the Best Food Waste Management Alternative by a Spherical Fuzzy AHP Based Approach. In: Kahraman C., Cevik Onar S., Oztaysi B., Sari I., Cebi S., Tolga A. (eds) Intelligent and Fuzzy Techniques: Smart and Innovative Solutions. INFUS 2020. Advances in Intelligent Systems and Computing, vol 1197. Springer, Cham.
19. Łuczak A., Wysocki F.: Application of the Analytic Hierarchic Process in Analysis Motivation System of a Transport Firm (in Polish). Journal of Agrobusiness and Rural Development 4(10), 47–60 (2008).

20. Abidin, S.N.Z., Jaffar, M.M: Analytic hierarchy process (AHP) as a tool in asset allocation. AIP Conference Proceedings 1522, 384 (2013).
21. Saaty, T.L., Vargas, L.G.: Models, Methods, Concepts and Applications of the Analytic Hierarchy Process. Kluwer Academic Publishers, Norwell (2001).

Section 3: Materials for bio-medical engineering applications

State-of-the-art and innovative approaches in biomaterials and surface treatments for artificial implants

Vasily Efremenko[0000-0002-4537-6939], Oleksandr Cheiliakh[0000-0003-0805-0443], Oleksandr Azarkhov[0000-0003-0062-0616], Bohdan Efremenko[0000-0003-0438-6433], Yuliia Chabak[0000-0003-3743-7428] and Vadym Zurnadzhy[0000-0003-0290-257X]

Pryazovskyi State Technical University, Mariupol, 87555, Ukraine
vgefremenko@gmail.com

Abstract. The chapter presents the information on the state-of-the-art in biomaterials used for artificial orthopedics implants and the technologies for their manufacturing and treatment. The main groups of these materials are metallic substance (steels and alloys), bioceramics (bioinert and bioactive) and bioglasses. They should meet the requirements of bio-stability, biocompatibility, bioactivity, and desired mechanical properties (Young's modulus, strength, hardness, fatigue resistance, resistance to brittle fracture and wear). The chemical composition, structural components and the properties (physics-chemical and mechanical) of the up-to-date biomaterials are described in the chapter. The technologies of surface treatment and the coating deposition applied for artificial implants are covered as well. The chapter could be used as the teaching materials for the students in Bio-Engineering specializing in artificial implants fabrication.

Keywords: Biomaterials, Biocompatibility, Alloys, Bioceramics, Bioglass.

1 Requirements for artificial implant materials

The rapid aging of the world's population leads to increasing demand for the use of implants. Implants (from Latin «in + plantare») are a class of medical products used for implantation in a living organism as prosthetics (replacements of a missing organ) or as an identifier (e.g., the skin-fed chip with pet information). Implants always have an inorganic composition as opposed to a transplant. Implants applications are dentistry (implantation of teeth), otorhinolaryngology (cochlear implants), neurosurgery (brain implants), cardiology (heart valves (mechanical, biological prostheses), stents), ophthalmology (retinal implants), orthopedics (implants of paralyzed limbs, artificial joints, etc.), identification implants (identity setting, etc.).

Implants are made of biomaterials. The biomaterial is a synthetic or natural material used in a medical product (device) or in contact with biological systems (organisms). Biomaterials are designed to replace damaged organs and tissues of the body. Classification of inorganic biomaterials (for implants): (a) metals and alloys, (b) biopolymers, (c) bioceramics, (d) bio-glasses, (e) composite materials.

The human body is a very complex system that has ability for self-regulation and self-organization. It is very sensitive to external intervention, which leads to the biological and physicochemical interactions "implant material/body" as a result of direct contact with the technical system. During the life process, the implant embedded into the body is negatively affected by internal fluids and biological substances, pressure and temperature, mechanical stresses, and friction. All these lead to the degradation of implant material contributing to the reduction of its durability and functional properties. On the other hand, implant material may destructively effect the body's biological system through mechanical, toxic, carcinogenic, and energy factors.

The body's response on implant materials are (a) corrosion, (b) biodegradation (the destruction of a material as a result of the activity of living organisms with the loss of its basic specific properties), (c) mechanical destruction, (d) wear in the joints.

Corrosion resistance of metallic materials depends on its ionization potential The ionization potential for some metals are Li –3.040; Ca –2.870; Na –2.710; Mg –2.360; Al –1.660; Ti –1.630; Zn –0.763; Cr –0.744; Co –0.277; Ni –0.250; Pb –0.126; H 0.000; Cu +0.337; Ag +0.799; Pd +0.987. High corrosion resistance is attributed to the metals with positive ionization potential (Pd, Au, Ag, Cu). The more electronegative the element, the worse its corrosion resistance.

The requirements for **bio-stability, biocompatibility**, and **bioactivity** are imposed on implant material [1].

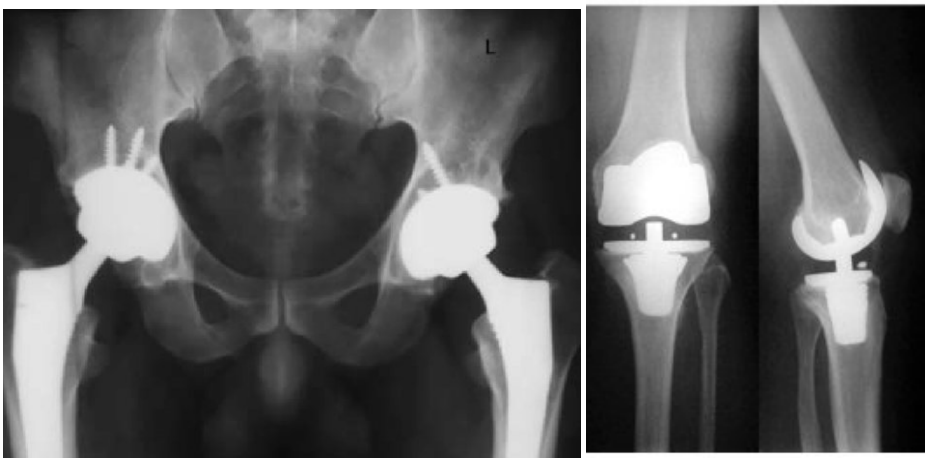
Bio-stability (bio-inertness) is the ability of a material to be not rejected by the immune system and to withstand structural changes and decomposition being in a biological environment. There should be no negative impact on the biological system. The bio-stability of metallic materials is largely determined by their susceptibility to corrosion, which depends on the standard chemical potential (the higher potential the higher corrosion resistance), the surface quality (the more rough surface, the higher the susceptibility to corrosion), the purity of the chemical composition (the purer, the higher corrosion resistance).

Biocompatibility is the ability of a material to be incorporated into a patient's body without causing clinical side-effects. **Bioactivity** is the ability of a synthetic material to provide the required positive reaction (healing) by the body in contact with non-living material. Bioactivity implies the interaction with surrounding tissues to form a direct bond with them, exhibiting ***osteoconductivity*** and (or) ***osteoinductivity***. Osteoconductivity is the ability of materials to adhere and bind osteogenic cells, provide biological fluxes, forming bone-binding cells. Osteoinductivity is the ability of a material to generate cells (chondrocytes, osteoblasts) from surrounding non-bone tissues leading to the accelerated healing of injured organ.

2 Overview of metallic bio-materials

2.1 Introduction to metallic bio-materials

The greatest share of orthopedic implants (up to 80 %) is made of metallic materials [1, 2]. These materials are indispensable in the bones' and joints' treatment (Fig. 1), the fracture risk of which increases significantly with age. Apart from bioinertness and biocompatibility, the materials for implants must meet certain requirements for chemical properties (absence of unwanted chemical reactions with the body's environment), mechanical characteristics (strength, Young's modulus, resistance to brittle destruction and wear) and biological properties (absence of rejection by the immune system, promoting the regeneration of bones' tissue). Recent studies have shown that modern materials for endoprosthetic require further improvement to improve their biocompatibility and biofunctionality. Traditional metallic biomaterials for implants are represented by three groups of alloys: Ti-based alloys, corrosion-resistant steels, and Co-Cr alloys. The latter have the highest wear resistance, stiffness, and strength as compared to stainless steels and Ti-alloys. Stainless steels tend to be superior in plasticity and cyclic strength over Ti-alloys and Co–Cr alloys. The metallic materials for the implants are not limited by the above alloys. Also, active research and development of new types of metallic biomaterials based on Nb, Zr, Ta, with improved biocompatibility and the high potential for use in implantology, are ongoing in various countries.



(a) (b)

Fig.1. X-ray images of metallic implants for (a) hip [3] and (b) knee [4]

2.2 Ti-based alloys

Ti-based alloys are the most commonly used biomaterials for the endoprosthesis [5] (Fig. 2). The reason for this is the highest (among other metals) biological compatibility of titanium in combination with its high corrosion resistance. Also, titanium alloys have an advantage over stainless steels and Co-Cr alloys in specific strength. The latter is due to the low specific weight of titanium (4.505 g/cm^3), which is important for the endoprosthesis. Currently, the widely used alloy Ti–6Al–4V is not biomedically applied due to the vanadium toxicity. Therefore the non-vanadium analogs have been developed such as Ti–6Al–7Nb and Ti–5Al–2.5Fe [6]. With the emergence of information that aluminum causes Alzheimer's disease (also a neurotoxic effect of aluminum is revealed), the further development of Ti-based alloys presumes the absence of aluminum in their composition.

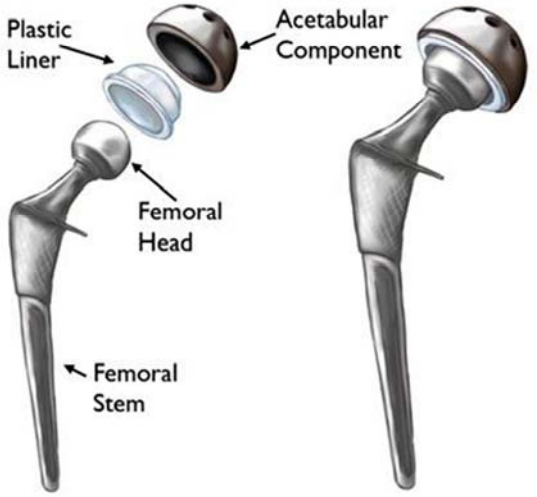


Fig.2. Hip prosthesis design [7].

Low-modulus β-Ti alloys. Currently, β-titanium alloys for implants have attracted particular attention from researchers regarding the problem of stress shielding associated with an uneven stress distribution between the implant and the bone [8]. Because the metallic implant's Young's modulus is slightly higher than Young's modulus of bone tissue, tension is perceived primarily by the implant, not the bone. This leads to bone atrophy with weakening or loss of the implant or bone's fracture after its removal. Therefore, the metallic materials should have Young's modulus as close as possible to the bone, which needs developing suitable metallic materials. In general, the titanium's elasticity modulus is the lowest among the major metallic biomaterials [1]. For example, Young's modulus of Ti-6Al-4V alloy (110 GPa) is significantly inferior to Young's modulus of corrosion-resistant steel (180 GPa) and cobalt-based alloy (210 GPa).

As is known, titanium alloys are divided by structure into three classes: α, β, and

α+β. Among them, the lowest Young's modulus is attributed to the alloys with β-structure. This is caused by the crystalline lattice type of β-alloys (BCC), which has a lower atomic density than that of the α-titanium lattice (HCP). Thus, one of the main approaches for biomedical Ti-alloys is the development of β-structure alloys alloyed by non-toxic elements (Nb, Zr, Mo, Ta, Fe, etc.). Such alloys Ti–16Nb–10Hf, Ti–13Nb–13Zr, Ti–12Mo–6Zr–2Fe, Ti–15Mo–5Zr–3Al, Ti–35.3Nb–5.1Ta–7.1Zr, Ti–29Nb–13Ta–4.6Zr (TNZT) have already been developed and tested. The Young's modulus values for β-type titanium alloys range from 40 to 80 GPa. The data presented in [9] convincingly prove the advantage of β-alloy TNTZ over (α+β)-alloy Ti–6Al–4V and SUS 316L stainless steel, introduced as intramedullary rods during healing a rabbit shin's fracture. The application of low-modulus β-alloy provided the most positive dynamics in indicators of animal's broken bone remodulation.

A new approach in this framework is the development of relatively inexpensive low-modulus materials alloyed with cheap elements such as Fe, Cr, Mn, Sn. The alloys of alloying systems Ti–12Cr, Ti–Mn–Fe, Ti–Sn–Cr, Ti–Mn, Ti–Cr–Sn–Zr, Ti–(Cr, Mn)–Sn are currently being developed. Since the crystal's orientation effects on Young's modulus of alloy, its level can be reduced by forming a special crystalline texture: Young's modulus of a monocrystallicβ-alloy (TNTZ), oriented in the ⟨100⟩ direction, has Young's modulus (35 GPa) close to the upper part of the human bones range (about 30 GPa).

The disadvantage of β-alloys is the low (compared to α- and (α+β)-alloys) strength and fatigue resistance values, which limits their use in implantology. This problem is being solved by applying cold plastic deformation with high reductiondegrees, which provides the 1,5-fold increase in β-alloys' strength while maintaining a low elasticity modulus [1]. However, such treatment does not increase fatigue resistance. An alternative approach is the application of heat treatment (quenching with subsequent aging) to precipitate α-phase and ω-phase from the β-solid solution, which causes a strengthening effect. However, this results in Young's modulus growth. Short aging holding is optimal since it increases fatigue resistance with keeping Young's modulus less than of (α+β)-alloys. The dynamic strength of β-alloys can also be increased by the introduction of the dispersed particles, for example, TiB and Y_2O_3; in this case, cold deformation increases not only the strength but also the fatigue resistance [10].

Ti-alloys having a self-tunable elastic modulus. A separate group of Ti-alloys are alloys with self-tunable Young's modulus, which have been actively developed recently. Their development was due to a problem related to the spinal locking devices, which have to bend during surgery to reproduce the physiological curvature of the patient's spine [1]. As shown above, implants of low Young's modulus material must be used for the rapid healing of the bone. However, such materials have an increased ability to spring when bending, which complicates the adaptation of the device for the spine fixation in a limited space inside the patient's body. Thus, the Young's modulus of a Ti-alloy for such devices should be low enough to avoid bone atrophy, but high enough to be less springy for the spine surgery devices.

The main direction in the self-tunable titanium alloys development is the using of deformation-induced phase transformation's, which occurs in Ti-alloys with an initial low Young's modulus. Self-tunable titanium alloys include alloys of Ti–Cr, Ti–Mo, Ti–

Zr–Cr, Ti–Zr–Mo and Ti–Zr–Mo–Cr alloying systems [11]. Cold rolling increases Young's modulus of the Ti-12Cr alloy from 60-70 GPa (quenched state) to 85 MPa due to the ω-phase formation [1]. In contradistinction to the Ti-12Cr alloy, a similar treatment does not change Young's modulus of the TNTZ titanium alloy, which is not prone to deformation-inducedβ→ωtransformation.

Ti-Zr alloys with high calcium phosphate deposition resistance. Implants include a separate group of removable implants (such as screws or implants installed in a growing child's body) that are removed after some time after surgery. Removal can cause problems with calcium phosphate formation on the implant surface [1]. The formation of this substance is an important mechanism that uses the body for healing fractures, so the acceleration of this process is even stimulated by special technologies for surface treatment of fixed implants. The implant's accretion with a bone substance (calcium phosphate) secures it to the bone and complicates removal after treatment completion, which can lead to bone fracture. Therefore it is necessary to suppress the calcium phosphate deposition on surface ofthe removable implants.

This problem is solved by the use of Ti-Zr alloys. Titanium and zirconium have unlimited solubility in each other, while the introduction of zirconium increases the strength of titanium alloys. Zirconium has the same low toxicity to the body as titanium, i.e. it shows high biocompatibility. An important advantage of zirconium is its ability to prevent calcium phosphate deposition on titanium alloys, which is achieved with a content of at least 25 % Zr. On this principle, a series of Ti-Zr biomedical alloys such as Ti–Zr–Nb–Ta, Ti–Zr–Nb, Ti–30Zr–5Mo, Ti–Zr–Al–V have been developed [1]. A series of titanium alloys with a self-tuning Young's modulus (Ti–30Zr–7Mo, Ti–30Zr–5Cr, Ti–30Zr–3Cr–3Mo) are easily separated from the bones due to the alloying of 30 % zirconium [12].

Titanium alloys with the superelasticity and shape memory effect. The widely known Nitinol alloy (50% Ni – 50 % Ti), which has a pronounced shape memory effect and has been used in the manufacture of intravascular implants (stents), is now recognized as dangerous due to nickel toxicity. Therefore, studies on the development of nickel-free Ti-based alloys with superelasticity are continuing. According to the alloying system, these alloys are divided into four main groups [1]: Ti–Nb (Ti–Nb–O, Ti–Nb–Sn, etc.), Ti–Mo (Ti–Mo–Ga, Ti–Mo–Ge, etc.) [24], Ti–Ta (Ti–50Ta–4Sn, Ti–50Ta–10Zr, etc.), Ti–Cr (Ti–7Cr–1.5Al). All alloys have a β-structure, where the β-phase is unstable to reverse martensitic transformation under deformation. Adding a small amount of oxygen or nitrogen increases the elastic characteristics of Ti-alloys by reducing the start temperature of the martensitic transformation. An additional increase in the superplasticity effect is achieved by forming a special texture by the use of thermo-mechanical treatment, as well as multi-stage heat treatment, during which the dispersive particles of the α-phase precipitate making strengthening effect.

2.3 Stainless steels

Apart from Ti-based alloys, there are other metallic materials that are widely used for the implants in orthopedics. The traditional groups of such materials include stainless steels are widely used in various technologies [1, 5]. Stainless steels have the longest

term use among metallic materials for implants. The austenitic stainless steels (typical representatives are SUS 316/316L steels) make a wide range of biomedical products used in implantology such as plates, wires, screws, stents, catheters, electrical stimulants, etc. SUS 316 steel has the following chemical composition (wt.%): 16-18 Cr; 10-14 Ni; 2-3 Mo; ≤0,08 C; ≤2 Mn; ≤0,75 Si; ≤0,10 N.

The main advantage of stainless steels is the excellent corrosion resistance combined with the increased structural strength. A significant drawback is the high nickel content, which has high biotoxicity. In this regard, studies on the development of nickel-free stainless steels for bioapplications are continuing. These studies are conducted in the direction of nickel replacement with other strong austenitic forming elements like manganese and nitrogen. Examples of such steels are Fe–(15…18)Cr–(10…12)Mn–(3…6)Mo–0.9N, Fe–15Cr–(10…15)Mn–4Mo–0.9N, Fe–21Cr–9Ni–3Mn–0.41N(Rex 734), Fe–18Cr–18Mn–2Mo–0.9N, Fe–(19…23)Cr–(21…24)Mn–(0.5…1.5)Mo–0.9N (BioDur108), Fe–(16…20)Cr–(12…16)Mn–(2.5…4.2)Mo–(0.75…1)N [13]. These steels are made by liquid metallurgy methods using refining technologies, such as electroslag remelting or electroslag remelting under tension.

As the biocompatibility of manganese is still insufficiently proven, there is a promising trend to produce high-chromium steel with high nitrogen content by solid-phase saturation of steel with nitrogen. In this case, ferrite stainless steels such as Fe–24Cr or Fe–24Cr–2Mo are used as the starting material. The final product with the required shape and size are manufactured from these steels, after which long-term heating in a nitrogen gas environment is conducted. Nitrogen diffuses into steel, its content increases up to 0.9-1.0 %, leading to the γ-phase stabilization [1]. It is not possible to obtain steel with such a high nitrogen concentration by conventional metallurgical technology (smelting and deformation) due to the high material's brittleness. The nitrogen diffusion saturation method can only be used for thin-walled products. Studies have been conducted to show the prospect of nitrous nickel-free stainless steel as a material with high biocompatibility and non-toxicity.

2.4 Co-based alloys

Cobalt-based alloys are characterized by high corrosion resistance and superior wear resistance [1, 5], which makes them an indispensable material for the artificial joints manufacture. Cobalt belongs to polymorphic metals: it has a ε-structure (HCP lattice) at room temperature and a γ-structure (FCC lattice) in the high-temperature region. Due to a large number of dislocation slip systems, the γ-structure provides higher plasticity being therefore desirable for cobalt-based alloys. To obtain the γ-structure, nickel is usually added to the cobalt, which increases the packing defects' energy in the crystalline lattice, which stabilizes the γ-phase. The alloys of the Co–Cr–W–Ni system (ASTM F90) have a γ-structure at room temperature, which provides them with high technological plasticity and facilitates the implants' manufacture.

Since nickel relates to potentially dangerous for health elements, the nickel-free cobalt alloys for biomedical use have been developed. Among them, the most famous are the Co–Cr–Mo alloys (ASTM F75). One of these alloys is "Vitallium", containing 65Co–30Cr–56Mo. The absence of nickel in the alloy composition causes high energy

of packaging defects, and therefore in the structure of the alloy at room temperature, there is a ε-phase, which reduces its technological plasticity. This disadvantage is compensated by the addition of 0.2 % nitrogen, which allows the get complete γ-structure. However, the resulting γ-structure is unstable to deformation martensitic transformation γ→ε, which reduces the ability to cold plastic deformation. Therefore, hot plastic deformation is used for forming products of Co-Cr-Mo alloys, which is also used to refine the structure and enhance mechanical properties. For this purpose, the Co–27Cr–5Mo–0.16N alloy can also be subjected to a heat treatment during which the «eutectoid (ε+Cr_2N)→γ» transformation takes place, accompanied by grain refinement and mechanical properties enhancement.

2.5 Zr-based alloys

Zirconium alloys are increasingly used as biocompatible materials, mainly due to their low magnetic susceptibility. The problem is that in the case of the metal implants presence in the patient's body, the artifacts (areas with a distorted picture) may appear during magnetic resonance imaging on MRI (Magnetic Resonance Imagine). The reason is the difference in the magnetic susceptibility of metals and living tissues [1]. The magnetic susceptibility of Ti ($3.2 \cdot 10^{-6}$ cm^3/g) is higher than that for living tissues but much lower than for ferromagnetic Fe and Co. For Zr, this number is even lower to be of $1.3 \cdot 10^{-6}$ cm^3/g.

Zr and Ti are in one group and they have similar physics-chemical properties, i.e. Zr is promising for biomedical use. In this regard, Zr-based alloys (Zr–Nb, Zr–Mo) have been developed [14]. The introduction of 3-9 wt.% Nb reduces the magnetic susceptibility of zirconium to $(1.1\ldots1.2) \cdot 10^{-6}$ cm^3/g [10]. In this niobium concentration range Zr–Nb alloys have different microstructures: α′ (3 % Nb), (α′+ω+β) (6 % Nb), and (ω+β) (9 % Nb), among which ω-structure has the lowest magnetic susceptibility. Molybdenum has a similar effect: it is found that with the introduction of 3 % Mo, the magnetic susceptibility of the Zr-alloy becomes less than $1.1 \cdot 10^{-6}$ cm^3/g, which is the lowest value for known metal alloys.

2.6 Ta-based and Nb-based alloys

Ta and Nb are non-toxic elements used for alloying titanium and zirconium alloys. They exhibit similar high biocompatibility and corrosion resistance. The biocompatibility of tantalum is close to that of titanium [1], which is why Ta has been used since the 1940s in dentistry and orthopedics (as the radiographic bone markers, vascular clamps, cranial plates, etc.). Successful use of porous tantalum in arthroplasty of joints, spine surgery has been reported when bone grows into tantalum pores, forming a strong composite. In this regard, porous tantalum is actively investigated in the direction of determining its mechanical properties under static and dynamic loading conditions. Porous tantalum in the form of commercial Hedrocel™ material has been found to have a microhardness of 240-393 HV, a tensile strength of 63±6 MPa, bending strength of 110±14 MPa, and compression strength of 60±18 MPa. The plasticity of this material is inferior to other metallic biomaterials, but it is higher than the plasticity of bone or alternative materials

such as ceramics or ceramic-filled composites. The Young's modulus of the composite «C-frame/Ta-foam» is 1.15±0.86 GPa to be close to the bone modulus (1.08 ± 0.86 GPa) [1].

Another way of using tantalum is to make the stents. Comparison of Ta-stents with the stents made of Nitinol, stainless steel SUS 316L, and Co-Cr alloy revealed that Ta-stents and Nitinol stents produce the least pronounced artifacts during MRI and angiography. Unlike other metals that are electropositive, tantalum exhibits electronegativity. Therefore, it was expected to reduce the stent's ability to thrombus formation, since blood elements (fibrin, platelets) also exhibit electronegative properties. However, it has been found that tantalum does not outperform other metallic biomaterials for blood compatibility, which explains the activation of blood coagulation by the presence of negatively charged surfaces [1]. On the other hand, the dopping of TiO_2-films of with the Ta^{5+} ion increases its blood compatibility.

Nb-based biomedical alloys are currently being developed. The Nb–2Zr alloy subjected to equal channel angular pressing acquires high mechanical properties in combination with high corrosion resistance in body fluids [1]. A promising Nb–28Ta–3.5W–1.3Zr stent alloy has been developed having high corrosion resistance and low magnetic susceptibility, which significantly reduces the manifestation of artifacts in MRI.

2.7 Biodegradable alloys

Mg-based alloys. There are different implants in the form of locking structures that are not intended for long-term use: after reaching a certain result in treatment, they should be removed from the patient's body. Removal of a temporary implant is associated with undesired surgery causing trauma and risk of complications. In this regard, a new direction in Bio-Engineering has been actively developed in recent decades, which is to design the biodegradable materials for the manufacture of temporary implants [1, 15]. The concept of biodegradable materials is that the implant is not removed from the patient's body after performing certain tasks, but remains in it, gradually degrading (dissolving) due to chemical interaction with the internal environment of the body. The designing of biodegradable metallic materials is currently an ambitious scientific challenge that attracts the attention of material scientists from different countries. Because of the problems that arise, the development is underway in several directions that are currently actively developing.

A significant disadvantage of metallic Fe-based, Ti-based and Co-based orthopedics biomaterials is their high strength, which is much higher than the bones' strength. When using metal implants, the «stress shielding» effect can appear, when external loads is mainly concentrated on the implant while the bone remains weakly loaded. This leads to a decrease in bone density after its growth with repeated destruction [1].

In addition to durable metallic materials, calcium phosphate ceramics and bio-soluble polymers are increasingly used in implantology: polyglycolic (PGA) and polylactic (PLLA, PDLA) acids, their copolymers (PLGA), hydroxy derivatives of alkanoic acids (PHA). The disadvantages of these materials are low and unstable strength (tensile strength is 20-70 MPa, Young's modulus is 1.2-6.9 GPa), which makes it impossible to

use them in manufacturing loaded implants. These materials are also brittle, non-durable substances.

An alternative to durable metallic biomaterials, on the one hand, and non-durable biopolymers, on the other, are magnesium-based alloys. Magnesium refers to light metals (density – 1.74 g/cm^3). Its Young's modulus and tensile strength are 41-45 GPa and 113 MPa, respectively, to be higher than that of the polymers, and are as close as possible to the mechanical properties of human bone tissue (3-20 GPa and 30-150 MPa, respectively), compared to other metallic biomaterials.

Magnesium and its corrosion products have high biocompatibility [15]. Magnesium is important for the metabolism processes of the human body as a cofactor of about three hundred enzymes formation. Magnesium corrosion products in the body are Mg^{2+} ions, hydroxide ions OH^- and hydrogen (H_2). Mg^{2+} ions promote rapid healing of the injured tissues without causing cellular toxicity; their excess is excreted through the urinary tract. Magnesium alloys are easily soluble in water and other corrosive environments, making them potentially attractive biosoluble materials. The first use of magnesium alloys as implants date back to the late 19th century [15].

The disadvantages of pure magnesium are its high brittleness, the hydrogen release during the magnesium dissolution (hydrogen accumulates near the implant and adversely affects the bone tissue formation), undesirable increase in the body's alkalinity level. Therefore, the task arises to develop the magnesium alloys with improved mechanical properties and corrosion resistance.

The possibility of using commercially available magnesium alloys of various alloying systems such as WE43 (alloy systems of Mg–Y–REM–Zr), AZ91 (Mg–Al–Zn), AZ31 (Mg–Al–Zn–Mn), LAE442 (Mg–Li–Al–REM–Mn, etc), as biosoluble materials is being intensively investigated [1, 15]. For example, alloy AZ31 contains about 2.5 % Al, 0.9 % Zn, 0.3 % Mn, and the Mg is balance. As these alloys contain chemical elements with unproven biocompatibility, Mg-based alloys consisting of non-toxic elements (Zn, Ca) have been developed. The screws of commercial biosoluble magnesium alloys (Magnezix®, Sintellix, Milagro®, DePuy Mitek (the latter made of Mg-Y-REM-Zr alloy) have emerged on market in recent years.

One of the major drawbacks of Mg-based alloys is the high dissolution rate in the physiological environment. Rapid implant's degradation leads to the formation of gas cavities or breaks at the bone/implant interface. In this regard, studies on improving the corrosion resistance of biomedical Mg-alloys are being conducted. One of the way is to deposit the protective coatings. These technologies include [1, 15]: (a) obtaining a Al_2O_3 layer using plasma implantation of Al and O atoms and ions, (b) electrochemical deposition of fluorinated hydroxyapatite (FHA) and Ca-P coating, (c) impulse deposition of a Ca-deficient hydroxyapatite coating, (d) the formation of «$CaHPO_4 \cdot 2H_2O$/polycaprolactone» composite coating, (e) hydroxyapatite deposition by magnesium anode dissolution method, (f) the deposition of biodegradable polymers, etc. Recent studies have found potentially high biocompatibility of graphene and its derivatives as a protective coating on magnesium alloys. These coatings have been reported to promote the adhesion and reproduction of human osteoblastic cells and facilitate the mesenchymal stromal cells differentiation into osteoblasts. The composite «Graphene oxide + Hydroxyapatite» coating forms, when initially deposited graphene oxide being held in the

body's fluid simulates the hydroxyapatite deposition. This coating significantly increased the corrosion resistance of AZ91 magnesium alloy.

Another direction is increasing the physical and chemical properties of Mg-based alloys by adding alloying elements (Pd, Sc, Zr, Nd). The addition of 0.1 % Sc reduces the dissolution rate of the standard ML-10 alloy in saline (Venofundin), while the addition of 0.05 % Ag increases it. It was found that the addition of 1.25–1.3 % Zr and 2.9–3.1 % Nd provides an increase in the stability of the magnesium alloy's mechanical properties under the prolonged holding in saline (Gelofuzin) [16]. The absence of a toxic effect for Mg-Zr-Nd alloy's degradation products on a living organism was proved and the positive dynamics of reparative osteogenesis was established.

Fe-based alloys. Pure iron and Fe-based alloys were proposed for use as bio-soluble materials in the early 2000s [1]. These alloys have a higher strength and Young's modulus than magnesium alloys. However, implants made of iron take years to completely dissolve, meaning that their corrosion rate is too slow for the implant to disappear without long-term side effects. Iron oxidation products (Fe^{2+} and Fe^{3+} ions) are important for vital activity and do not exhibit toxicity at expected concentrations. However, due to the very low biodegradation of pure Fe, such stents exhibited adverse reactions associated with the accumulation of insoluble iron hydroxide, preferably in the cells at the implantation site. Afterwards, the migration of iron-rich cells throughout the body was observed, causing chronic inflammation.

To increase the iron's biodegradation rate, Fe-Mn alloys have been developed that use the deformation-induced ε-martensite. ε-martensite areas are distributed in the austenite enhancing the mechanical properties and accelerating the alloy's corrosion rate [19]. However, the biodegradation rate of Fe-Mn alloys remains an order of magnitude lower than Mg-based alloys. The alloying of Fe-Mn alloys with palladium in combination with heat treatment further increases the strength and corrosion rate due to the intermetallic inclusions (Fe, Mn)Pd precipitation [20].

Zn-based alloys. The first data on the use of zinc alloys for biosoluble implants have emerged over the last few years [1]. Like magnesium and iron, zinc is a vitally needed element in the human body; it is a part of many enzymes and other proteins, ensuring the optimal flow of metabolic processes, as well as the growth, divide and function of cells. From these positions, zinc ions released from the implant during degradation can normally integrate into the body's metabolic activity without causing toxic effects. The zinc electrode potential (−0.762 V) is intermediate between magnesium (−2.372 V) and iron (−0.444 V), so pure zinc corrodes with favorable kinetics, i.e. faster than iron but slower than magnesium alloys. This is due to the formation of the passive layers on the surface. The degradation products of zinc alloy (Zn^{2+} ions) are considered to be sufficiently biocompatible and even to have anti-inflammatory effects. Zinc alloys have a high ability to manufacturing and processing; unlike magnesium alloys, they can be smelted in an atmosphere of air.

The addition of iron controls the biodegradation rate of pure zinc [17]. The addition of 1.3 % Fe doubled the zinc corrosion rate due to the micro-galvanic effect caused by the δ-phase ($Zn_{11}Fe$) formation. Histological examinations performed 14 weeks after the introduction of the implant to the rat showed no adverse systemic effects in the tissues adjacent to the implant.

The use of zinc implants is constrained by its low mechanical properties (tensile strength (UTS) is 50-140 MPa, elongation (EL) is 0.3-5.8 %). Currently, zinc alloys with a higher properties complex are being actively developed. The main direction is the introduction of alloying elements that provide solid-state hardening and intermetallic phases strengthening. Alloys of various alloying systems: Zn-Mg, Zn-Ca, Zn-Sr, Zn-Al, Zn-Cu, Zn–Al–Mg–Bi, Zn-Li, Zn-Ag, etc. have been developed [1]. In the latter two cases, the enhancement of strength is ensured by the intermetallic compounds α-$LiZn_4$ and ε-$AgZn_3$, respectively. Currently, the following properties of zinc alloys have been achieved: Zn–5Al (UTS=300 MPa, EL=16 %), Zn–4Li (UTS=440 MPa, EL=13,7 %), Zn–7Ag (UTS=287 MPa, EL=32 %) [1]. The resources for further improving mechanical properties are heat treatment an d cold plastic deformation. Prospects for further research in this area are to conduct long-term clinical investigations and analyze their results to prove the safety for humans of designed biosoluble materials based on magnesium, iron and zinc.

As shown above, the field of metal alloys for biomedical use shows high dynamics in the direction of the new materials development with improved biofunctional properties and biocompatibility. This leads to formation of a new branch of Materials Science – Bioengineering Metal Science.

3 Bioceramics

Ceramics came from ancient times meaning the products and materials obtained by sintering clays and their mixtures with mineral additives. Plasticity of clay was used by man at the dawn of human civilization, and the first clay products were the sculptures of humans and animals that have come from the Paleolithic Era. However, the annealing process which gives the clay products hardness, water-resistance and fire resistance, began to be widely used only in the Neolithic Era. Terracotta architectural details, tiles, water pipes, etc. were made both in ancient Greece and in ancient Rome, where the manufacturing of bricks was specially developed. The word “ceramics” came from the ancient Greek language (“keramos” means “burnt clay”, “keramici” means “pottery”).

For the centuries ceramics industry traditionally produced glass, cookware, constructional and refractory materials. After the Second World War ceramics have found new applications in high-tech advanced industries of electronic engineering. Ceramics are the most important elements of computer components, including capacitors, substrates of integrated circuits, thermistors, etc. The concept of “ceramics” has recently been transformed. In addition to traditional materials made of clay, materials derived from pure, simple and complex oxides, carbides, nitrides and the like began to be attributed to ceramics. Now ceramic is understood as any polycrystalline materials obtained by sintering non-metallic powders of natural or artificial origin. The most important components of modern ceramics are the oxides of Al, Zr, Si, Be, Ti, Mg, the nitrides of Si, B, Al, the carbides of Si and B, their solid solutions and composites.

Ceramics have a wide range of physical and chemical properties as compared with metals and polymers. The ceramics are not oxidized and are stable in a much higher

temperature range than that of metals. The modulus of elasticity of ceramic fibers is much higher than that of metals. Among the ceramics, it is easy to find materials with large or small (even negative) coefficients of thermal expansion. Ceramics include the-materials belonging to dielectrics, semiconductors, and conductors, and even superconductors. The fabrication of ceramic is less energy-consuming and more environmentally friendly compared with the metallic materials. Ceramic materials have higher corrosive resistant, they are more biocompatible than metals and polymers, and this allows them to be used for implantation of artificial organs, and as structural materials in biotechnology and genetic engineering.

Up-to-date ceramics can be divided into (a) structural ceramics, (b) functional ceramics, and (c) bioceramics. Bioceramics are ceramics that can be used as biomaterials due to their properties [18]. Bioceramics and bioglass belong to ***ceramic biomaterials***. Bioceramics are widely used for dental and bone implants. Artificial joints are usually coated with bioceramic materials to reduce wear and prevent inflammatory tissue reactions. Other bioceramics applications are the elements of pacemakers, artificial kidney apparatus and respirators. Bioceramics are also intended for use in extracorporeal circulation systems (in dialysis) or in special bio-reactors.

The main advantages of bioceramics are bioinertness in the human body, high hardness and wear resistance, which makes it useful for bone implants, dentures, and artificial joints. Such features as good esthetic appearance and electrical insulation are also attractive for specific biomedical applications. The global market for medical ceramics and ceramic components is constantly growing.

Bioceramics should have the structural similarity of bone tissue, should ensure the absence of undesired chemical reactions with tissues and body fluids, should exclude the reactions from the body's immune system while they have to provide the fusion of the implant with bone tissue, and the stimulation of bone formation (osteosynthesis). Also, the mechanical characteristics of bioceramics should be close to those of the bone, otherwise, the differences in elastic modulus may result in implant loss due to resorption of bone in contact with the implants.

3.1 Biological and biostable ceramics

By its reaction to the effect of the human body, bioceramics are divided into (a) bioinert ceramics and (b) bioactive ceramics [19, 20]. Bioinert ceramics have high chemical inertness in the human body environment and high strength causing some negative consequences. The chemical inertness of the ceramics results in the bone not being able to grow into the implant and the place of contact is filled with a fibrous connective tissue that does not provide strong contact. The high strength of ceramics leads to increased stiffness of the implant. Thus implant will bear almost the entire mechanical load while the bone will be not loaded enough. In the absence of the usual load, bone cells start the process of bone dissolution, which leads to osteoporosis which is a disease of reducing bone mass due to the growth of their porosity. Therefore the risk of bone fracture in contact with the implant increases significantly.

Bioactive ceramics promote rapid healing of the fracture due to the involvement of the implant into this process. When the implant is dissolved, then its active elements and substances are released to be involved in the repairing of the bone tissue.

By structure, bioceramics are divided into monolithic and porous (Fig. 3). Monolithic ceramic is a compacted material without pores ensuring the highest mechanical strength and, accordingly, it has a high resistance to brittle fracture. Porous ceramics have a large number of pores (cavities) results in its low mechanical strength. However, the porous surface of bioceramics provides a larger contact surface between the biomaterial and the growing bone, leading to the formation of higher number of chemical bonds; as a result, the bone grows into the pores of the ceramics. Therefore, bioceramics are attempted to be manufactured with macropores (pore size greater than 100 μm) by introducing pore-forming agents that are either volatile or easily water-soluble compounds (for instance, naphthalene, sucrose, $NaHCO_3$, NaCl, gelatin, polymethyl methacrylate etc.). The increase in the size of macropores from 150 to 1220 μm does not accelerate bone healing in contact with implant. In addition to macropores, any ceramics have micropores (pore size less than 10 μm), which are formed in the course of ceramics sintering. The size of the micropores depends on the sintering conditions (temperature and duration of the process). The foaming is also used in the fabrication of porous ceramics.

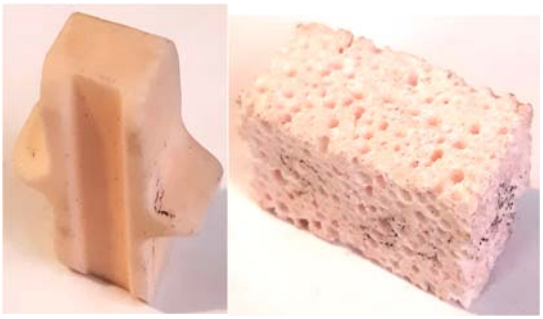


Fig.3. The appearance of Al_2O_3: (a) monolithic and (b) porous

Biological ceramics. The biological ceramics are bones and tooth's elements (dentin and enamel) [21]. The bones are the constituents of a skeleton. ***Bones*** are made of bone tissue, which is a type of connective tissue. The bones are composed of cells (osteocytes) and an intercellular substance consisting of collagen, glycoproteins, and mineral components (mainly crystals of hydroxyapatite $Ca_{10}(PO_4)_6(OH)_2$). Due to its composition, the bone has both flexibility and strength. The tensile strength of a bone is about the same as of copper, and 9 times greater than that of lead. The bone can withstand the compression load of 10 kg/mm² (similar to cast iron). A pressure of 110 kg/cm² is required to break the human rib. Under external load increase, the bone responds with an increase in strength while a bone loses its strength with the load decrease. ***Dentin*** (from the Latin. *dens*, *dentis* - tooth) is mineralized tissue of the tooth. It consists mainly of hydroxyapatite (70-72%), impregnated with calcium salts and imbued with dentinal tubules and collagen fibers (fibrils). The thickness of the dentin layer ranges from 2 to

6 mm. Dentin hardness reaches 58.9 kg/mm². Dentin forms the most part of the tooth. Dentin coating is different, depending on its location: the tooth crown is coated by tooth enamel, and at the root is coated by tooth cement.

Tooth enamel is a solid, mineralized tissue that covers the outside of the tooth crown and protects dentin and pulp from external irritants. The thickness of the enamel layer is maximal in the area of mounds of chewing teeth, where it reaches 2.3-3.5 mm and minimum in the area of the tooth neck (about 10μm). On the approximate surfaces of the permanent teeth, it is usually equal to 1-1.3 mm. The enamel layer on temporary teeth is not exceeding 1 mm. Enamel is the hardest tissue of the human body. It contains 95% of mineral substances (mainly hydroxyapatite, fluor-apatite, carbonate apatite, etc.), 1.2% organic, 3.8% water (free and bound to crystals, and organic components).

Biostable (bioinert) ceramics. The main varieties of biostable ceramics are ***alumina*** (corundum, aluminum oxide α-Al_2O_3) and ***zirconia*** (zirconium oxide ZrO_2) [18]. These types of ceramics are not dissolved in the body environment and therefore are considered biostable (bioinert). The most characteristic of their feature is good wettability with aqueous solutions, which allows them to adapt well to bone tissue. The wetting angle of saline fluids for Al_2O_3 is 45°. This promotes better cell adhesion to the ceramic surface. For steel and polymer, the wetting angle is 72-87° and 80° respectively, which impairs adhesion. Al_2O_3 and ZrO_2 have high friction resistance due to their high hardness and strength, which is important for artificial joints. Therefore, the surface of the friction joints is covered with a layer of the mentioned ceramics.

Corundum is a natural mineral that belongs to the class of oxides/hydroxides. It has a hexagonal crystal lattice in which oxygen ions occupy positions at the nodes and aluminum ions are in octahedral pores where each aluminum ion is surrounded by six oxygen ions. The type of bond between atoms is ionic. This causes high strength and low ductility of this material. In nature, corundum is in the form of minerals of bluish or yellowish-gray colours. There are also varieties of corundum with different colors: red (ruby), blue (sapphire), yellow (oriental topaz), and others. The properties of corundum are: melting point - 2072 °C, density - 3.97 g/cm^3,bending strength - 500 MPa, tensile strength - 4100 MPa, hardness - 2000 HV, Young's module - 380 GPa. Corundum has been used in medicine since 1969, mainly for the manufacture of articular prostheses. Corundum covers a cup of the acetabulum in the hip joint. Corundum has been found to have exceptional wear resistance that reaches the highest values in the corundum-corundum friction pair. The products of corundum wear are less damaging to the body as compared with a metallic surface.

Bioceramic Al_2O_3 products are made of powders by pressing and sintering. The powder is obtained from the mineral bauxite $Al(OH)_3$ which is decomposed at high temperature (1150 °C): $2Al(OH)_3 = Al_2O_3 + 3H_2O$. Obtained Al_2O_3 then is ground to a size less than 1 μm and compacted. The corundum is manufactured by three-staged solid-state sintering. In the first stage, the corundum particles are in contact with each other, but they are not physically coupled. The compacted powder is heated to a temperature of about 2/3 of the corundum melting point. At this stage, the “bridges” between the particles appear connecting them. The area of adhesion gradually increases, while the volume of pores decreases, the density of the material increases. The particles coalesce to reduce the surface energy of the entire system. In the last step, the individual particles

become indistinguishable. At the same time in the third stage, there is an increase of grains, which can lead to a decrease in strength.

Al_2O_3 ceramic materials have a very low ability to plastic deformation. If a crack originated in corundum, its development will be very easy, since plastic deformation is not possible. Therefore, high brittleness of ceramics is its main drawback. It is possible to improve the resistance of the ceramic to brittle fracture by controlling the sintering process and reducing the number of pores.

High-temperature modification of zirconium oxide has a cubic face-centered crystal lattice in which oxygen ions occupy tetrahedral pores and zirconium ions are located in the nodes of the lattice [19]. The bonding type is ionic. This causes high strength and low ductility of this material. Depending on the temperature, ZrO_2 may have a different crystalline lattice: monoclinic at room temperature, which turns to tetragonal at 900-1220 °C and then to cubic at 2370 °C. Under cooling, when the tetragonal lattice becomes monoclinic, there is a 3-5% increase in volume, which leads to ceramics cracking. To prevent this, ZrO_2 is stabilized by the addition of oxides (Y_2O_3, HfO_2, Al_2O_3), which inhibit the mentioned transformation. More than 95% of zirconium oxide has the structure of the metastable tetragonal phase. The properties of ZrO_2: melting point – 2715 °C, density – 6.08 g/cm^3,bending strength – 500-1000 MPa, tensile strength – 2000 MPa, hardness – 1200 HV, Young's module –210 GPa.

If a crack is originated in zirconium oxide, its development occurs with the transformation of the tetragonal phase into a monoclinic phase ahead of the crack. This transformation consumes the crack energy, which means that crack rapid development is delayed. As a result, zirconium oxide has a higher toughness as compared to aluminum oxide. When heated to 60-500 °C, a spontaneous transformation of the tetragonal phase into a monoclinic can occur, which negatively affects the mechanical properties of ceramics. In addition to its higher toughness, zirconium oxide has the advantage that it can be used to make articulated heads of small size (up to 22 mm) and it is adjustable for various bushing materials (Ti-alloys, stainless steel, Co-Cr alloy).

In dentistry, zirconium oxide is used to make dental crowns. These crowns have the following advantages: (a) hypoallergenic (zirconium is the most biocompatible and hypoallergenic material used in dentures), (b) minimum thickness (0.3 to 0.4 mm), so the doctor sharpens the tooth only slightly minimizing the risk of inflammatory processes, (c) high strength (they are 30 times stronger than ceramics, which allows them to be used not only for the front teeth but also for the chewing teeth), (d) natural appearance, (e) service life is not less than 15 years.

3.2 Ca-P-based bioactive ceramics

Calcium phosphate cement. In 1986, Brown and Chow designed CaP cement (CPC), which consisted of a combination of tetracalcine phosphate (TECP) and anhydrous calcium hydrophosphate ($CaHPO_4$). These materials have favorable osteoproductive properties as bone substitutes [18]. CPC cement has a weakly crystalline structure similar to the mineral phase of normal bone tissue. To prepare CPC the combination of the powders is mixed in a solution of sodium phosphate. The paste turns into a solid material in 10-15 minutes and it has a compressive strength close to the normal trabecular

bone. CPC cement has low tensile stress (of 1-10 MPa) but it has a higher compressive strength of 10-100 MPa. CPCs are not resistant to shear forces; therefore they are not suitable for diaphyseal fractures.

Hydroxyapatite (HA). This is a mineral $Ca_{10}(PO_4)_6(OH)_2$ belonging to apatite group, hydroxyl analog of fluorapatite and chlorapatite [19, 22]. The complete formula of natural apatite is $(Ca, M)_{10}(PO_4,Y)_6(OH,X)_2$, where M are metal cations (Mg^{2+}, Na^+, K^+, Sr^{2+}, Ba^{2+}, etc.), Y are anions (CO_3^{2-}, H_2PO^{4-}, HPO_4^{2-}, SO_4^{2-}, etc.), X are F^-, Cl^-, CO_3^{2-} and others.It is the main mineral component of bone (about 50% of the total bone mass) and teeth (96%) (Fig. 4). Its main components - calcium and phosphorus - are two vital elements responsible for mineralization, integrity, and firmness of bones. HA stimulates osteogenesis. The synthetic hydroxyapatite is used in medicine, as a replacement of parts of the lost bone (in traumatology and orthopedics, in bone surgery), and as a coating of implants, capable of creating new bone. In dentistry, hydroxyapatite is used in kinds of toothpaste as an element that remineralizes and strengthens tooth enamel.

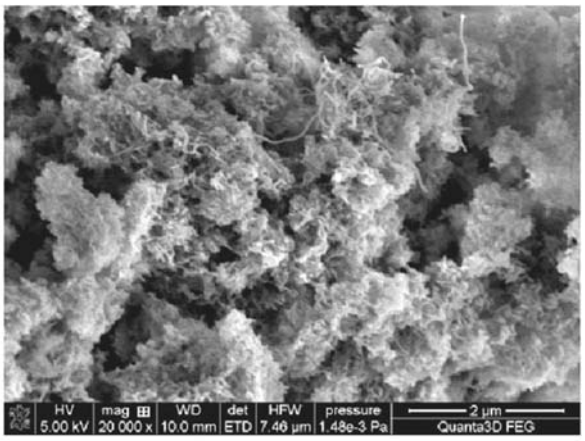


Fig.4. The appearance of crystalline hydroxyapatite [23]

Biogenic HA is a complete analog of the mineral particles of human bone tissue, it contains chemical elements in the ionic forms in which they are in a human body. HA regulates the metabolism of calcium-phosphorus in the body. When introduced into the body HA does not cause rejection reactions, activates osteosynthesis (formation of new bone tissue), increases the activity of osteoblasts, and stimulates the processes of reparative osteogenesis in other places. HA dissolves in bone, while it supplies the building materials (Ca, Mg, P) necessary for the regeneration of damaged bone tissue. HA prevents the development of an inflammatory reaction in the bone wound. After filling the bone cavities, HA will be replaced by new bone tissue. HA refers to low-toxic substances, it does not give deferred side effects.

Hydroxyapatite can be produced in dense and porous forms. Dense HA are manufactured using the technology of cold or hot powder pressing. The final product has a porosity of less than 5% with pores of less than 1 μm. Such material is used as filler for

defective bone in dental or craniofacial surgery. The porous form of HA is currently of growing interest as a potential material for bone regeneration. The main factors for the efficiency of porous HA are the size and distribution of pores. There is macro-porosity (pore diameter greater than 100 μm) and micro-porosity (pore size less than 5 μm). Micro-pores provide mechanical (physical) adhesion of tissues and implants, which increases spatial implant stability. The porous structure plays the role of a skeleton, between the elements of which a new tissue accumulates, and this skeleton gradually dissolves into the bone.

The pore size should not be less than approximately 100 μm. Pores smaller than 10 μm delay bone ingrowth; at 10-75 μm fibro-vascular growth begins to be stimulated; pores of 75-100 μm stimulate the growth of the non-mineral component of the bone; pores greater than 150 μm provide growth of the mineralized bone fraction. The pairing of pores with each other is an important process. At maximum coupling, the most favorable development of regeneration is ensured since the number of "blind zones" where the process can regress is reduced.

The main disadvantage of HA is the low mechanical properties, which do not allow the designing the large implants that can withstand a required load. To avoid this drawback, composite materials such as titanium-calcium phosphate coating are being developed. In this frame, more and more attention is paid to developing the modified HA in the polymer matrix, where the polymer matrix plays an important role in the process of retaining the structure of hydroxylapatite at high stress and dynamic impacts. HA is modified by various additives to enhance its properties. Such an additive is silicon, which promotes bone growth since silicon is involved in the formation of cartilage.

Another area of use of HA is to apply it as a coating on the surface of metal prostheses and implants made of titanium and its alloys or corrosion-resistant steels. Such a coating has to provide the ingrowth of cells and blood vessels into the structure of the implant with subsequent bone formation. The deposition of calcium-phosphate coatings on metallic implants greatly accelerates the process of engraftment and reduces the risk of implant rejection. Due to this HA is not only in composition but also in morphology coincide with natural bone tissue.

Tricalcium phosphate (TCP) is another widely used ceramic material which has a high rate of degradation and the ability to form a strong bond with the bone [21]. TCP ceramics have better biodegradability than other biomaterials, including HA. The hollow TCP frame is obtained by 3D-printing. The main mechanism of TCP bioactivity is the partial dissolution and release of Ca ions and phosphate ions thus forming a biological apatite precipitate on the surface of the bioceramic frame. The bending strength and toughness of the bioceramic β-TCP framework are higher than that of the bioceramic HA but still lower than that of human cortical bones. Therefore, bioceramic β-TCP frames cannot be used for bearing implants. In addition, the degradation rate of TCP exceeds the growth rate of new bone tissue.

Biphasic calcium phosphate ceramics (BCP) have high biocompatibility, bioactivity, and osteoconductivity and are combined with HA and β-TCP to modulate the mechanical and biological characteristics of HA for bone regeneration [18, 21]. Compared to pure HA and pure β-TCP, BCP ceramic exhibits a controlled degradation rate, better biocompatibility, and enhanced bone regeneration capacity. The best results are

obtained with high porosity of ceramics (50%). Due to its low mechanical properties, BCP ceramics also cannot be used for bone tissue regeneration.

Calcium silicate-based ceramics $CaSiO_3$ (CS)is anew bioactive material for bone regeneration that can induce the bone formation of a layer of HA on the surface by dissolving it in body fluid. The 3D-printing technique is used to obtain highly homogeneous CS structures with a controlled pore structure. The results showed that CS scaffolds have excellent compressive strength, satisfactory apatito-mineralization ability, and a high level of bone defect healing, i.e. 3D-printed CS scaffolds have significant potential for bone regeneration.

Tricalcium silicate (Ca_3SiO_5, C_3S) is a type of CS, being the bioactive silicate ceramics. This silicate ceramic has hydraulic properties and spontaneous consolidation in the aqueous medium. The 3D-printed frames of C_3S have been successfully tested for slow drug release. C_3S ceramic scaffolds with nano-needle surfaces have an increased ability to regenerate bone tissue. The major disadvantage of Ca_3SiO_5 ceramics is their high dissolution and degradation rates, leading to high pH in the body's environment, which may impair cell growth.

To enhance the properties of bioceramics, such elements as Zn, Sr, Mn, and Mg are added [18]. The dopping HA ceramics by ***Zn*** enhances HA dissolution, promotes osteoblast differentiation from bone marrow, which enhances bone formation by promoting osteoblast mineralization and inhibiting osteoclast differentiation. The bioactive ceramics with ***strontium*** $Sr_5(PO_4)_2SiO_4$ showed an increased ability to activate bone repair processes. ***Manganese*** is an important trace element for metabolism in the human body and has several advantages as an alloying element for tissue engineering. It stimulates cell adhesion, increases the bioactivity of cartilage matrix protein, and increases osteogenic activity. Ceramic TCP scaffolds doped with manganese have demonstrated their effectiveness for the regeneration of both cartilage and subchondral bone tissues due to the synergistic effect of Mn ions and Ca ions released from the scaffolds. ***Magnesium*** is a vital element in bone repair because it affects the formation of new blood vessels in the inner regions of implanted scaffolds. The corresponding ceramics $Ca_7MgSi_4O_{16}$ (BRT-H), made by the 3D-printing, was developed.

3.3 Bioactive glass and glass ceramics

Bioactive glass (bioglass, BG)is a biologically active material based on silicate glass consisting of a vitreous matrix and microcrystals used for bone repair [24]. Bioactive glasses are ceramics that can interact with body tissues. After integration, in contact with saliva or any physiological fluid, the bioglass promotes the formation of carbonized hydroxyapatite. Bioglass was invented by Larry L. Hench and it was first described in 1971. Clinical trials of BG were successfully passed in 1980-1995, and in 1995-2005 it has already been actively used in practice. Since 2005, BGs have entered the stage of innovation with the expansion of their properties and applications. Studies have shown that the dissolution products that result from the degradation of bioglass can stimulate not only the process of osteogenesis but also the chondrogenesis that leads to cartilage

formation. The process of new bone formation takes 7 days, and after 4 weeks the bone is completely restored.

At the initial stage, the outer layer of bioactive glass dissolves in the physiological fluid, attracting to its surface from the environment ions Ca^{2+}, PO_4^{-}, CO_3^{2+}, which contributes to the formation of osteoinductive layer of hydroxyapatite (calcium phosphate) and differentiation of precursors required for bone formation (Fig. 5). Subsequently, the inner layer of HA slowly resolves under the action of osteoclasts and physiological fluid, releasing chemical elements that promote bone growth.

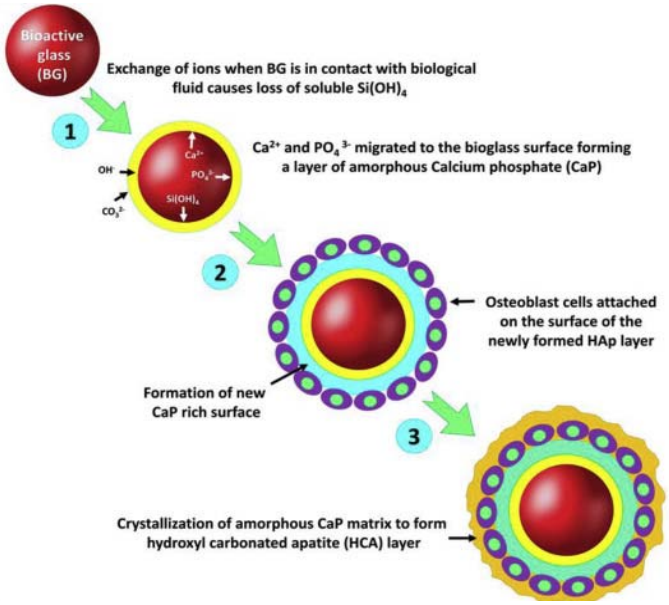


Fig.5. Illustration of the mechanism of bioglass action [25].

Bioactive glass is mainly made from silicon dioxide (SiO_2) with the addition of other oxides. When the proportions of the glass-forming substance and the alkaline components are changed, then the properties of the bioglass change from maximum bioactivity to bioinertial as follows from Fig. 6. The glass composition falling to the area S is osteoproductive, binds to both soft tissues and bone; the hydroxyapatite layer is formed in a few hours (area E falls to Bioglass compositions). The glass of area A is osteoconductive and but it does not bind to soft tissues; the formation of a layer of hydroxyapatite takes from one to several days. Area B: if SiO_2 is more than 65% then the glass is not bioactive, almost inert, encapsulated in fibrous tissue. Area C: glass is non-bonding (reactivity too high). Area D: if SiO_2 is less than 35% then the glass is not formed.

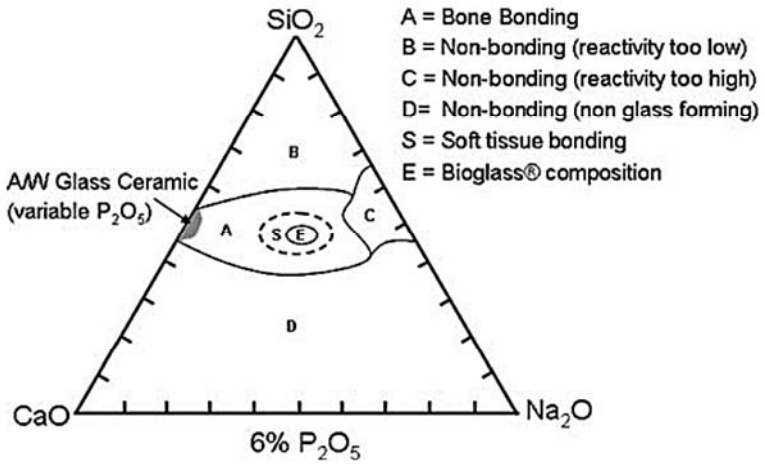


Fig. 6. Property areas depending on BGs' chemical composition [26].

The most known BG is "Bioglass 45S5", made of silicon dioxide, sodium oxide, calcium oxide, and phosphorus pentoxide in proportions: SiO_2 (45 %), Na_2O (24.5 %), CaO (24.5 %), P_2O_5 (6%). Composition of other BG brands are: "Bioglass 58S": SiO_2 (58 %), CaO (33 %), P_2O_5 (96 %), "Bioglass 70S30C": SiO_2 (70 %), CaO (30 %), "Bioglass S53P4": SiO_2 (53 %), Na_2O (23 %), CaO (20 %), P_2O_5 (4%) (inhibits bacterial development).

The mechanical strength of BGs, including fatigue resistance and fracture toughness, are significantly (10-100 times) lower than that of natural bone tissue. This limits the possibility of using a bioactive glass structure for the reconstruction of damaged bone. Bioglass is used as the main material only in cases where the implant does not carry the significant mechanical load (examples - implantation of electrodes for the restoration of hearing in case of damage to the auditory nerve or restoration of the root of the teeth). Usually bioglass is combined with polymers and metals. According to a certain production technology, bioactive glass can be obtained in the form of the desired porous structure with required cell size and orientation. Such glasses can serve as a filler or coating in polymers that gradually decompose and are replaced by the natural tissue of the host organism. The elasticity of the obtained composite materials is close to the elastic constants of the bone.

Slow cooling of the melted glass according to special temperature schedule allows to partially crystallize the glass and obtain mixed, glass-crystalline materials – ***biositals***, which have higher mechanical properties compared to glasses. The heat treatment of bioglass reduces the content of volatile oxides of alkali metals and leads to the formation of apatite crystals in the glass matrix. The obtained glass-ceramic material has higher mechanical strength, but lower biological activity.

Bioglass is obtained in various forms: particles, granules, powder, pellets. The main methods of obtaining bioactive glass and its composites are [24]: a) powder sintering; b) sol-gel process; c) melt quenching; d) self-propagating high-temperature synthesis; e) microwave irradiation; g) laser spinning.

The powder method consists of three main stages: preparation of powder, molding the workpiece by pressing, heat treatment to increase the density and strength of the material. Hot pressing and isostatic pressing are most commonly used. During the stamping process, the atoms of the polycrystalline material are diffused and the viscous flow of amorphous glass occurs.

The sol-gel process is to convert sol to gel, which is used to produce foamy and porous bioactive glass-ceramics. During the sol-gel process, the hydrolysis of silicon oxide with the formation of a colloidal solution occurs and its subsequent polymerization in the condensation reaction with the formation of a gel takes place. The gel is dried at 120-140 °C and stabilized at 700 °C. This method allows to achieve high molecular homogeneity and purity of the product.

The microwave method is that the precursors are dissolved in deionized water and placed in an ultrasonic bath for irradiation. As a result of irradiation, the powder is obtained, which is subsequently dried and heated.

Laser spinning: A small amount of raw material is melted by a high-energy laser to produce a super-fine filament, which is then lengthened and cooled by a powerful gas stream. The advantage of technology is high process speed (nano-fibers are formed in a few microseconds). The method allows to obtain glass nanofibers with a diameter of ten to hundredths of a micron.

The bioglasses are under the constant improvement of their composition and performance. The directions in this improvement are:

(a) ***boron bio-glasses*** have boron oxide B_2O_3 as the basis. These glasses are able to create a HA like a classic bioglass and stimulate the origin of new tissues. They are characterized by a high module of elasticity. The chemical composition of boron bioglass "DermaFuse" is B_2O_3(53%), CaO (20%), K_2O (12%), Na_2O (6%), MgO (5%), P_2O_5 (4%). "DermaFuse" promotes healing of chronic wounds due to the formation of a structure that resembles a special kind of protein - fibrin, which stops platelets and promotes the formation of a blood clot. Also, this material neutralizes bacteria of various kinds (cat wand, salmonella, staphylococci),

(b) ***polyester glass*** combines quartz and polycaprolactone (biodegradable polyester with a low melting point). The physical properties of polycaprolactone are very close to those of cartilage, it has sufficient flexibility and strength. Constructions of this glass, made using a 3D-printer, enhance the growth and regeneration of cartilage cells. The biodegradable implant can support the weight of the patient and provides the ability to walk without the need for additional metal plates or other implants,

(c) ***alkali-free bioglass***. Although "Bioglass 45S5" have been clinically applied to more than 1.5 million patients, this material has some drawbacks caused by high alkalis content: (a) high rate of dissolution, which causes rapid resorption, which can adversely affect the balance of bone formation, leading to the formation of a gap between the bone and the implant; (b) poor sintering ability and early crystallization through a narrow range between the glass transition temperature (~550 °C) and the start of crystallization (~610 °C). This impedes compaction and worsens the mechanical strength of the material; (c) cytotoxic effect caused by high doses of sodium, which is leached into the implant environment. To address these shortcomings, a new series of alkali-free bio-

glass were developed based on calcium dioxide $CaMg(Si_2O_6)$, calcium fluoride phosphate $Ca_5(PO_4)_3F$, and tricalcium phosphate $3CaOP_2O_5$ in different proportions. A composition "70-Di-10FA-20TCP" ($CaMg(Si_2O_6)$ (70%), $Ca_5(PO_4)_3F$ (10%), $3CaO{\cdot}P_2O_5$ (20%)) allows making a skeleton for bone tissue of any desired size.

4 Ceramic coating deposition methods

The strength of bonding bone tissue to the implant can be increased using the coatings consisting of materials close to the bone. The most commonly used coated material is the bioactive ceramic due to the similarity of its chemical composition to the composition of bones. Applying a layer of hydroxyapatite provides better adaptability of the implant to the body and faster healing of fractures. From other hand, some implant materials need to have a biologically inert coatings on their surfaces. One of such materials is Nitinol (titanium nickelide: 46-52 at.% Ti and 48-54 at.% Ni), which is used in implant surgery having a unique memory shape effect (it is also applied for the manufacture of retainers in the treatment of spinal injuries and degenerative-dystrophic diseases, staple implants in cardiac surgery, etc.). However, the diffusion and accumulation of nickel ions in soft tissues can lead to negative effects, in particular, the development of tumors. The potentially dangerous materials should be covered by the coating to prevent their negative effect on the human body. For the protection of the implant against the corrosive-active biological environment and for a better adaptation of bone tissues to a foreign body, the optimal option is to create composite bioinert (with polymer) or bioactive (with calcium hydroxyapatite, phosphates, antibiotics, etc.) layers on the surface of metal implants. Recently, Mg-based implants, which are prone to biodegradation, have become increasingly used. However, their extremely low corrosion resistance requires the use of protective coatings to slow down the corrosion process. There are many technologies for a deposition HA and other ceramics coatings on metal substrates [2, 27].

Physical vapor deposition (PVD). There are three types of PVD: vacuum, ion spraying, and magnetron sputtering. The method is based on the evaporation of the material under a beam of high-energy ions and the subsequent deposition of vaporous material on the substrate (Fig. 7). The atomic nature of the process allows using a wide variety of substrates, but the process itself is time-consuming because mass transfer occurs by separate atoms. The coatings obtained by this method have a thickness of 3-5 μm, which is much less than the thickness of the coatings obtained by other methods. This is due to the extremely low deposition rates. The deposition rates of calcium phosphate PVD coatings are typically 0.2-0.4 μm/h. The disadvantage of the PVD method is the impossibility of coating deposition on complex-shaped parts.

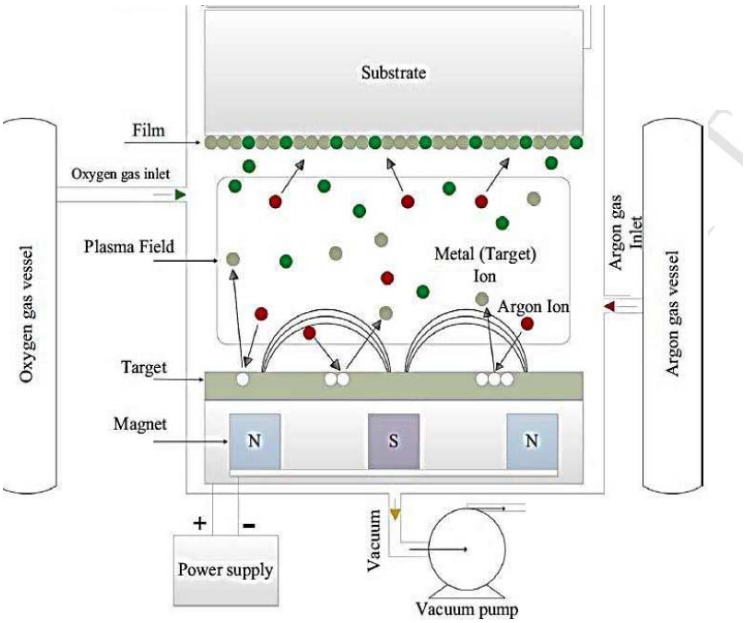


Fig.7. The scheme of PVD (magnetron sputtering) [2]

High-velocity oxidative flame (HVOF). This method is a variation of the physical vapor deposition method. The difference is that the ceramic particles are heated by the flame to relatively low temperatures and applied at higher speeds than during physical vapor deposition. These coatings have a significant volume fraction of the crystalline phase, which is associated with higher particle deposition rates and low flame energy, which leads to less powder decomposition and, as a result, to a larger volume fraction of the crystalline structure. These coatings have high porosity since the particles are in the flame for too little time and they do not enough melt to obtain a dense coating.

Chemical vapor deposition (CVD). The method allows to obtain hard, chemically resistant coatings. The starting material is gaseous or volatile chemical compounds that are either transferred by plasma to the surface of the substrate or form a coating substance by interaction with gas. Plasma consisting of gas and electron ions is formed in a constant or alternating electric field.

Electrical discharge machining (EDM). This method is basically used for machining metal products when the material is removed due to successive spark discharges between the tool and the workpiece in the dielectric medium. As a result of the discharge, a small portion of the electrode material melts and forms a plasma stream in the presence of a dielectric medium (Fig. 8). Due to the presence of powder in the dielectric medium, the plasma channel is composed of substances that arise from the decay of the dielectric medium and the erosion of the electrode. This phenomenon changes the chemical composition of the molten layer. Thus it can be used for deposition of the functional material to protect the surface of the substrate.

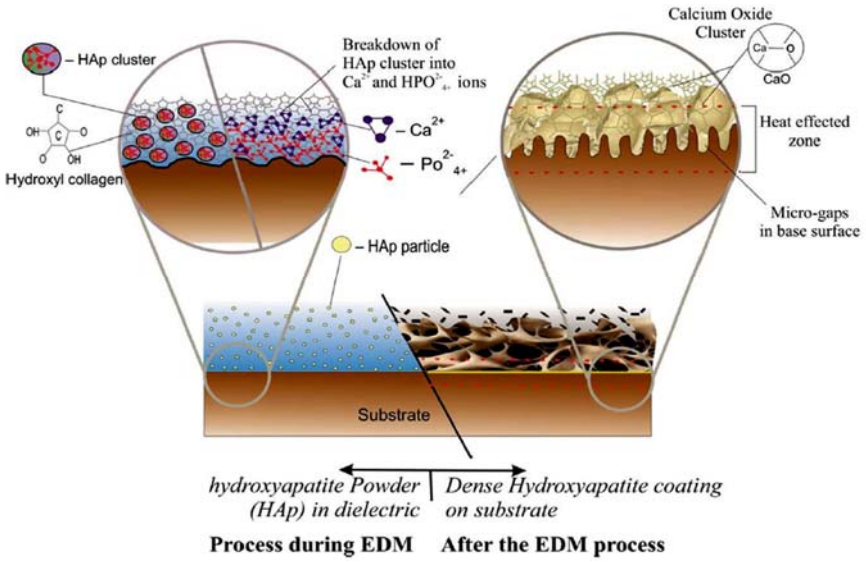


Fig.8. The scheme of EDM process involving HA powder [27].

Currentless coating methods. This method of coating (deposition from suspensions with subsequent sintering) is one of the varieties of deposition method based on sedimentation. The essence of the method is the formation of coatings by (a) deposition of the dispersed phase from the suspension on the surface of metals and alloys without using an external DC source and (b) subsequent sintering at the temperature of dispersed particles crystallization. When sintering, a connection between the coating and the substrate appears. This technique is also called a sol-gel method. It is used to produces high-crystalline calcium-phosphate coatings. The coating is formed on the surface of a dense or porous metal substrate by immersing it in a viscous slurry (with the content of the dispersed phase of 60-70%), followed by annealing at a 900-1000ºC. HA is commonly used as the dispersed phase, the medium is water, and as a binder of fine-ground calcium phosphate bioglass. The thickness of the calcium-phosphate coatings obtained by this method is 200-300 μm. The most important parameters affecting the formation of HA and the strength of the coating are the ratio $[Ca]^{2+}/[PO_4]^{3-}$ (for bone tissue the ratio is 1.67), the thickness of the coating, and the sintering temperature. The advantages of this method are: (a) the calcium-phosphate coatings have a stoichiometric phase composition; (b) there is the ability to control the thickness and porosity of coatings. The disadvantages of this method are: (a) high annealing temperature may adversely affect the strength of the metal properties and impair the adherence between the substrate and coating leading to the brittleness of the coating, (b) restrictions on the application of calcium phosphate coatings on titanium may be established since at temperatures above 625-673 K in titanium recrystallization processes begin leading to a decrease in Ti mechanical properties, (c) there is a problem of inconsistency of the

values of thermal linear expansion coefficients of metals and ceramic (coating) materials. For reliable adhesion of the coating to the matrix, the difference in this coefficient values should not exceed 15%. Otherwise, the coating is prone to peeling.

Plasma coating deposition. Plasma spraying is the process of formation of coating using a plasma flux (Fig. 9). Plasma is a partially or fully ionized gas having conductivity and high temperature. The essence of plasma spraying is that sprayed powder material is introduced into the plasma flux to be heated and partially melted. Thus the two-phase (gas and melt) stream of material is targeted to the substrate. The particles impact on the substrate making the interaction with the surface to form a coating. The plasma process consists of three main stages: (a) generation of plasma flux; (b) introduction of the sprayed material into the plasma flux, it's heating and acceleration; (c) interaction of the plasma flux and the molten particles with the substrate. The simplest variant of plasma generation is an arc electric discharge between two electrodes, which is blown by the gas in the axial or perpendicular direction. The average mass temperature of heated gas is 10000 K when working with monatomic gases (Ar) and 4000-5000 K when working with diatomic gases (N_2, H_2). Low-temperature plasma deposition allows to use different materials (metals, alloys, ceramics, plastics, and their various combinations), to deposit them in several layers, receiving coatings with special characteristics to improve the quality of coatings which become more uniform, stable, of high density, of good adhesion to the substrate surface. Phosphate-calcium coatings applied by plasma deposition have good biocompatibility with osteoblasts. The method of plasma spraying allows to flexibly vary the structure parameters and the properties of the calcium-phosphate coating over a wide range.

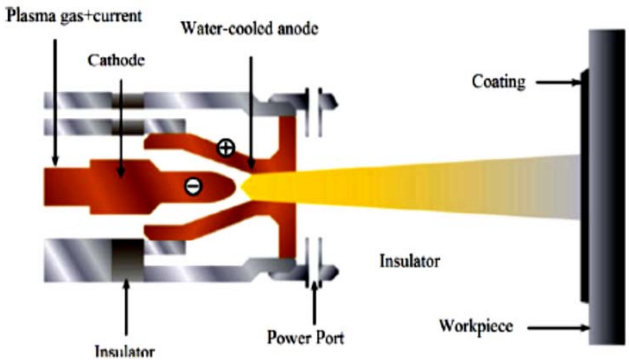


Fig.9. The scheme of coating deposition using steady plasma flux [2].

The alternate method is pulsed-plasma deposition. Plasma pulse is generated by arc discharge inside the chamber of electrothermal axial plasma accelerator (EAPA) connected with electric circuit as shown in Fig. 10 [28]. The electric circuit consists of a capacitive storage and a launch circuit. The storage capacitor (C1) is 1.5-3.0 mF, the operating voltage is up to 5 kV, the maximum accumulated energy is 19-37 kJ. The launch circuit includes elements C2, C3 and L. A high-current pulsed arc discharge is initiated between the cathode (A) and anode (B), to be limited by a narrow dielectric

channel of EAPA. The discharge duration is ~ 1.5 ms, the maximum current reaches 4 kA. In this case, the plasma flow occurs in the discharge channel due to the intense evaporation of the electrodes and of the wall substance. As a result, the pressure in the channel rises for a short time to 100-150 atm. In this case, a pulsed injection of a dense gas-plasma clot occurs through the annular anode into the substrate surface. The operating mode of the plasma accelerator is gas-dynamic. Plasma carries the materials of expandable cathode creating the coating on substrate surface. According to the estimates obtained, the temperature and density of the plasma are approximately estimated as 10-15 thousand degree (K). Reaching the substrate plasma flux results in melting of surface ensuring good bonding connections between the substrate and the coating (Fig. 11). As soon as the coating mostly consists of cathode materials, the coating composition can be easily controlled by selecting the cathode with appropriate chemical composition.

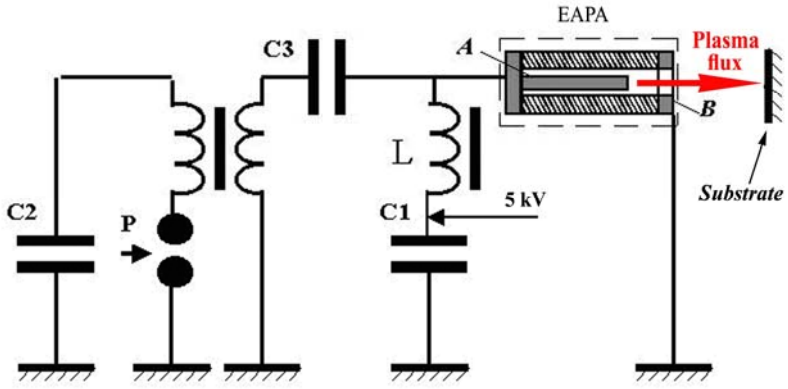


Fig.10. The scheme of electric circuit for pulsed plasma deposition [29].

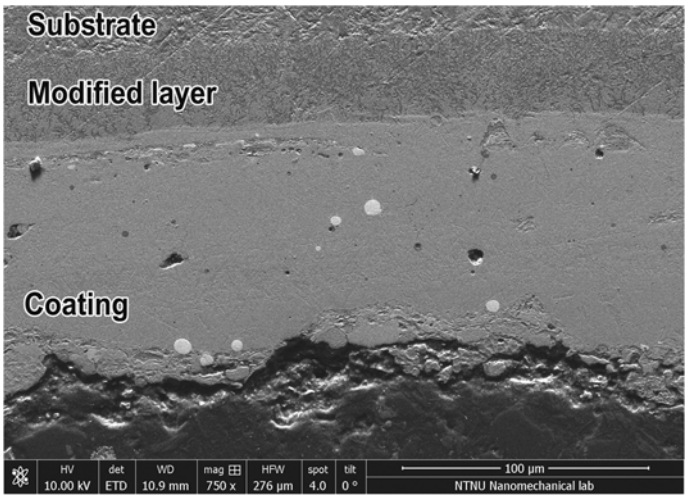


Fig.11. The microstructure of Fe-C-W coating obtained by pulsed-plasma deposition.

The main disadvantage of the plasma deposition is the low cost-effectiveness of the sputtering process because of the high loss of deposited material.

Laser deposition. The deposition of a biocompatible coating by pulsed laser ablation (PLAD) method allows obtaining a homogeneous composition and structure of the film with high adhesion strength to the substrate. Depending on the parameters of the process, it is possible to vary the surface roughness. Laser coatings can be used as intermediates between the substrate and the plasma-deposited layer. The method was successfully used to obtain films of bioactive glasses on titanium substrates. However, this method is rarely used for the formation of phosphate-calcium coatings. Laser deposition methods differ by the input the powder material into a laser beam (gravity, injection, pre-placing powder on the surface) (Fig. 12).

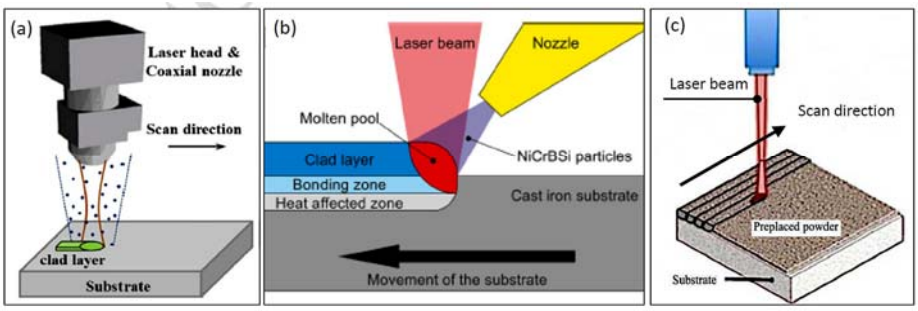


Fig.12. Scheme of laser deposition using (a) gravity, (b) powder injection, and (c) pre-powder application [2].

The coatings obtained using an Nd:YAG-laser has an amorphized structure formed by the melting and partial chemical decomposition of the material. Materials with this

structure have a high hardness. The substrate temperature is important for forming the coating structure. The deposition of the substrate at room temperature leads to greater amorphization of the coating compared to the heated substrate. As the temperature increases, the evaporation of the material and its thermal decomposition increases.

Isostatic compression. This method is a two-stage process: (a) the cold isostatic pressing of a ceramic powder on a metal substrate and (b), to obtain a dense coating, the substrate and the coating are encapsulated and pressed at 900-1000 ºC. Isostatic pressing is of limited use due to the difficulties with the sealing.

Electrophoresis method. This method is based on the deposition of electrically charged particles from solutions. It is suitable for the deposition on parts of complex shape. The low energy of the particles and the room temperature of the process lead to the formation of coatings with asymmetric porosity, which negatively affects the biological properties. Electrophoresis coatings have low adhesion to the substrate. Additional difficulties are due to the electrodeposition of undesired impurities and the heterogeneity of the coating, resulted from variable current density because of changes in the concentration of the particles.

Electrochemical deposition. According to this method, the coating forms as a result of electrochemical reactions under applied direct electric current. The formation of the coating occurs due to the oxidation of anode material, as well as due to the deposition of the electrolyte components. The coatings obtained by this method have a high complex of corrosion resistance, wear resistance, and hardness. The electrochemical method is applied for the deposition of the coating of 10-30 μm thick on objects of complex shape. The advantage of the method is its simplicity, reliability, and availability for wide applications. The disadvantage is that the coatings obtained do not meet the increased requirements for the durability and reliability of products. In particular, electrochemical deposition leads to decrease in fatigue strength.

Oxide coatings and calcium-phosphate coatings obtained by this method have the maximum adhesion in comparison with calcium-phosphate coatings obtained by other methods. Micro-arc discharge formation of coatings in aqueous electrolyte solutions as a method of depositing calcium-phosphate coatings on the titanium surface has become widespread in the last decade. The formation of the coating in the micro-arc discharge is associated with high-temperature chemical processes in the local area due to the oxidation of the substrate material, as well as due to the precipitation of ultra-dispersed phase from the electrolyte. The coatings obtained by this method have a good range of physics-chemical properties: high corrosion resistance, wear resistance, hardness. The micro-arc method allows ceramic calcium-phosphate coatings up to 10-30 μm in thickness to be deposited on complex objects. The main advantage of the micro-arc method is the formation of a chemical bond between the coating and the metal substrate. This allows to get a higher (in comparison with other methods) value of adhesive strength.

Biomimetic approach. Recently, biomimetic (means “imitation of nature”) approach has been widely used, that is obtaining coatings under conditions close to the physiological environment of the organism. The connection of the implant to the bone tissue develops through the stage of formation of a biologically active layer of carbonate-apatite on the surface of the implant. The formation of such a layer is initiated by the transition of calcium ions through the special fluid, which simulates the body

fluid in its composition. Calcium-phosphate coatings can be obtained by directly immersing the substrate in a calcium-saturated solution. The thickness of the biomimetic layer increases with time, the speed of its formation increases with the degree of saturation of the solution. The method works at physiological temperature, so the formation of apatite occurs in conditions close to those observed in the body. This method allows the formation of coatings on implants of complex geometry and gives the necessary binding of the coating to the bone tissue. Since the initial surface of the titanium implants does not have sufficient bioactivity to form a calcium phosphate coating, the substrate surface is pretreated to create a bioactive TiO_2 layer. The chemical composition, pH of the solution, temperature, and other parameters of the experiment may vary to obtain coatings of the desired composition and crystallinity. Biomimetic apatite coatings can also be formed on inert material, including polymers. Hydroxylapatite coatings are very effective for the osteointegration of metal implants with bone tissue. Titanium implants with such coatings are used in dentistry and orthopedics. The coating must have a developed system of open interconnected pores of sufficient size (preferably more than 150 μm) to provide the biological fluxes required for the osteointegration process.

Conclusions

As can be seen from the data presented in this section, biomaterials represent a large group of different solid substances of various structures and physical and mechanical properties. They are the basis for the development, manufacture and use of artificial orthopedic implants. The value of these materials cannot be overestimated given the growing attention to improving the health care and quality of life of patients. All types of biomaterials and their processing are actively developed to better meet the high standards of biocompatibility and long service life of the implants. All of this underlines the importance of biomaterials study by the students specializing in Biomedical Engineering. This section is recommended for use as a reference for studying the disciplines "Biomedical Materials", "Biotribology" and other engineering disciplines related to the design of artificial implants.

References

1. Niinomi, M., Nakai, M., Hieda, J.: Development of new metallic alloys for biomedical applications. Acta Biomaterialia 8, 3888–3903 (2012).
2. Ibrahim, M.Z., Sarhan, A.A.D., Farazila, Y., Hamdi, M.: Biomedical materials and techniques to improve the tribological, mechanical and biomedical properties of orthopedic implants – A review article. Journal of Alloys and Compounds 15,636-672 (2017).
3. Virk, J.S.: Pseudotumor as a rare complication following metal on metal total hip Arthroplasty. Journal of Postgraduate Medicine Education and Research 52(1), 26-30 (2018).
4. Bergschmidt, P., Bader, R., Finze, S., Schulze, C., Kundt, G., Mittelmeier W.: Comparative study of clinical and radiological outcomes of unconstrained bicondylar total knee endoprostheses with anti-allergic coating. The Open Orthopaedics Journal 5, 354-60 (2011).

5. PEEK Biomaterials Handbook. Editor Kurtz, S. 2nd edn.William Andrew (2019).
6. Semlitsch, M. Classic and new titanium alloys for production of artificial hip joints. Titan 2,721–740 (1986).
7. Aherwar, A., Singh, A.K, Amar Patnaik: Current and future biocompatibility aspects of biomaterials for hip prosthesis AIMS Bioengineering 3(1), 23-43 (2016).
8. Niinomi, M., Hattori, T., Morikawa, K., Kasuga, T., Suzuki, A., Fukui, H, Niwa, S. Development of low rigidity β-type titanium alloy for biomedical applications. Mater. Trans. 43, 2970–2977 (2002).
9. Sumitomo, N, Noritake, K, Hattori, T. Experiment study on fracture fixation with low rigidity titanium alloy: plate fixation of tibia fracture model in rabbit. J. Mater. Sci. Mater. Med. 19, 1581–1586 (2008).
10. Kasano, Y, Inamura, T, Kanetaka, H, Miyazaki, S, Hosoda, H. Phase constitution and mechanical properties of Ti–(Cr, Mn)–Sn biomedical alloys. Mater. Sci. Forum 654–656, 2118–2121 (2010).
11. Zhao, X.L., Niinomi, M., Nakai. M., Heida, J. Beta type Ti–Mo alloys with changeable Young's modulus for spinal fixation applications. Acta Biomaterialia 8, 1990–1997 (2012).
12. Zhao, X.L., Niinomi, M., Nakai, M. Relationship between various deformation induced products and mechanical properties in metastable Ti–30Zr–Mo alloys for biomedical applications. J. Mech. Behav. Biomed. Mater. 4, 2009–2016 (2011).
13. Thomann, U.I., Uggowitzer, P.J. Wear–corrosion behavior of biocompatible austenitic stainless steels. Wear 239, 48–58 (2000).
14. Suyalatu, Kondo, R., Tsutsumi, Y. Effects of phase constitution on magnetic susceptibility and mechanical properties of Zr-rich Zr–Mo alloys. Acta Biomaterialia 7, 4259–4266 (2011).
15. Lin, D.-J.; Hung, F.-Y.; Lee, H.-P.; Yeh, M.-L. Development of a novel degradation-controlled magnesium-based regeneration membrane for future guided bone regeneration (GBR) therapy. Metals 7, 481 (2017).
16. **Kafri**, A., **Ovadia**, **S., Goldman**, J., **Jaroslaw W., J., Drelich**, **E. A.**: The suitability of Zn–1.3%Fe Alloy as a biodegradable implant material. Metals 8(3),153 (2018).
17. Shalomeev, V., Aikin, N., Chorniy, V., Naumik, V. Design and examination of the new biosoluble casting alloy of the system Mg–Zr–Nd for osteosynthesis. Eastern-European Journal of Enterprise Technologies 1(12), 40-48 (2019).
18. Kafri, A., Ovadia, S., Goldman, J., Drelich, J., Aghion, E. The suitability of Zn–1.3%Fe alloy as a biodegradable implant material. *Metals 8*(3), 153 (2018).
19. Sabu, T., Preetha, B., Sreekala M.S.: Fundamental Biomaterials: Ceramics. Woodhead Publishing (2018).
20. Dorozhkin, S.V.: Current state of bioceramics. Journal of Ceramic Science and Technology 9 (4), 353-370 (2018).
21. Antoniac, I.: Bioceramics and biocomposites: From research to clinical practice. 1st edn. Wiley-American Ceramic Society (2019).
22. Dorozhkin, S. V. Calcium Orthophosphate-Based Bioceramics and Biocomposites. John Wiley & Sons (2016).
23. Skwarek, E. Bolbukh, Yu. Janusz, W., Tertykh, V.: Hydroxyapatite composites with multiwalled carbon nanotubes. Adsorption Science and Technology 35(5): 026361741770063 (2017).
24. Gurbinder, K.: Bioactive Glasses: Potential Biomaterials for Future Therapy. Springer (2017).
25. Oliver, J.N., Su, Y., Lu, X., Kuo, P., Du, J., Zhu, D.: Bioactive glass coatings on metallic implants for biomedical applications. Bioactive Materials 4, 261-270 (2019).

26. Fiume, E., Barberi, J., Verné, E., Baino, F.: Bioactive Glasses: From parent 45S5 composition to scaffold-assisted tissue-healing therapies, Journal of Functional Biomaterials 9(1): 24- (2018).
27. Devgan, S., Sidhu, S.S.: Evolution of surface modification trends in bone related biomaterials: A review. Materials Chemistry and Physics 233, 68–78 (2019)
28. Efremenko, V.G., Chabak, Yu.G., Lekatou, A., Karantzalis, A.E., Shimizu, K., Fedun, V.I., Azarkhov, A.Yu., Efremenko, A.V.: Pulsed plasma deposition of Fe-C-Cr-W coating on high-Cr-cast iron: Effect of layered morphology and heat treatment on the microstructure and hardness. Surface and Coatings Technology 304, 293–305 (2016).
29. Kolyada, Y.E., Fedun, V.I., Onishchenko, I.N., Kornilov, E.A.: The use of a magnetic switch for commutation of high-current pulse circuits. Instruments and Experimental Techniques 44(2), 213–214 (2001).

New biodegradable magnesium based alloy for osteosynthesis

Vadim Shalomeev[1][0000-0002-6091-837X], Nikita Aikin[1][0000-0001-9513-2804], Vadim Chorniy[2][0000-0002-0730-4184]
[1]National University "Zaporizhzhia Polytechnic", Zhukovsky str., 64,Zaporizhzhia, 69063, Ukraine
fitone14@gmail.com
[2]Zaporizhzhia State Medical University, Maiakovskyi avenue 26, Zaporizhzhia, 69035, Ukraine
gr@radiocom.net.ua

Abstract. The development of biodegradable implants for osteosynthesis is a promising direction in traumatology, which makes it possible to exclude repetitive surgeries. The most promising biodegradable materials are magnesium based alloys. They are biocompatible, their biodegradation products do not cause a toxic effect, but their application in osteosynthesis is limited, mainly due to insufficient physical and mechanical properties as well as high biodegradation rate. Using the Design of Experiments the new cast biodegradable magnesium alloy was developed. The chemical composition of the new alloy was: Zr - 1.2 ... 1.3 %, Nd - 3.1 ... 3.2 %, Zn – up to 0.7 %, the rest - Mg. The alloy prepared by sand clay casting had high mechanical properties (UTS = 266 - 274 MPa; TEL = 4.3 - 5.1 %) and adjusted biodegradation rate (UTS = 188 MPa; TEL = 3.2 % after 3 months exposure to gelofusine). Pre-clinical tests on animals were used to determine the toxicity of the new magnesium based alloy as well as the reaction of animal organism to the biodegradation process. The tests showed no significant negative effects of biodegradation products neither on regenerative processes of bone tissue, nor on overall well-being of testes animals.

Keywords: osteosynthesis; biodegradation; alloying elements; design of experiments; ultimate tensile strength; total elongation; chemical composition; pre-clinical tests

1 Introduction

A large number of surgical operations involving the use of metal implants are performed around the world. Implants made of titanium or steel are traditionally used in medicine. Such long-lasting implants are, in fact, foreign bodies that carry a high risk of local inflammation. In addition, they slow the body regeneration processes and complicate further treatment. Implant removal operations are performed to avoid such negative consequences. These operations are quite expensive and do not exclude the necessity of another surgical intervention. Magnesium based alloys are promising biodegradable materials for the production of biodegradable implants. Implants made from magnesium alloys have a number of advantages over bioinert metal alloys, polymers and bioceramics. Such implants are non-toxic, non-carcinogenic, with similar mechanical properties to those of the cortical bone. Their use excludes repeated surgical

intervention, which is of great social and economic importance. However, the mechanical properties of magnesium alloys are insufficient for the production of complex metal structures (screws, plates, rods). Therefore, an important task is to improve the mechanical properties of magnesium alloys.

2 Literature review and problem statement

Various metals and alloys, such as stainless steels, titanium and its alloys, cobalt-based alloys, zirconium and tantalum, are the most widely used materials for the manufacture of implants [1]. Economically available among these materials are stainless steels, titanium and titanium alloys, while the high cost and scarcity of pure tantalum and zirconium limit their application. The use of metallic materials often causes the "stress-shielding" effect induced by their high mechanical properties that far exceed human bone tissue properties (Table 1) [2 - 5]. In addition, although metallic materials generally have good corrosion resistance, stainless steels, cobalt alloys and partially titanium alloys have disadvantages in terms of biocompatibility caused by the presence of highly toxic alloying elements (chromium, molybdenum and nickel). Accumulation of these toxic elements in living body may lead to metallosis [4, 5].

Table 1. Properties of various metals and alloys used in the manufacture of implants [2 - 5]

Material	Ultimate tensile strength, MPa	Total elongation, %	Modulus of elasticity, GPa
Stainless steels	530 – 1000	20 – 45	200
Pure titanium	200 – 550	15 – 24	102
Titanium alloys	750 – 1.100	10 – 20	105 – 115
Cobalt alloys	600 – 1.793	8 – 50	235 – 240
Pure zirconium	330	32	94.5
Pure tantalum	285 – 650	5 – 30	186
Human bone tissue	30 – 150	1.4 – 3.1	3 – 20

Excellent biocompatibility and biodegradability of certain polymers makes them good alternative to metals. They also don’t induce the stress-shielding" effect due to their much lower mechanical properties compare to metallic materials [3].

Good examples are, approved by the US Food and Drug Administration, Polyglycolic acid and polylactic acids as well as their copolymers [6]. The difference between these polymers lies in their biodegradation rates.

Polyglycolic acid (PGA) biodegrades rapidly and loses most of its mechanical strength after about 4 – 7 weeks after implantation *in vivo* [7]. On the opposite, Polylactic acid, or polylactide (PLA) and its varieties Poly-L-lactide (PLLA) and Poly-D-lactide (PDLA) have much slower biodegradation rate. For example, biodegradation of the PLLA does even not occur within the first 2 years after implantation *in vitro.*

After that, the entire period of biodegradation takes up to 5 years [8]. The cause of this are high crystallinity and hydrophobicity of the polymer [9, 10]. PDLA has lesser than that of PLLA strength and resistance to hydrolysis due to its lower degree of crystallinity thus it has faster biodegradation rate [8, 10].

Copolymers of L-lactide and glycolide (PLGA) due to the amorphous structure do not release crystalline particles with biodegradation. The rate of PLGA degradation lies between PGA and PLLA. While in body tissues, PLGA implants biodegrade in about 18 months with few and weak inflammatory reactions [11, 12]. Synthetic origin of PLGA serves as a disadvantage leading to sterile sepsis due to it lowering the pH at the site of implantation [13].

Polyhydroxyalkanoates (PHA) are of natural origin and have high biodegradability and excellent biocompatibility. PHA polymers have lower molecular weight, which makes them more sensitive to biodegradation [13]. When using PHA, no adverse biological reactions such as sterile sepsis were observed. The biodegradation rate of PHA polymers is lower than almost all other polymers except PLLA [13, 14].

Despite the advantages of biodegradable polymers, their low and unstable mechanical strength makes hard finding them application as load bearing implants (Table 2) [3, 15, 16].

Table 2. Properties of various polymers used for the manufacture of implants [3, 15, 16]

Material	Ultimate tensile strength, MPa	Total elongation, %	Modulus of elasticity, GPa
PLA	49.6	2.4	3,6
PDLLA	54	16	1.5
PLLA	28 – 57	23	1.2 – 2.7
PGA	55 – 70	1.0	6.5 – 6.9
PLGA	41.4 – 55.2	-	1.4 – 2.8
PHA	20 – 43	5 – 20	2.5

Another widely used group of materials in osteosynthesis are calcium phosphate-based ceramics such as natural hydroxyapatite $Ca_5(PO_4)_3OH$ and artificial tricalcium phosphate $Ca_3(PO_4)_2$ [17]. The ratio of Ca/P trace elements in such materials resembles the mineral phase of bone tissue (Ca/P = 1,666 for $Ca_5(PO_4)_3OH$ and Ca/P = 1.5 for $Ca_3(PO_4)_2$), and a certain chemical composition of the surface facilitates the adsorption of proteins, increasing osteoinductive properties [17, 18]. These types of ceramics are similar in mechanical properties.

The use of hydroxyapatite (HAP), although it has excellent biocompatibility, is limited due to its high fragility, which does not allow its use in the treatment of large bones or those that carry loads. Many studies indicate that to achieve maximum osteoinductivity, HAP implants must be porous [18]. The presence of pores in implants reduces their mechanical properties, which is a significant disadvantage of hydroxyapatite as a material for osteosynthesis [20 - 22]. HAP is practically insoluble in a neutral medium and slowly degrades under natural conditions, mainly through the mechanisms of cellular resorption. The biodegradation rate of HAP is so insignificant

that in the early stages of research it was believed that HAP does not have biosoluble properties [22]. The only difference of tricalcium phosphate (TCP) is its much higher biodegradation rate compare to HAP [21].

HAP is used to manufacture of polymer-based composite materials such as PLLA-HA (poly-L-Lactide based) and POC-HA (poly-1.8-octanediol citrate based) (Table 3), in order to increase their biocompatibility, osteoconductivity and mechanical properties [20].

Table 3. Mechanical and physical properties of polymer-based composites (POC-HA and PLA-HA) [20]

Characteristic	Value	
	POC-HA	PLA-HA
Ultimate tensile strength, MPa	21.4 – 334.8	35 – 154.1
Modulus of elasticity, GPa	0.023 – 0.027	1.1 – 2.6

Magnesium is a very promising biodegradable material. The density of magnesium is 1.74 g/cm^3, which is 1.6 and 4.5 times less than the density of aluminum and steel respectively. The low weight of magnesium is a great advantage over other metals, as light-weight implants cause less discomfort when used [23]. The strength and crack resistance of magnesium is greater than that of ceramic biomaterials such as hydroxyapatite, while its density, modulus of elasticity and tensile strength are closer to the properties of natural bone than in the case of other common metal implants (Table 4) [23 – 24]. Additionally, magnesium and its corrosion products have excellent biocompatibility [25].

Table 4. Comparison of mechanical properties of magnesium and human bone tissue [23, 24]

Material	Ultimate tensile strength, MPa	Total elongation, %	Modulus of elasticity, GPa	Density, g/cm^3
Pure magnesium	113	2 – 3	41 – 45	1.74
Bone tissue	30 – 150	1.4 – 3.1	3 – 20	1.8 – 2.1

Despite the advantages of magnesium, there are problems that complicate its widespread use in osteosynthesis:

- low complex of mechanical properties and high fragility complicate the use of load-bearing magnesium implants [26];
- as a result of low corrosion resistance in media containing chlorine (including human body fluid and blood plasma), the loss of mechanical properties of magnesium implant is faster than bone restoration [27, 28].
- the release of hydrogen bubbles upon biodegradation of magnesium that can accumulate in the gas pockets near the implant slowing bone tissue growth [27, 28].
- rapid corrosion of magnesium can lead to an increase in the pH of the internal environment of the body [28].

There is a possibility of eliminating listed disadvantages by using magnesium-based alloys, which have a higher level of mechanical properties and better corrosion resistance, which leads to a lower biodegradation rate, lowering the amount of released hydrogen bubbles and allowing the body to adjust the pH levels. However, among existing ones, there are no suitable alloys that would fully meet all of listed criteria. Thus, the aim of this study is to choose the best alloying system for a new biodegradable magnesium alloy, optimize its chemical composition to improve its mechanical properties and slow down the biodegradation rate and, lastly, to perform its pre-clinical testing.

To accomplish the aim, the following tasks have been set:

– to choose the most suitable alloying system in accordance with the established criteria and to study the effect of the selected alloying elements on the structure formation and the mechanical properties of the alloy;

– to conduct mechanical testing and use the testing results to construct mathematical models that describe the influence of the studied alloying elements on the mechanical properties of the new alloy;

– use the derived dependences to optimize chemical composition of the new alloy;

– to perform preclinical approbation of implants made from the designed biodegradable alloy.

3 Materials and methods of the study

Preheated charge materials were loaded to the crucible induction furnace IPM-500 with a capacity of 0.5 tons, power of 140 kW, performance of 230 kg/h and melted. After melting the alloy was poured into extraction crucibles at 650…730 °C. Then, the crucibles were installed in distributing furnace with the rated capacity of 150 kg, in which the alloy was adjusted by chemical composition and refined with VI-2 flux at 740…760 °C. Next, increasing ligature additives containing Zr, Nd, Zn were introduced into the melt, heated, and kept the melt at 730 °C. After that, the samples were poured into molds. The charge of the studied alloy contained of: primary magnesium ingots Mg 90, Mg 95, Mg 96, zinc ingots Z2, Mg-Nd foundry alloy, Mg-Zr foundry alloy. To study the effect of the chemical composition on the mechanical properties of the alloy, standard samples for mechanical tests were poured into the sand-clay form. The influence of alloying elements was studied within the following limits: 0.4 - 1.5 % Zr; 2.2 - 3.4 % Nd; 0.1 - 0.7 % Zn (Table 5).

The study of the metal was performed after heat treatment in Bellevue-type thermal shaft furnace, with the power of 112 kW and performance of 95 kg/h, as well as in PAP-4M thermal furnace type, with performance of 50 kg/h. The heat treatment mod was: hardening at 540 ± 5 °C for 8 hours with cooling in air + aging at 200 ± 5 °C for16 hours with cooling in air. INSTRUN 2801 universal test machine was used to determine the ultimate tensile strength (UTS) and total elongation (TEL) of the alloy samples. The tests were conducted in accordance with acting standards before and after aging in gelofusine (artificial blood substitute) during different periods of time.

Table 5. Chemical composition of samples and their coded values in the matrix of experimental design

Sample No. (Experiment No.)	Content of Zr, % (X_1)	Content of Nd, % (X_2)	Content of Zn, % (X_3)
1	0.4 (–1)	2.2 (–1)	0.1 (–1)
2	1.5 (1)	2.2 (–1)	0.1 (–1)
3	0.4 (–1)	3.4 (1)	0.1 (–1)
4	1.5 (1)	3.4 (1)	0.1 (–1)
5	0.4 (–1)	2.2 (–1)	0.7 (1)
6	1.5 (1)	2.2 (–1)	0.7 (1)
7	0.4 (–1)	3.4 (1)	0.7 (1)
8	1.5 (1)	3.4 (1)	0.7 (1)
9	0.95 (0)	2.8 (0)	0.4 (0)
10	0.95 (0)	2.8 (0)	0.4 (0)
11	0.95 (0)	2.8 (0)	0.4 (0)

The macro- and microstructure of the studied alloys were studied using the light microscopy («Neophot 32», «OLYMPUS IX 70») with the magnification up to 350 times. Metallographic samples for microstructure analysis were studied after heat treatment. The etching agent consisted of 1% nitric acid, 20% acetic acid, 19% distilled water, 60% ethylene glycol.

Mathematical processing of mechanical tests results was carried out in accordance with the standard method of experiment design [29] using STATISTICA - the program for the analysis and visualization of scientific and statistical data.

4 Results of the study

4. 1 Selecting the alloying system and analysis of the influence of alloying elements on mechanical properties of the new alloy

The main criteria for choosing the material for implants are:
- optimal combination of mechanical strength and ductility;
- high biocompatibility, the absence of toxic effects on the body;
- biodegradability;
- the ability of the material to keep the appropriate level of mechanical properties during long-term exposure to biocorrosion (average duration of fracture consolidation - 3 months);
- economic affordability.

As it was shown before, magnesium has the biodegradability and biocompatibility, thus, the ability of a magnesium alloy to meet listed criteria depends on the nature of its alloying elements and their ability to form complex alloyed solid solutions and intermetallic compounds in the result of heat treatment. The latter is determined by the proximity of their atomic radii, which, in accordance with the Hume-Rothery rule, should differ by not more than 15 % [30]. The larger ratio leads to a decrease of the binding energy of solvent and alloying atoms, high crystal lattice distortion and a decrease in solubility of alloying elements. Also, according to [31], the important

condition for the solubility of an element in the base metal is the difference in the electronegativities of elements that should not exceed 0.2...0.4.

Among the existing alloying systems, only Mg–Nd–Zr fully meets the listed requirements. ML10, NZ30K and WE-43 are the most usable alloys of the given system (Table 6) [32–34].

Despite the fact that Mg-Zr-Nd alloys have the lowest biodegradation rate among all of the alloying systems [34], it is still too fast to use it in long-term treatments, as the mechanical properties of these alloys are reduced below the bone tissue level. Exposure of the samples to gelofusine for 3 months (average time of fracture consolidation) leads the loss of half the implant's strength (Table 7).

Table 6. Physical-mechanical properties of alloys of the system Mg–Nd–Zr [32–34]

Alloy	Ultimate tensile strength, MPa	Total elongation, %	Modulus of elasticity, GPa
ML10	235	>3	44
WE-43	250	>2	44
NZ30K	230	>3	44
Bone tissue	30 – 150	1.4 – 3.1	3 – 20

Table 7. Mechanical properties of the Mg–Nd–Zr (ML10) alloy after exposure to gelofusine

Basic		1 month		2 months		3 months	
UTS, MPa	TEL, %	UTS, MPa	TEL, %	UTS, MPa	TEL, %	UTS, MPa	TEL, %
235	3.0	178	2.6	146	2.3	115	1.2

According to the analysis of Mg-Nd and Mg-Zr phase diagrams, increasing the Nd content in the alloy leads to alloying of the solid solution, as well as to the formation of larger amount of β"-phase (Mg_3Nd) particles after heat treatment, which should have a positive effect on strength (UTS) [35]. The increase in the amount of Zr in the alloy should have a positive effect on the ductility due to the increase in the number of α-Zr particles, which serve as nucleation centers, providing grain grinding and increasing elongation (TEL) [36]. Increasing the amount of Zn leads to additional alloying of the solid solution. Z also participates in the formation of the hardening phase, but its content in small quantities does not lead to visible changes in the microstructure [37]. We studied both separate and joint effect of changing the content of all three elements on the mechanical properties of the alloy (UTS and TEL). The mechanical tests results are given in Table 8.

Table 8. Mechanical tests results

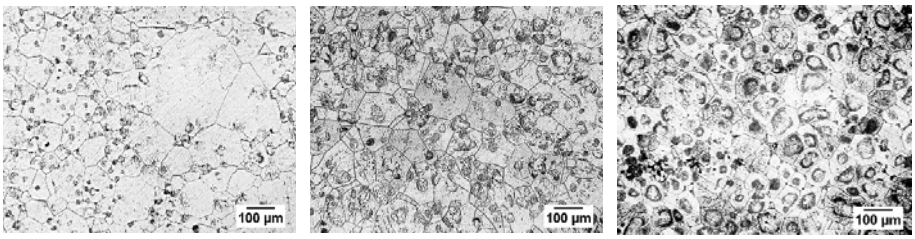
Sample No.	Content of Zr, %	Content of Nd, %	Content of Zn, %	UTS, MPa	TEL, %
1	0.4	2.2	0.1	230	2.6
2	1.5	2.2	0.1	236	5.4
3	0.4	3.4	0.1	298	2.7
4	1.5	3.4	0.1	258	3.9
5	0.4	2.2	0.7	232	2.8
6	1.5	2.2	0.7	237	5.5
7	0.4	3.4	0.7	300	2.9
8	1.5	3.4	0.7	260	4.1
9	0.95	2.8	0.4	242	3.3
10	0.95	2.8	0.4	232	3.1
11	0.95	2.8	0.4	236	2.9

The test results show that neodymium had a positive effect on the tensile strength of the alloy (samples 1 and 3; 5 and 7). Zirconium, in turn, provided an increase in relative elongation (samples 1 and 2; 5 and 6). Zinc had a minimal but positive effect on both properties (samples 1 and 5; 2 and 6; 3 and 7; 4 and 8). The combined effect of neodymium and zirconium was not additive, so at high content of both elements the mechanical properties of the alloy decreased (samples 2, 3 and 4; 6, 7 and 8).

4. 2 Analysis of microstructures

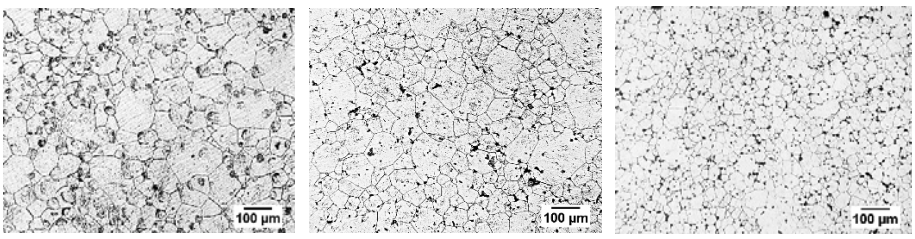
The microstructure of the alloy with the minimal amount of alloying elements of neodymium and zirconium (Fig. 1, a; Fig. 2, a) consisted of equiaxed grains of solid solution of neodymium, zirconium and zinc in magnesium, round clusters of zirconide particles, undissolved during quenching $(Mg, Zn)_{12}Nd$ phase particles (eutectoid) and dispersed β"-phase (Mg_3Nd) phase particles, which were concentrated mainly around undissolved particles of eutectoid and zirconides, forming round clusters. The increase in the neodymium content led to an increase in the number of both undissolved and dispersed particles of the hardening phase, which led to an increase in the size of the clusters (Fig. 1, b, c). With an excessive amount of neodymium, the number of secondary phases continued to increase, and at the grain boundaries, especially in the joints, there are noticeable residues of eutectics $(Mg, Zn)_{12}Nd$ (Fig. 1, c) undissolved during hardening, which indicates its excessive release.

The microstructure of the samples with a high content of zirconium (Fig. 2) had a smaller grain and an increased amount of Zn_2Zr_3 zirconides, which formed more dispersed clusters. The effect of grain grinding increased with increasing zirconium content.



a *b* *c*

Fig. 1. The microstructure of magnesium alloy samples with different content of Nd: *a* – 2.2 % Nd; *b* – 2.8 % Nd; *c* – 3.4 % Nd



a *b* *c*

Fig. 2. The microstructure of magnesium alloy samples with different content of Zr: *a* – 0.4 % Zr; *b* – 0.95 % Zr; *c* – 1.5 % Zr

Thus, the results of metallographic analysis of samples with different content of alloying elements are correlated with the theoretical analysis of phase diagrams. An increase in the amount of neodymium led to an increase in the amount of hardening phase, and an increase in zirconium - to the grain grinding.

4. 3 Mathematical processing of experimental data

To study the influence of neodymium, zirconium and zinc on tensile strength (UTS) and relative elongation (TEL), we compiled an experiment design matrix (Table 9). The influence of zirconium was encoded by number X_1, neodymium – X_2, and zinc – X_3, respectively. The joint effect of zirconium and neodymium is encoded by number X_{12}, zirconium and zinc – X_{13}, neodymium and zinc – X_{23}, all elements – X_{123}.

As a result of mathematical data processing, the regression equations describing the influence of chemical elements on the tensile strength (1) and the relative elongation (2) of the alloy are obtained:

$$\text{UTS} = 256.4 - 8.6x_1 + 22.6x_2 - 11.4x_1x_2. \quad (1)$$

$$\text{TEL} = 3.7 + 1x_1 - 0.3x_2 - 0.4x_1x_2. \quad (2)$$

Analysis of the obtained regression equations showed that the increase in neodymium content led to a significant strengthening and reducing the ductility of the alloy, and the increase in zirconium reduced the strength and increased the ductility.

The combined effect of neodymium and zirconium had a negative effect on both the tensile strength and elongation. The effect of zinc was negligible in both cases.

Table 9. Experiment design matrix (2^3)

Experiment No.	X_1	X_2	X_3	X_{12}	X_{13}	X_{23}	X_{123}	UTS, MPa	TEL, %
1	–1	–1	–1	1	1	1	–1	230	2.6
2	1	–1	–1	–1	–1	1	1	236	5.4
3	–1	1	–1	–1	1	–1	1	298	2.7
4	1	1	–1	1	–1	–1	–1	258	3.9
5	–1	–1	1	1	–1	–1	1	232	2.8
6	1	–1	1	–1	1	–1	–1	237	5.5
7	–1	1	1	–1	–1	1	–1	300	2.9
8	1	1	1	1	1	1	1	260	4.1
9	0	0	0	0	0	0	0	242	3.3
10	0	0	0	0	0	0	0	232	3.1
11	0	0	0	0	0	0	0	236	2.9

After decoding the regression equations, we derived the following dependences:

$$\text{UTS, MPa} = 188.0 + 16.3 * \text{Zr (\%)} + 27.0 * \text{Nd (\%)} – 8.6 * \text{Zr (\%)} * \text{Nd (\%)}; \quad (3.3)$$

$$\text{TEL, \%} = 3.0 + 1.7 * \text{Zr (\%)} – 0.3 * \text{Zr (\%)} * \text{Nd (\%)}. \quad (3.4)$$

To determine the optimal ratio of alloying elements and obtain the highest set of mechanical properties, optimization was performed using the Derringer desirability function using the obtained dependences. According to the results of the calculation, the optimal desirability criterion was 0.48, which corresponded to the following values of the content of alloying elements: Zr - 1.2 ... 1.3%, Nd - 3.1 ... 3.2% (Fig. 3). The expected values of mechanical properties are as follows: UTS = 268 MPa; TEL = 4.7%.

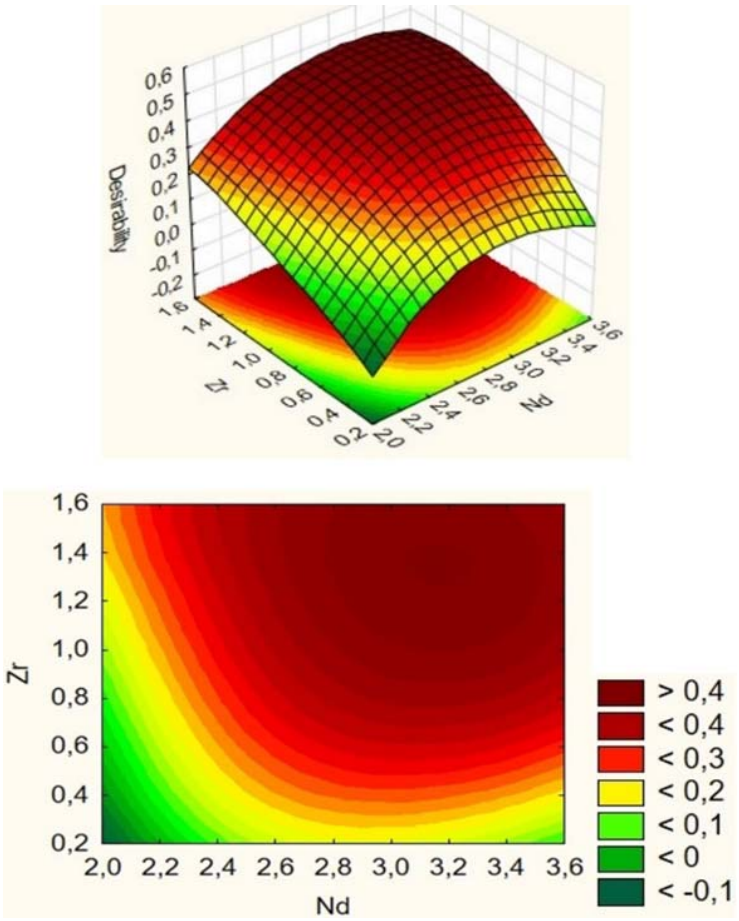


Fig. 3. The results of optimizing the chemical composition of the magnesium alloy using "STATISTICA"

4. 4 Pre-clinical tests

Implants of various designs from the developed alloy were produced for pre-clinical testing.

The developed alloy had a much smaller microstructure compared to the typical industrial alloy. The average grain size of the alloy was 131 μm (compared to 160 – 180 μm of industrial alloy). The grain boundaries were clear, no eutectic secretions along the grain boundaries were observed. There were a large number of small clusters of secondary phases that were evenly distributed over the volume of the microstructure. The its mechanical properties significantly outperformed the available Mg–Zr–Nd alloys (Table 10).

Table 10. Mechanical tests results of the alloy with an optimal content of alloying elements

Content of Zr, %	Content of Nd, %	Content of Zn, %	UTS, MPa	TEL, %
1.25	2.98	0.61	266	4.3
1.3	3.05	0.69	271	5.1
1.28	3.1	0.54	274	4.6

Exposure of the samples of developed alloy to gelofusine for 3 months showed that its mechanical properties are at appropriate level. (Table 11). Given the results obtained, the alloy is recommended for further preclinical testing.

Table 11. Mechanical properties of alloys with different chemical composition after exposure to gelofusine

Material	Basic		1 month		2 months		3 months	
	UTS, MPa	TEL, %	UTS, MPa	TEL, %	UTS, MPa	TEL, %	UTS, MPa	TEL, %
ML10	235	3,0	178	2,6	146	2,3	115	1,2
Developed alloy	270	4,7	246	4,4	220	4,0	188	3,2

Note: average values

Pre-clinical tests of the developed alloy were carried out on animals according to the “Regulations on the use of animals in biomedical research”.

The experimental animals were 20 white outbred male rats weighing 220 – 270 g divided into 2 groups: the experimental group, in the thigh muscle mass of which an implant made of developed alloy was inserted, and the control group, which did not undergo surgery. The experimental group consisted of 14 animals, the control group of 6 animals. Manipulations were performed according to the regulations of animal research. Animals of both groups were kept in standard vivarium conditions for 6 months.

To identify possible signs of intoxication, animals were regularly weighed (2 times a month), their motor and research activity was monitored, and they were constantly monitored for the pattern of water and food consumption, the urinary system, hair and mucous membranes. The conditions and feeding habits of laboratory rats met current standards.

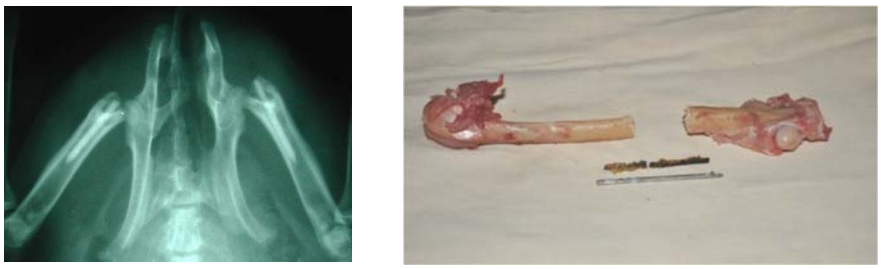
Registration of weight dynamics in the study by the method of "open field" showed a decrease in this indicator by 7% during the first 2 weeks after surgery. Subsequently, there was a significant increase in weight and a significant improvement in the appetite of animals from the experimental group, which gives reason to believe that the initial decrease is the result of surgical trauma and pain at the site of intervention, rather than toxic effects. In addition, regular monitoring of the condition of the hair and mucous membranes did not reveal abnormalities. The study of urine tests of experimental animals also showed no signs of toxic effects of biodegradation products of the implant.

A study of the orienting-exploratory behavior under the conditions of the "open field" showed that on the second day after the operation, rats with implants showed a decrease in the total index of motor activity and exploratory behavior by 60.6%. On the 14th day after the operation, the activity of animals from the experimental group slightly increased, but was significantly lower (by 35.3%) than in the intact group rats. This phenomenon can be explained by an operating trauma. In the future, the indices of the motor activity of the research and intact groups did not differ significantly. This is especially significant in periods from 2 to 6 months after surgery, when we can expect the maximum toxic effect of the biodegradation of the magnesium-based alloy.

As a result of the studies, a significant (relative to the intact group) increase in the content of all fractions of medium-weight molecules (MWM) in the plasma of white rats with implants has been revealed. Thus, the fraction which has an absorption maximum at 254 nm, increased 1.19 times in the experimental group; at 272 nm - 1.3 times, and at 280 nm - 1.27 times. This increase only indicates that the immune system of animals with implants is in a reactive state and corresponds to a slight release of biologically active substances into the bloodstream by immunocompetent cells. With endogenous intoxication, this indicator usually increases dozens of times, which was not observed in our case.

The source of extracellular nucleic acids in the blood can be necrosis or disintegration of the nucleus-intensive cellular elements of the blood. In the research and intact groups, there were no significant differences in the index of nucleic acids in the plasma of animals, and at the end of the six-month period after the operation, the content of nucleic acids in the plasma completely normalized. Therefore, there is no reason to believe that the products of biodegradation of the implant provoke cell death.

The study of the effect of magnesium alloy degradation on the process of regenerative osteogenesis after a fracture performed on 12 mature rabbits. We conducted a simulation of fractures of the upper third of both femoral bones. Intramedullary osteosynthesis with fixatives from the developed magnesium alloy was performed in the main group (Fig. 4). In the control group of animals, osteosynthesis was performed with 12X18H10T (AISI 304 analogue) stainless steel rods.



a b

Fig. 4. Osteosynthesis of a rabbit with the implant made from the developed alloy: a) radiograph; b) removed implant

The animals were taken out of the experiment in 2 weeks, 1.4 and 6 months from the time of the operation. After removing the animals from the experiment, the femur was sawn through the fracture zone along with the implant. For the collection of biological material, a slice was made along the reparation zone.

Evaluation of the biological material established that the growth of fractures of the femoral bone in animals when using implants made from magnesium alloys did not differ significantly from the processes of physiological reparation that was observed in control animals. Consolidation of the fracture occurred at the same time and without any pathological changes.

The use of magnesium implants did not disrupt the formation and development of new blood vessels in the fracture zones. In all cases, a network of microvessels was distributed, the density of which increased with the intensity and duration of the reparative process. Already in the early periods after the fracture of the formation of bone tissue, the formation of cavities of various sizes took place, on the inner surface of which endothelial lining appeared.

Using stainless steel implants lead to a disturbance in the blood supply of compact bone osteons. This caused the death of precursor cells with their subsequent resorption replacement with connective tissue fibers. Such changes were not observed when using magnesium implants

In the later periods of observation (6 months), the structure of a part of the bone in the magnesium implant area did not differ from the normal bone tissue structure. The formed bone callus was slightly different from the bone tissue only by a certain random arrangement of trabecula of bone

So, as a result of an experimental morphological study, it has been established that the use of the developed magnesium alloy for osteosynthesis of fractures did not violate the processes of reparative bone tissue regeneration.

5 Discussion of results

The results of mechanical tests of samples with different contents of alloying elements in the magnesium alloy show that the change in mechanical characteristics depends on such microstructure parameters as the average grain size, the number and distribution of particles of secondary phases. The alloying with neodymium leads to an increase in the strength characteristics of the alloy due to the release of a larger amount of dispersed particles of the β"-phase. However, excessive alloying of the alloy with neodymium leads to excessive precipitation of both secondary phase particles at the grain boundaries after heat treatment and an increased amount of non-equilibrium eutectic precipitates that do not dissolve during heat treatment and also concentrate at the grain boundaries. This results in a decrease in the ultimate strength. At the same time, an increase in the amount of zirconium leads to both grain refinement and the release of a large amount of zirconides that form clusters. These particles are the center around which the particles of the remaining phases, including the strengthening phase, gather. Excessive alloying with zirconium leads to the precipitation of a large amount of zirconides at the grain boundaries, leading to a decrease in both strength and ductility. Excessive alloying of the alloy with both elements is accompanied by a significant decrease in both properties. At this stage, the amount of zirconides, β"- phase particles and undissolved eutectic at the grain boundaries reaches a critical value, leading to softening of the alloy. Moreover, the softening effect overpowers the hardening by solid solution alloying and grain-boundary hardening.

The optimal ratio of neodymium and zirconium at the level of Zr - 1.2 ... 1.3 %, Nd - 3.1 ... 3.2 % corresponds to the state of the microstructure with the maximum number of dispersed rounded clusters consisting of particles of zirconides and β"-phase, the minimal grain size and the absence of softening effect.

Although the effect of zinc was found to be insignificant in regression analysis, its presence stabilizes the β"-phase. Also, judging by the results of mechanical tests, larger amounts of zinc always led to a slight increase in properties. Thus, the amount of zinc for a new alloy is recommended to be set at 0.7 %. Considering that a larger amount of zinc can lead to intoxication, further studies are required on its maximum possible concentration in Mg-Zr-Nd alloys without the appearance of negative effects.

The acquired regression equations have certain limitations despite a sufficient accuracy of the results obtained. First, they show the change in mechanical properties only within the predefined limits: Nd: 0 – 3.4 %, Zr: 0 – 1.5 %. Second, a procedure of mathematical modeling implies the presence of errors in calculations. Consequently, the results derived will be somewhat different from those obtained practically. Lastly, Mg–Zr and Mg–Nd phase diagrams are not thoroughly studied, especially in solid solution area, that leads to different numbers in maximal solubility depend on the sources. Thus, there is an opportunity to further refine the chemical composition, based on practical application.

The analysis of the orientation-research activity of the animals showed a positive trend for the animals of the experimental group, due to a gradual increase in the motor and orientation-research activity of the animals in the experimental group, starting from the 2nd month after the operation. For almost all criteria, no significant differences were found between the control and experimental groups. The high level of the general neurological status of animals confirms the absence of neurotoxicity of the biodegradation products of the new magnesium alloy.

The positive results of experiments conducted allow us to recommend the use of implants made from the designed biodegradable alloy of the system Mg– Zr–Nd for clinical studies.

6 Conclusions

1. Mg-Zr-Nd alloys are the most promising group for manufacturing biodegradable surgical implants. They have the highest mechanical properties along with lowest biodegradation rate.

2. The existing industrial alloys still have insufficient mechanical properties after prolonged exposure to biocorrosive environment. There is a possibility to increase the mechanical properties of Mg-Zr-Nd alloys by optimizing the chemical composition.

3. Empirical dependences that describe the influence of the chemical composition of the magnesium alloy on its mechanical properties have been established. The optimal chemical composition of the new alloy was determined: Zr - 1.2 ... 1.3%, Nd - 3.1 ... 3.2%, Zn – up to 0.7%, the rest - Mg. It provides the increased complex of mechanical properties: UTS = 266 - 274 MPa; TEL = 4.3 - 5.1 %. After exposure for 3 months in artificial blood substitutes, the alloy maintained a level of mechanical properties close to the properties of bone tissue: UTS = 188 MPa; TEL = 3.2 %, followed by biodegradation.

4. The results of preclinical tests showed that the developed alloy had no toxic or carcinogenic effects on living organisms, had a positive effect on osteogenesis and reparative processes.

Results from the experiments conducted allow us to suggest a favorable prognosis about the possibility of using implants made from the designed biodegradable alloy of the Mg–Zr–Nd system in humans.

References

1. Hayes J.S., Richards R.G.: The use of titanium and stainless steel in fracture fixation. Expert Rev. Med. Devices 6, 843 – 853 (2010).
2. ASTM F1586-13e1, "Standard Specification for Wrought Nitrogen Strengthened 21Chromium-10Nickel-3Manganese-2.5Molybdenum Stainless Steel Alloy Bar for Surgical Implants (UNS S31675)", https://www.astm.org/Standards/F1586.htm, last accessed 2020/12/15
3. Wu S., Liu X., Yeung K., et al.: Biomimetic porous scaffolds for bone tissue engineering. Materials Science and Engineering 80, 1 – 36 (2014).
4. Muzychenko P.F.: Problemy biomaterialovedenija v travmatologii i ortopedii. Travma 1, 94 – 98 (2012).
5. Seall C.K., Vince K., Hodgson M.A.: Biodegradable surgical implants based on magnesium alloys – A review of current research. Processing, Microstructure and Performance of Materials 4, 1 – 4 (2009).
6. Chen C., Shyu V.B., Chen Y., et al.: Reinforced bioresorbable implants for craniomaxillofacial osteosynthesis in pigs. British Journal of Oral and Maxillofacial Surgery 51, 948 – 952 (2013).
7. Vasenius J., Vainionpaa S., Vihtonen K., et al.: Comparison of *in vitro* hydrolysis, subcutaneous and intramedullary implantation to evaluate the strength retention of absorbable osteosynthesis implants. Biomaterials 11, 501 – 504 (1990).
8. Pihlajamaki H., Bostman O., Hirvensalo E., et al.: Absorbable pins of self-reinforced poly-L-lactic acid for fixation of fractures and osteotomies. The journal of bone and joint surgery 74-B, 853 – 857 (1992).
9. Lauer G., Pradel W., Leonhardt H., et al.: Resorbable triangular plate for osteosynthesis of fractures of the condylar neck. British Journal of Oral and Maxillofacial Surgery 48, 532 – 535 (2010).
10. Bergsma J.E., Bruijn W.C., Rozema F.R., et al.: Late degradation tissue response to poly(L-lactide) bone plates and screws. Biomoterials 16, 25 – 31 (1995).
11. Bostman O.M.: Osteoarthritis of the ankle after foreign-body reaction to absorbable pins and screws. The journal of bone and joint surgery 80-B, 333 – 338 (1998).
12. Schumann P., Lindhorst D., Wagner M., et al.: Perspectives on resorbable osteosynthesis materials in craniomaxillofacial surgery. Pathobiology 80, 211 – 217 (2013).
13. Turesin F., Gursel I., Hasirci V.: Biodegradable polyhydroxyalkanoate implants for osteomyelitis therapy: *in vitro* antibiotic release. Journal of Biomaterials Science, Polymer Edition 12, 195 – 207 (2001).
14. Chen G., Wu Q.: The application of polyhydroxyalkanoates as tissue engineering materials. Biomaterials 26, 6565 – 6578 (2005).
15. Philip S., Keshavarz T., Roy I.: Polyhydroxyalkanoates: biodegradable polymers with a range of applications. Journal of Chemical Technology and Biotechnology 82, 233 – 247 (2007).
16. Berins L.M.: SPI Plastics Engineering Handbook of the Society of the Plastics Industry. 5th ed. Springer, US (2000).

17. Barinov S.M.: Keramicheskie i kompozicionnye materialy na osnove fosfatov kal'cija dlja mediciny. Uspehi himii 79, 15 – 32 (2010).
18. Barakat N., Khil M.S., Omran A.M., et al.: Extraction of pure natural hydroxyapatite from the bovine bones bio waste by three different methods. Journal of materials processing technology 209, 3408 – 3415 (2009).
19. Heise U., Osborn J.F., Duwe F.: Hydroxyapatite ceramic as a bone substitute. International Orthopedics 14, 329 – 338 (1990).
20. Qiu H., Yang J., Kodali P., et al.: A citric acid-based hydroxyapatite composite for orthopedic implants. Biomaterials 27, 5845 – 5854 (2007).
21. Li J., Lu X.L., Zheng Y.F.: Effect of surface modified hydroxyapatite on the tensile property improvement of HA/PLA composite. Applied Surface Science 255, 494 – 497 (2008).
22. Shikinami Y., Okuno M.: Bioresorbable devices made of forged composites of hydroxyapatite (HA) particles and poly-L-lactide (PLLA): Part I. Basic characteristics. Biomaterials 20, 859 – 877 (1999).
23. Staiger M.P., Pietak A.M., Huadmai J., Dias G.: Magnesium and its alloys as orthopedic biomaterials: A review. Biomaterials 27, 1728 – 1734 (2006).
24. Ejdenzon M.A.: Magnij. Metallurgija, Moscow (1969).
25. Karpov V.G., Shahov V.P.: Sistemy vneshnej fiksacii i reguljatornye mehanizmy optimal'noj biomehaniki. CCT, Moscow (2001).
26. Song G.: Control of biodegradation of biocompatible magnesium alloys. Corrosion Science 49, 1696 – 1701 (2007).
27. Witte F., Kaese V., Haferkamp H., et al.: *In vivo* corrosion of four magnesium alloys and the associated bone response. Biomaterials 26, 3557 – 3563 (2005).
28. Song G., Song S.: A possible biodegradable Magnesium implant material. Advanced engineering materials 9, 298 – 302 (2007).
29. Samarskij A.A., Mihajlov A.P.: Matematicheskoe modelirovanie. Fizmatlit, Moscow (2001)
30. Shalomeev V., Aikin N., Chorniy V., Naumyk V.: Design and examination of the new biosoluble casting alloy of the system Mg-Zr-Nd for osteosynthesis. Eastern-European Journal of Enterprise Technologies 1(97), 40 – 49 (2019).
31. Shalomeev V., Tsivirco E., Vnukov Y., Osadchaya E., Makovskyi S.. Development of new casting Magnesium-based alloys with increased mechanical properties. Eastern-European Journal of Enterprise Technologies 1(82), 4 – 10 (2016).
32. GOST 2856-79. Splavy magnievye litejnye. Marki. Izdatel'stvo standartov, Moscow (1981).
33. Xingwei Z., Jie D., Wencai L., Wenjiang D.: Microstructure and mechanical properties of NZ30K alloy by semicontinuous direct chill and sand mould casting processes. China foundry 8, 41 – 46 (2011).
34. Turowska A., Adamiec J.: Mechanical properties of WE43 magnesium alloy joint at elevated temperature. Archives of metallurgy and materials 60, 2695 – 2701 (2015).
35. Gill L., Lorimer G.W., Lyon P.: Microstructure/property relationships of three Mg – RE – Zn – Zr alloys. In: Proceedings of the 6th international conference: Magnesium alloys and their applications, January 21, 421 – 426 (2004).
36. Koltygin A.V.: Analiz vozmozhnyh fazovyh prevrashhenij pri kristallizacii i ih vlijanie na lituju strukturu v splave ML10. Metallovedenie i termicheskaja obrabotka metallov 8, 25 – 28 (2013).
37. Choundhuri D., Srinivasan S.G., Gibson M.A., et al.: Exceptional increase in the creep life of magnesium rare-earth alloys due to localized bond stiffening. Nature communications 8, 1 – 9 (2000).

Modern trends in application of smart materials in biomedical engineering

Oleksandr Cheiliakh[1], Yan Cheylyakh[1], Natalia Mak-Mak[1], Inna Malysheva[1], Dariya Burova[1], Janusz Mikula[2]

[1]Pryazovskyi State Technical University, Universitertska Str., 7, Mariupol, Ukraine, 87555
[2]Cracow Politechnika im. T. Kostyushko, Jana Pawla II nr 37, Cracow, 31864
aleksandr.cheylyakh@gmail.com

Abstract. Among the many materials used in biomedical engineering, smart biomaterials hold a special place. They are able to adequately respond to changing external conditions and to improve their operational properties. Moreover, in such materials unique properties and operational capabilities are achieved, which cannot be achieved in conventional materials. Review analysis of modern trends in the application of the groups of smart materials of different physical nature and functional purpose in biomedical engineering is carried out, implemented in them during the operation of physical phenomena, transformations and effects, which leads to the formation of conditioned properties. The application features of smart materials with different physical natures, providing the formation of properties in the process of their operation under the action of phenomena (transformations) realized in them were researched. This chapter will be used to update the enrichment of the number of disciplines in the Biomedical Engineering, Material Science specialties etc.

Keywords: Smart Biomaterials, Biomedical Engineering, Physical Phenomena, Properties.

1. Introduction

Biomedical material science has been developing intensely lately. Among the variety of materials applied in bio-medical engineering smart bio-materials (hereinafter SBM) capable of adequate reacting to changes in outer conditions, improving themselves, repairing themselves or alter their functional characteristics directly in the process of exploitation occupy quite a prominent place [1]. It's very important there that SBM should have a biological compatibility, i.e. co-existence with a living organism for execution of a certain, specified function, not causing any inflammation of tearing away by the organism. Reaction by the organism should not contradict to safe and efficient functioning of SBM.

As examples of application of SBMs in bio-medical engineering we may mention artificial heart valves, tissues "plasters", contact lenses, artificial eye lenses, artificial human organs: middle ear bones, cochleae, joints, bone plates, inner bone rods, bone cement, skin, urination ways, inserted skull plates, shunts of intracranial liquid, tooth

fillings, tooth dentures and the like. These SBMs could be metallic and non-metallic [2].

2. Metallic SBM

Among metallic IBMs wide application in the bio-medical engineering industry have found ***alloys, possessing memory effect***, like nitinol – NiTi (an alloy, consisting of nickel and titanium). NiTi alloy was developed in 1960 by William F. Buehler, while nitinol, the first orthodontic alloy was represented by Andraeson [3]. These alloys are capable of changing their shape, acquiring the original intended one, set under the influence of heating in human's organism (up to ~37 °C) owing to the process of reverse phase (martensitic) transformation.

For alloys possessing the memory effect [2-7] the degree or a relative value of restoration of parts' shape depend sufficiently upon the chemical composition of the alloy, temperature intervals of direct (M_s and M_f critical points) and reverse (A_s and A_f points) phase transformations, the microstructure and the loading conditions in which they are realized. A typical operating cycle for nitinol, such material (~55 % Ni, and ~45 % Ti) is represented in Fig.1. Deformation is carried out through a direct martensitic transformation at b-c stage (see Fig.1) - reorientation of martensite crystals (the effect of martensitic non-elasticity) and remains after the loads have been.

The shape memory effect is revealed at heating (inside human organism) at c-d stage (see Fig.1), where the material restores independently its shape due to the reverse transformation of martensitic into austenite and can reach sufficient effects (800…1,300 MPa). The kinetics of martensitic transformations possesses a clear-cut hysteresis (Fig. 1e). Depending upon the alloy's chemical composition the interval of M_s…M_f and A_s…A_f points is largely narrow and is equal to ~30…80 °C, alloying reduces theses temperatures drastically. Beside titanium nickel alloy (Ni50Ti50), such materials also comprise the alloys of Au-Cd, Cu–Zn-Al, Ag-Cd systems, on the basis of copper Cu-Al-Ni with 15 % Ni, 36 % Al; on nickel basis with 16 % Al et al. and also polymers with shape memory [8].

Alloys with the effect of shape memory, being unique in properties have found worthy application in medicine. For instance, they are used for manufacturing filters to be inserted into blood-vessels, fasteners for fixation of weak veins, rods for correction of vertebral column at scoliosis, a great number of orthopedic implants of different designs and configurations, orthopedic arcs and brackets fixations used for correction of teeth row and great many other useful and necessary medical devices [9].

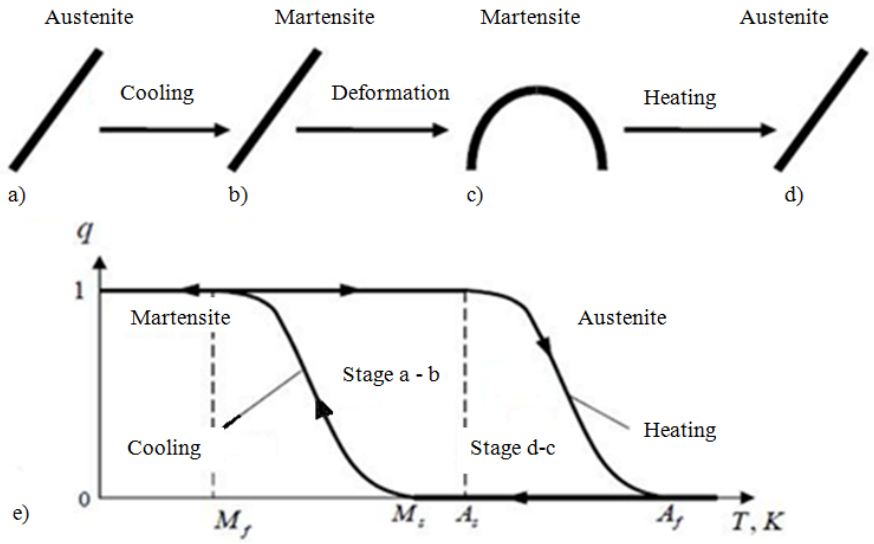


Fig. 1. The diagram of deformation of a rod with the memory shape effect (a-d) and dependence of martensitic volume share (q) upon temperature (T) (e) [7].

Application of such smart alloys is explained by simple design of devices, due to super-elasticity phenomenon and their biological compatibility with living tissues [10-12]. Osseintegration - such reaction of human organism is characterized by reliable bone cohesion with an implant, without formation of a protective tissue capsule. This discovery led to creation of glue-free fixation of artificial joints, made of Ti6Al4V alloy possessing shape memory and teeth dentures, made of pure titanium. Simon's filters (Fig. 2), an example of application of alloys with shape memory, represent a net with nitilon threads, which is inserted into a blood-vessel (see the upper pictures in Fig. 2), then it gets straitened under the influence of body temperature, acquiring the original shape (see the lower picture in Fig. 2), assigned at martensitic state.

The effect of "self-unfolding" of the filter inside the blood-vessel is caused by reverse martensitic transformation at its heating up to 36…37 °C, i.e. exceeding the transformation temperature (A_s point is below the body temperature). As a result the filter acquires the original shape (prior to its introduction into the blood-vessel) – so, it is "self-unfolded" and executes its specified function of destruction of thrombi that originate in blood-vessels. In this unfolded state the filter entraps thrombi migrating in blood current, ensuring their complete resorption.

A whole variety of implants, orthodontic wires, fixing parts, buckles and brackets for osteosynthesis, joining plates and the like is made of nitinol alloys for treatment of bone fractures and scoliosis [8].

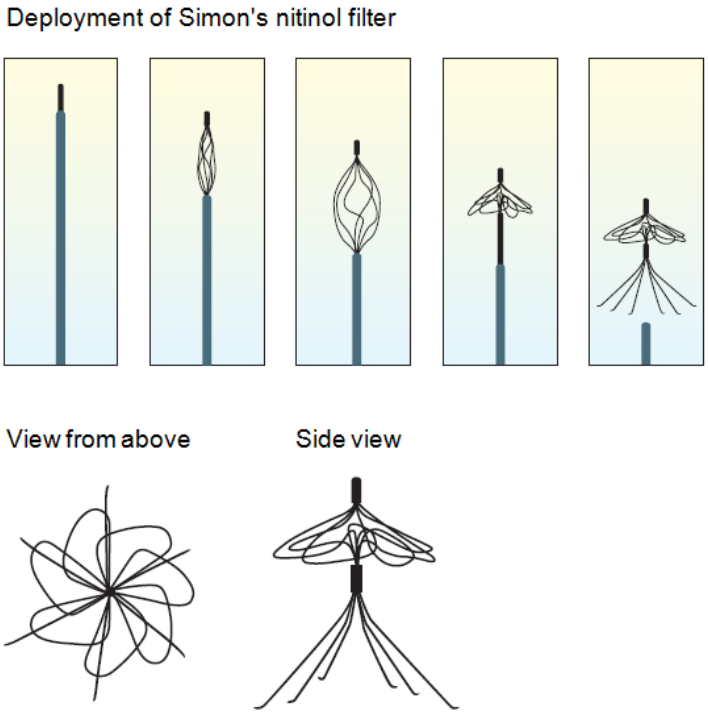


Fig. 2. Simon's nitinol filter: the upper pictures– the filer's unfolding in the catheter; the lower pictures – the filter in the unfolded state [8, 13]

Tooth bracket of orthodontic arc (see Fig 3), made of nitinol is efficiently applied for teeth leveling [3]. These alloys possess unique mechanical properties - low rigidity, super elasticity, high elasticity and a wide range of resilience [14, 15], as well as high corrosion resistance [9]. Wire is made at the temperature above TTR. As the wire is heated in human mouth above that temperature it memorizes its original shape and corresponds to it. That is why an outstanding alloy's effect is realized here-this property of wire is known as alloy with shape memory [14].

The installed arc always tries to restore its correct shape, so it "drags" the incorrectly arranged teeth. The arc's shape, density and section dimensions are changed throughout the entire treatment process. It is caused by the necessity to apply different pushing force upon the teeth row at different stages.

Super elastic NiTi wires possess excellent elasticity, as compared to other Ni Ti wires . They, too, can create constant forces at big sagging of wire [15], they will reveal the property of shape memory and generate light forces within the entire time period, ideal for orthodontic treatment [16].

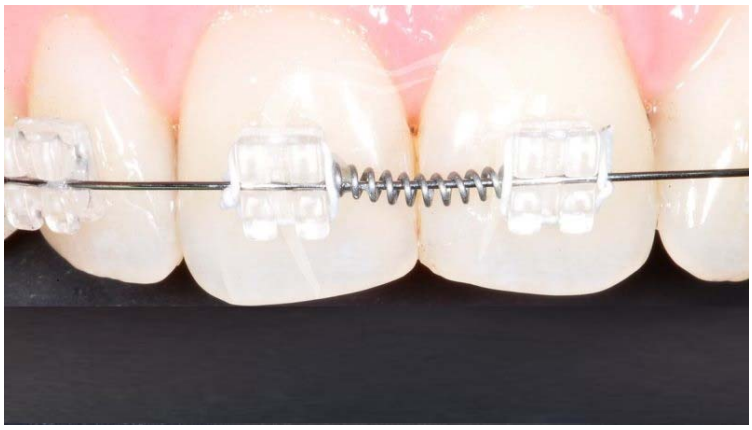


Fig 3. Orthodontic arc bracket, made of nitinol.

3. Non-metallic SBM

Non-metallic SBMs are quite diversified in their composition, designation and the field of application. As non-metallic SBM, used, for example, for teeth treatment ***color cement*** is widely used, it is also called "glass-ionomeric", it usually does not differ from the adjoining natural teeth in colour, being at the same time a source of fluoride, thus preventing further propagation of caries. Introduction of SBM into bone tissues promotes appearance of a new bone upon their surface and the effect of the area treatment without any surgical operation. Synthetic ***hydroxiapatite, bio-active glass*** and some types of ***glass-ceramics*** may be mentioned as an example of such SBM.

Some proteins are absorbed on the surface of such materials from the biologic environment that stimulate growth of bone cells, accelerating the process of healing. In some SBM (like, bio-glass, for instance) it is preceded by ionic exchange reactions on the introduction surface and calcium phosphate appears, it promoting formation of direct chemical bonding between bio-glass and the mineral phase of newly-formed bone tissue.

Bio-glass accelerates joining of the bone and the introduction [1]. If bio-glass is in liquid medium then calcium and phosphate ions are leached from the introduction, forming a surface layer, rich in calcium phosphate, that ensures good joining of bio-glass with the bone. Mutual joining is possible, by forming collagen finger-like offshoots [10]. The effect of joining of various types of glass and glass-ceramics with living bones is widely known. Some of them are used in surgery, for example, for forming of artificial middle ear bones [17], artificial cochleae, artificial joints, bone plates, inter-bone rods, inserted skull plates, bone cement, artificial skin, teeth filling and dentures [1], teeth holes (jaw-bone alveoli) [11, 12]. Bio-active glass are also used for making artificial vertebrae , iliac bones and granules for filling bone defects [10, 11, 17].

4. Ceramic SBM

Among ceramic SBM ***phosphate-calcium ceramics*** possesses outstanding introduction properties, because the bones mineral phase is similar to ***hydroxiapatite*** [$Ca_{10}(PO_4)_6(OH)_2$]. Still it possesses a less degree of crystallinity and has a whole set

of phosphate micro-inclusions like trycalciumphosphate, carbon-apatite and various ionic fluoride, magnesium, sodium admixtures, beneficial for human organism.

Fluoride-apatite [$Ca_{10}(PO_4)_6(F)_2$] differing from hydroxiapatite in its chemical structure is more stable at increased temperatures at acid impact. ***Tricalciumphosphate*** [$Ca_3(P0_4)_2$] exists in two types, called a- and /3-tricalciumphosphates. Both types possess high resistance to dissolving [18]. These materials are capable of forming a direct bonding with living bones without formation of protective fibrous tissue. Formation of a chemical bond of phosphate-calcium ceramics with bone is explained by partial solution of the near-surface layer of ceramics and formation of crystals of CO_3-apatite with included bio-molecules of surrounding liquid. Meanwhile it should be noted here that application of phosphate-calcium ceramics is limited, due to its poor strength. A new efficient way of increasing an operating mechanical load on SBM is creating a metallic substrate by plasma cladding and deposition of a layer, consisting of ceramic calcium phosphate upon it [19]. At that this metallic substrate ensures the specified strength, while the coating ensures the biological reaction of compatibility of the implant with human organism.

5. Smart coatings

A new step in modern bio-material engineering is creation of SBM of the third generation-smart coatings [1]. Formation of connecting tissue is a macro-process to be regulated by the processes of nano-scope, while the organism's reaction to the bio-material is determined by the properties of its surface [20].

One of such methods consists of chemical modification of the surface of biomaterials for absorption of some proteins, while another in inoculation of bio-active molecules on its surface. Modified surfaces can exert influence upon cell's behavior, controlling the nature of the protein layer, emerging on the surface in the biological; environment [1]. The cells are capable of "feeling" the emerged proteins with the help of receptors. As a rule, it is desirable that certain cells should be precipitated on the surface of their introduction, it is also desirable that they should grow and quickly divide. Peculiarities of the surface bio-modification are determined by the character of the required reaction of organism. If an implant undergoes an impact by bacteria, like, for instance, artificial vocal chords cells adhesion is not desirable. On the contrary adhesion is simply indispensable for orthopedic implants. By establishing the factors, determining the interaction of the sell and the introduction surface it is possible to modify deliberately the surface of SBM in order to control the bio-chemical reaction of organisms. Modified surfaces can exert influence upon cell's behavior, by controlling the nature of protein layer that is formed on the surface in the biological environment. The cells are capable of "feeling" the precipitated proteins with the help of receptors. As a rule it is desirable that certain cells should be precipitated upon the surface of its introduction, their growth and quick division are also desirable [21-23].

A space alternation of the surface chemical composition seems to be an interesting trend in developing engineering of living tissues. For example, sprayed nitrogen containing polymers stimulate precipitation of nerve cells and neurons growth. The influence of the surface chemical composition upon precipitation of certain cells can be used for controlling precipitating of required cells in certain areas [24].

Opportunities of ***biological modification of surfaces*** of SBM implants can be poly-functional, like for instance, anti-bacterial, ensuring reliable bonds with blood, or the bone.

6. Activated effects in liquid SBM

Application of the magnificent effect of ***magnet-activated (magnetorheological) liquids (MR-liquids)*** in bio-medical engineering is of great interest. Their main peculiarity is alternation of rheological properties, under the influence of magnetic field. They are used as suspensions of micro-particles of magnetic metal (iron) 3…8 μm in size, in the quantity equal to 20…40 vol. % in carrying liquid (mineral or synthetic oil, water, glycol or the like). When magnetic field is applied ferromagnetic micro-particles are rearranged in accordance with direction of the magnetic field - they are arranged in chains in direction of force lines (see Fig. 4), thus resisting fiercely the shift or the current in direction perpendicular to those chains and drastically increasing viscosity (including the yield strength) of the liquid in that direction. Accordingly, when the magnetic field is removed the particles chains are disintegrated, it leading to restoration of the original viscosity [8].

Such stimulated alternation of their properties is proportional to tensity of the magnetic field and disappears completely when the field is removed. The rate of MR-liquid response to application of magnetic field and transition into semi-solid state is exceedingly high, as well as restoration of the initial viscosity – the time of such transition is just 6.5 ms [8].

This effect of remarkable properties of MR-liquids can be used, for example, for manufacturing of "Motion Master" hydraulic shock absorber on the basis of MR-liquid or for manufacturing an artificial "Smart Magnetics" leg (see Fig. 5). The rate of reacting for this shock absorber is 20 times greater than for its mechanical or hydraulic analogues, while the response time is just 40 ms, it corresponding with response time of a real human knee joint [8]. Such "smart" artificial leg is capable of most precise imitating human movements, thus making man's life more convenient and comfortable.

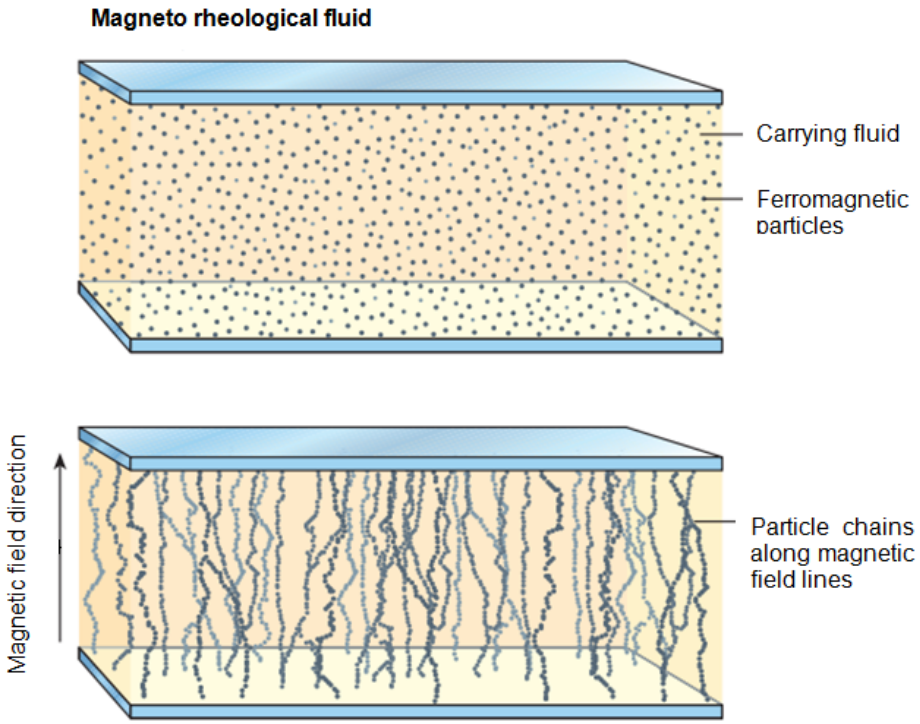


Fig. 4. The diagram of alternation in arrangement of ferromagnetic particles in MR-liquid when magnetic field is exerted [8]

The effect of collapse of polymer gels is used in polymer gels, capable of hundredfold changing (reducing) their volume [25, 26] at a slight alternation of external conditions (temperature, composition of dissolvent, light radiation, hydrogen index of the medium — pH). For example, the volume of such gel in swollen and collapsed state can be hundredfold different, some hydro-gels can withstand up to 2 kg of water per 1g of dry polymer. This happens as a result of appearance of an efficient attraction between polymeric chains of macro-molecules (hydrophobic interactions, hydrogen links) (see Fig. 6).

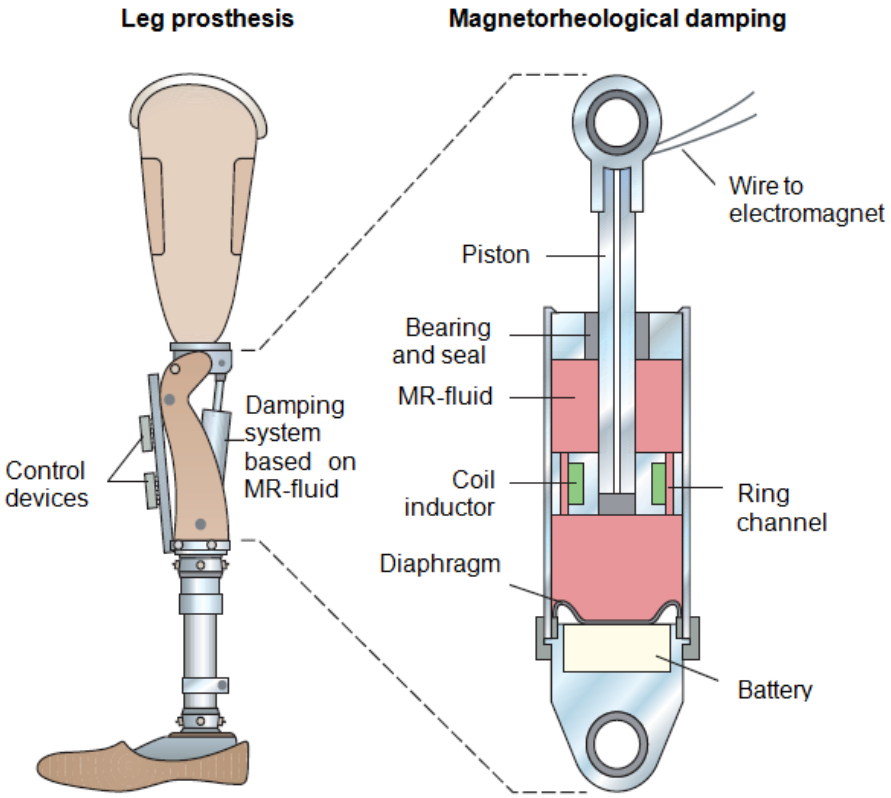


Fig. 5. An artificial knee joint of Smart Magnetix (Biedermann Motech) artificial leg equipped with hydraulic "Motion Master" shock absorber on the basis of MR-liquid (Lord Corporation) (see left picture) and the cross-sectional view of the hydraulic absorber (see the right picture) [8].

Polymeric gels being "on the verge" of collapsing are capable of exceedingly sharp and reversibly (it being very important) altering their volume in response to slight changes in the medium parameters. Due to it such gels are called responsive gels, or smart or intelligent materials, i.e. materials capable of reacting to slight environmental in previously specified way [27]. Polymeric gels are used as fillers in pampers, hygienic napkins, mild shoe inner soles taking the shape of foot etc. However, the main domain of their application is medicine. For example for manufacturing medicine capsules-the effect of "address delivery" of medicines [28], when the capsule gets swollen under the action of gastric juices and releases the medicine with pre-specified velocity.

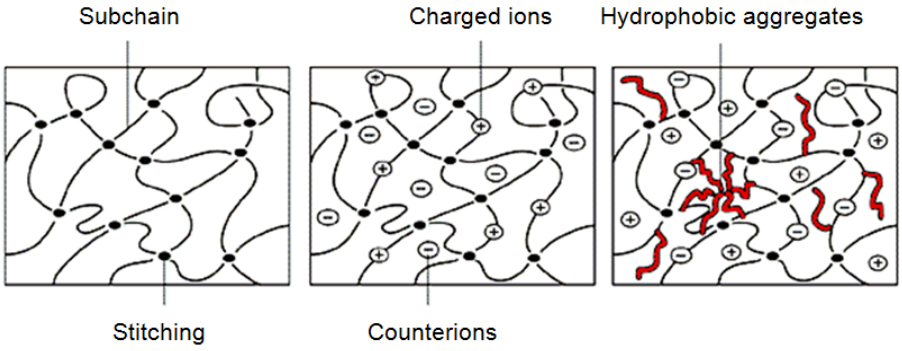


Fig. 6. The diagram of composition of three types of polymeric gels. Left to right: uncharged net, polyelectrolyte (charged links on polymeric chains and low-molecular anti-ions are formed in it by means of dissociation of ionogenic groups) and the net with hydrophobic groups, associating with each other in the liquid solution [26].

Further progress in creating self-regulating medicines requires, in particular, development of the ways of obtaining hydro-gels that could pass from collapsed state into swollen at different pH values. It will allow more fine controlling both the site and the rate of extracting the medicine, on the spot, required by the organism and the procedure of treating the ailing place. By selecting the length and the number of hydrophobic groups it is possible not only to obtain the gel with pre-specified pH values for transition from collapsed state into the swollen state, but also reach the required degree of swelling at a given acidity. By modifying the composition and the degree of gel linking it is possible to change its selectivity to molecular mass and the charge of absorbed dissolved compounds [27].

7. Bio-degrading materials

Bio-degrading materials –is a modification of SBM, capable of destroying with time, under the influence of biological factors, they either disintegrate or assimilated inside human organism or in nature, due to ferments action. Their degrading at the process of exploitation is based on passing chemical or bio-chemical reactions: hydrolyzation under the action of ferments and microorganisms or disintegration of macro-molecules into oligomers to be destroyed by bacteria. Their application is restricted to medicine and food and consumer goods packing. Bio-degrading materials are quite efficiently used in surgery for making seam threads and various cardiologic, dental, orthopedic, fixing and other implants. For instance, dissolved surgical threads made of polylactide acid, that disappears (it is dissolve din human organism) after it has carried out its fixing function.

Application of bio-degrading materials seems to be the foundation for creating the most promising and swiftly developing medical trend - creating organs from patient's tissue cells. A kind of skeleton is created from them with a spongy structure, in its meshes recipient's cells are implanted, capable of growing and thus capable of restoring the lost tissue. This gives an opportunity and conditions for this artificial formation to

be germinated with blood vessels with subsequent resorption of the bio-degrading material and its eventual replacement with own connecting restoring tissue. The described technologies are already used for restoration and replenishing of lost sections of cartilaginous and soft tissues. On the basis of matrixes, consisting of organic acids (like lactic or glycolic acids) such organs and tissues like skin, cartilages, bones, tendons, muscles, small intestines etc. are created in tissue engineering [1, 2].

8. Living tissue engineering

Living tissue engineering has become a new philosophy of developing up-to-date bio-materials – true SBM of the fourth generation, it being one of the quickly developing spheres of bio-medical engineering [1]. It lies in biological or technical methods of creating functional tissues, either replacing or improving functioning of ill or pathological parts of organism. In practice this idea is realized by growing living cells on biomaterial in the presence of biologically active molecules. After that living cells and the produced extracellular matrix together with the substrate are introduced into the organism as a uniform cell-biomaterial structure, capable of self-generating and self-restoring in the process of carrying out the vital functions.

9. Biologically modified SBM

One of the promising research trends is developing ***biologically modified SBM***, the surface of which will carry certain information for living cells that interact with this surface. The information may consist of determination where the cells should and should not be planted and specifying their orientation or differentiation [1].

Creation of truly smart substrates is but a task for the future. It requires further complex development of bio-materials engineering, biology and medicine. Development of new medical equipment, bio-materials, particularly SBM and artificial tissues is to play more important part in treating illnesses. Further development of bio-materials will be the result of joint efforts by specialists of materials engineering, biologists and doctors [2]. New SBM are likely to differ drastically from their analogues of past. They will become much smarter, in the sense of their interacting with biological environment, promoting self-restoration of physiological functions of organism and living tissues. The ultimate objective of treatment will be restoration of healthy tissue with simultaneous vanishing of the implanted bio-material [29].

10. Smart nanomaterials

New branches of science and engineering-nano-sensorics, nano-mechanics and nano-bionics are producing very interesting results of creating the newest ***smart nanomaterials***, such results require further consideration and it exceeds the bounds of our work.

11. Conclusions

For all described examples of application of physical phenomena under conditions of exploitation of SBM an action of disturbing factor for appearance of a certain physical phenomenon (effect) in the process of exploitation and as a result –obtaining the specified positive effect may be considered as common signs. Here, a dependence of its quantitative and qualitative indices upon the chemical composition, the nature of

matter or substance, its structure, the conditions and parameters of imitating actions and certainly practical exploitation is revealed.

Thus, the domain of created smart bio-materials is constantly growing, just like new functional opportunities of their application for healing, prosthesis and relieving those who need their application. No doubt, the future of successful developing of perspective bio-medical engineering belongs to such smart materials.

References

1. Worden K., Bullough W.A., Haywood J. Smart Technologies. - World Scientific Publishing, 271 p. (2003)
2. Cheiliakh O.P. and Cheiliakh Ya.O. Implementation of Physical Effects in the Operation of Smart Materials to Form Their Properties // Progress in Physics of Metals. – Vol. 20, No3. – Pp. 363-464 (2020) [in Russian].
3. Orthodontic archwire. Access mode: https://en.wikipedia.org/wiki/Orthodontic_archwire
4. Schwartz M.M. Encyclopedia of Smart Materials (New-York City : John Wiley & Sons: (2002).
5. Smart Materials. https://ru.wikipedia.org/wiki/%D0%A3%D0%BC%D0%BD%D1%8B%D0%B5_%D0%BC%D0%B0%D1%82%D0%B5%D1%80%D0%B8%D0%B0%D0%BB%D1%8B
6. Miyazaki S., Otsuka K. Development of Shape Memory Alloys // ISIJ International. – Vol. 29, No.5. – Pp. 353-377 (1989)
7. Lokhov V.A., Nyashin Y.I., Kuchumov A.G. Application of Shape Memory Alloys in Medicine. Rewiev of the Models Describing their Behaviour // Russian Journal of Biomechanics. – Vol. XI, № 3, P. 9-27 (2007) [in Russian].
8. Bhavsar R., Y.Vaidya N., Ganguly P., Humphreys A., Robisson A., Tu H., Wicks N., McKinley G.H., Pauchet F. Intelligence in Novel Materials // Oilfield Review, Spring. – P. 32-41 (2008). https://publik.tuwien.ac.at/files/publik_288820.pdf
9. Delay Law and New Class of Materials and Implants in Medicine. / V.E. Gunther, G.Ts. Dambaev, P.G. Sysoliatin et al. - Northampton, MA: STT. - 432 p. (2000)
10. Hench L. L., Am. Ceramic Soc., 74 (1991).
11. Vogel W. and Holland W., Angew W. Chem. Int. UK, 26 (1987).
12. Wilson J., Clinical Application of Bioglass, Glass: Current Issues (ed. A. F. Wright and J. Dupuy), (1985).
13. Machado L.G. and Savi M.A. Brazilian Journal of Medical and Biological, 36, No6 (June 2003).
14. TechXpress.net. “Shape Memory Alloy – Nitinol Shape Memory”. jmmedical.com. Retrieved (2017-02-25).
15. *Figueiredo A. M., Modenesi P., Buono V. "Low-cycle fatigue life of superelastic NiTi wires" // International Journal of Fatigue. - 31 (4). – P. 751–758* (*2009-04-01*). doi:10.1016/j.ijfatigue.2008.03.014
16. Teramoto Ohara A. Clinical importance of austenitic final point in the selection of nickel-titanium alloys for application in orthodontic-use arches // Rev. Odont. Mex. – Vol.20. – No3. – P. 162-169 (2016).
17. Hench L. L. and Wilson J. Science, **226** (1984).
18. Klein C. P. A. T., de Blieck-Hogervorst J. M. A., Wolke J. G. C. and de Groot K. Adv. Biomaterials, **9** (1990).

19. Klein C. P. A. T., Patka P., van der Lubbe H. B. M., Wolke J. G. C. and de Groot K. J. Biomed. Mat. Res., **25** (1991).
20. Ikada Y. Biomaterials, **15** (1994).
21. Shelton R. M. and Davies J. E. The Bone-Biomaterial Interface (ed. Davies, J. E.), Toronto University Press (1991).
22. Cooper E., Wiggs R., Hutt D. A., etc. J. Mat. Chem. - 7.- P. 435 (1997).
23. Haddow D. B., France R. M., Short R. D., etc. J. Biomed. Mat. Res. - 47. - P. 379 (1999).
24. Matsuzawa M. Brain Res. - 47. - P. 67 (1994).
25. Khokhlov A.R., Philippova O.E. // Solvents and Self-Organization of Polymers. NATO ASI Series, Series E: Applied Sciences / Ed. By Webber S.E., Munk P., Tuzar Z. Dordrecht; Boston; London: Kluwer Acad. Publ. - V. 327. – P. 197 (1996).
26. Philippova O.E., Hourdet D., Audebert R., Khokhlov A.R. // Macromolecules. - V. 30, №26. – P. 8278 (1997).
27. Galaev Yu. Smart' polymers in biotechnology and medicine, RUSS CHEM REV, **64** (5) (1995). DOI: https://doi.org/10.1070/RC1995v064n05ABEH000161
28. Wise D.L. (ed), Handbook of Pharmaceutical Controlled Release Technology (New-York City: Marcel Dekker: 2002).
29. Sittinger M. Tissue engineering and autologous transplant formation : practical approaches with resorbable biomaterials and new cell culture techniques / M. Sittinger et al. // Biomaterials. - Vol. 17, № 3. - P. 237-242 (1996).

Section 4: Applications and case studies

An Introduction to Tissue Engineering & Bio Printing

Rory Gibney [0000-0003-1670-155X] and Eleonora Ferraris [0000-0003-1618-365X]

Department of Mechanical Engineering, KU Leuven, Leuven, Belgium

Abstract

Organ transplants can be a life-saving treatment to recipients, however, the rate of organ transplant is dictated by the availability of donors and transplant recipients must undergo intense immunosuppressive drug treatment. Artificial hip and knee implants often return a full range of motion to patient's ailing joints, however there is often a time limit to this effect due to the implant's inability to regenerate itself like native tissue. Tissue engineering is a field of medicine in development that aims to generate tissue and ultimately organs for implantation using a patient's cells. It employs our understanding of aspects of biology, such as tissue development, principles of engineering, and material science to attach cells to a material and induce those cells to proliferate, differentiate, and generate material. This chapter aims to introduce the reader to the field of tissue engineering, one of the most commonly used materials in tissue engineering, collagen, and the concept of bioprinting and also aims to prepare the reader for the proceeding chapter on biofabrication of the cornea.

Keywords: Tissue engineering, biofabrication, bioprinting, collagen

1 Tissue Engineering

Tissue engineering (TE) is a relatively new form of engineering first described by R. Langer and J.P. Vacanti as 'an interdisciplinary field that applies the principles of engineering and life sciences towards the development of biological substitutes that restore, maintain, or improve biological tissue function or a whole organ'[3]. The paradigm of TE is that of a circular process, illustrated in **Fout! Verwijzingsbron niet gevonden.**, where i) a biopsy is taken from a patient and cells isolated, ii) those cells are then expanded in culture, iii) the cells are seeded onto a scaffold with signals, iv) the scaffold and cells are cultured until they mature into a tissue, and v) the tissue is implanted into the original patient. The idea being to replace, regenerate, or restore a patient's failing organ, or pathological tissue using the patient's own cells, negating the need for donors, transplants, and the associated immunosuppressive drugs.

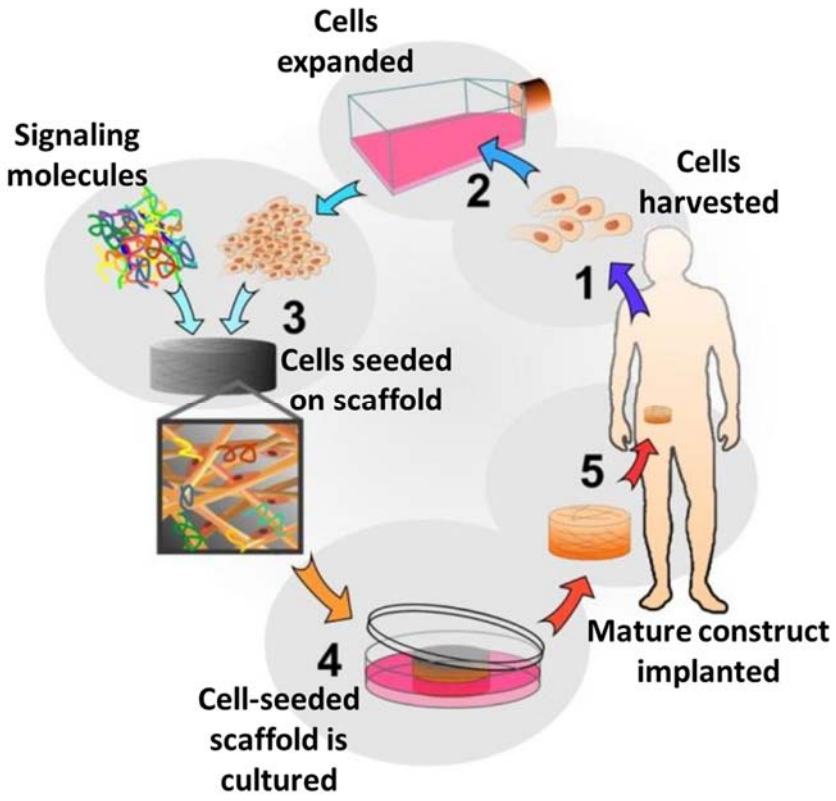


Fig. 1. Paradigm of tissue engineering illustrating: 1, the explant of cells, 2, the expansion of those cells in culture, 3, cells are seeded on a scaffold in the presence of signalling molecules (growth factors, hormones, proteins, or even mechanical stimuli), 4 culture of the construct to tissue maturation, and 5, implantation of the mature tissue construct. Adapted from [4].

The cells typically used in TE are stem cells which can be differentiated into many different cell types using chemical and physical cues such as growth factors, and mechanical stimulation, respectively. Stem cells come in different potencies (capacity for differentiation and proliferation), the most potent also being the most controversial, embryonic stem cells, however multipotent adult stem cells, such as mesenchymal stem cells, are the most commonly used stem cells in TE. Once the cells are isolated from a patient they are expanded in a tissue culture vessel containing culture medium. The vessel can be a tissue culture flask as depicted in step 2 of figure 1, or a petri dish, whilst the medium is a liquid, buffered at a suitable pH, and containing the necessary ions, and nutrients to sustain the cells. Under these conditions, cells are expected to adhere to a surface of the culture vessel and grow in a monolayer until they reach a desired level of “confluence”; the amount of available surface that is covered with adherent

cells. The cells are then harvested using trypsin, an enzyme that metabolises the proteins secreted by cells to attach to the vessel and each other, and re-suspended in fresh media. The re-suspended cells can then be counted, and subsequently used. Although this is a proven method for expanding cells, it is limited in two dimensions with cells only capable of re-producing in monolayers; thus, making it an inefficient and difficult way to generate 3-dimensional tissue.

A TE scaffold, put simply, should resemble a sponge; it is highly porous to allow media to access all volumes in the scaffold, and it has a high surface area to volume ratio. Most human cells (apart from hematopoietic cells) require a suitable substrate so that they can attach and begin interactions, otherwise the cells become rounded and inactive (quiescent). At the scale of a cell, the perimeter of pores of a sponge-like TE scaffold provides a continuous 2-dimensional substrate to which it can attach. Once attached to a TE scaffold the cells can then proliferate, differentiate and begin making their own matrix to take the place of the TE scaffold which should degrade slowly. One way of achieving this sponge-like morphology is through the use of a class of materials called hydrogels. Hydrogels are made of a polymer network swollen with water, an everyday example being the dessert, jelly, which is made of a gelatine swollen in water. Although on the surface jelly does not appear very sponge-like, on the microscale it is filled with pores. The porous, heavily hydrated polymer network of hydrogels is a useful as mimic of the extracellular matrix (ECM) for tissue engineering. However, there are many other ways of creating porous scaffolds, such as gas foaming, salt-leaching, and bio fabrication strategies such as bio printing and electrospinning.

1.1 The extracellular matrix

In order to understand cell attachment it is necessary to understand how cells interact with their extracellular matrix (ECM). The ECM is a network of proteins, glycoproteins, proteoglycans, and polysaccharides that gives all tissue its physical shape and acts as a communication network between cells. The composition of the ECM is specific to each tissue but some common components are the protein, collagen, and the glycoprotein, fibronectin. Cells interact with their surrounding ECM through proteins in the cell membrane (cell adhesion molecules, CAMs, such as integrins) attaching to complementary sites in the ECM components, as seen in Figure 2. The complementary sites are usually in the form of short amino acid sequences, such as RGD, which is naturally present in fibronectin. TE scaffolds often use extracted ECM components, or synthetic ‘motifs’ (synthetic short peptide sequences similar to RGD), to exploit their capacity for cell attachment, or modify artificial polymers, such as PEG, with RGD or similar peptides to increase their capacity for cell attachment. Other methods to improve cell attachment include increasing the hydrophilicity (or wettability) of a surface which is carried out on the plastic used to culture cells, TCP (tissue culture plastic), or coating a scaffold with in a thin layer of a biopolymer such as collagen.

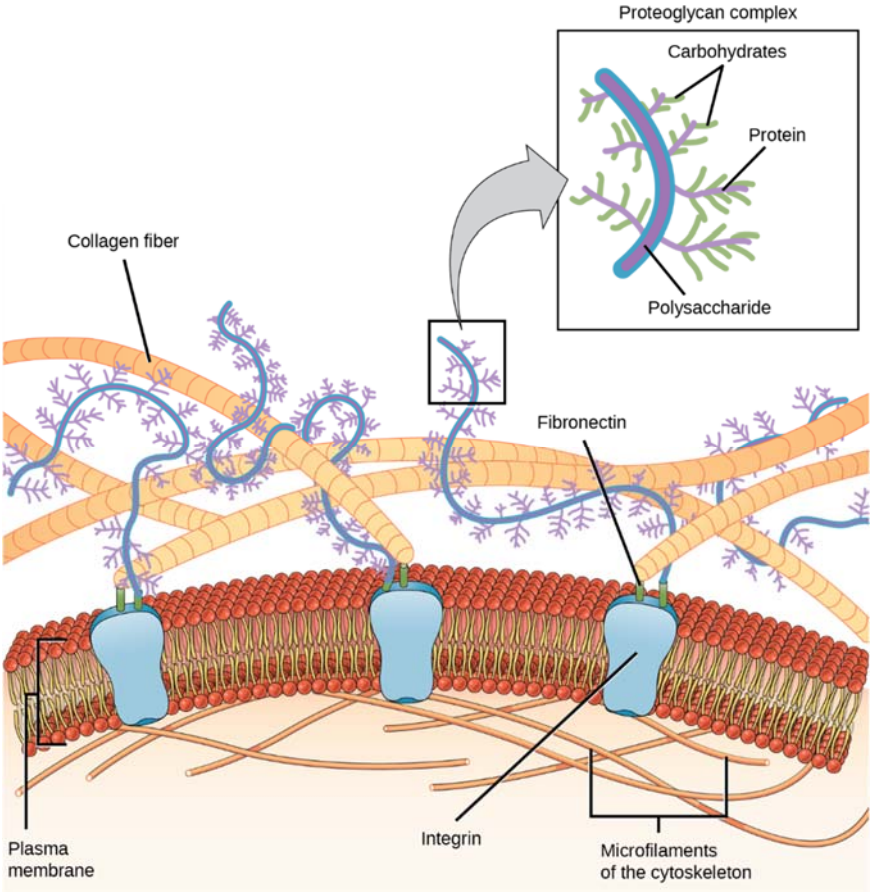


Fig. 2. An illustration of interactions between a cell and the extracellular matrix. The cell membrane (plasma membrane) is made of a phospholipid bilayer with integral proteins, integrins mediating the interaction of the plasma membrane with proteins such as collagen, and proteoglycans. Attachment is often through glycoprotein, fibronectin. [5] ©2017 Rice University

Once cells have attached to a scaffold then regulatory signals are often used to influence their behavior. These signals can be in the form of chemicals added to the cell culture media (growth factors, biomolecules), mechanical stimuli, or electrical stimuli. Growth factors are the most common signals. They are polypeptides or proteins (mostly) that are recognized by cells at their surface and induce intracellular signaling pathways that regulate the cell behavior and development. For instance, vascular endothelial growth factor (VEGF) can stimulate vasculogenesis (the production of blood vessels) by vascular endothelial cells through inducing their proliferation and promoting cell migration [6]. Vasculogenesis is one of the major challenges facing TE as it approaches questions of scalability, as capillaries need to be generated to deliver oxygen to the center of scaffolds. Mechanical stimuli can also be used on some cells such

as osteoblast. Electrical stimuli has been shown to enhance the differentiation and maturation of neurons and cardiomyocytes, respectively [7],[8]. This combination of cells, a scaffold, and signals illustrated in Step 3 of the paradigm in (**Fig. 2**) is often referred to as the "triad of tissue engineering" or "the three pillars of tissue engineering" [9]. When a level of automation is included through bio printing, electrospinning, or bio assembly, this is considered bio fabrication. The level of automation in bio fabrication has expanded relatively quickly from being employed solely for the manufacture of a scaffold, to being employed for scaffold fabrication, cell deposition, and fabrication of an encapsulating bio-reactor [10].

1.2 Collagen as a biomaterial for bio fabrication

The term biomaterial is used to refer to a wide range of materials including synthetic polymers, ceramics, and metals, along with biologically-derived materials, their unifying quality is that they are used to replace, restore, or regenerate tissue. However, the definition of a biomaterial is quite a contentious issue. A comprehensive and widely-accepted leading opinion paper on the definition of biomaterials was published by David F. Williams in 2009 [11]. The biomaterials used to build TE scaffolds must possess certain qualities, among which are that they must be biocompatible, and biodegradable. For a material to be defined as biocompatible, the material, its metabolites and any leachates from production, must not be toxic to cells. Biodegradation is necessary so that the construct can slowly degrade and be absorbed by the recipient whilst cells in the construct, or recruited by the construct, replace the biomaterial with their own matrix. As stated previously, cell attachment is necessary for cells to remain active, hence cells should be able to attach to a biomaterial for it to be considered effective for use in TE. This is relatively straightforward for ECM-derived biomaterials such as collagen and fibronectin since they possess CAMs, however, many biomaterials do not inherently possess CAMs. For these biomaterials the potential for cell attachment can be increased by increasing the hydrophilicity of their surface. Better still, they can be functionalized with CAMs such as the peptide sequence RGD.

Collagen is the most abundant protein in mammals. It is found in high concentrations in tendons, bone, the cornea, cartilage, and many other tissues in vertebrates. Rather than a single protein, collagen is a family of proteins: the collagen supra-family. All collagens possess a characteristic triple helix structure composed of three interwoven α polypeptide chains which may be identical (homotrimer) or a mix of two or more genetically distinct α chains (heterotrimer). The collagen supra-family includes 28 different types of collagen which all possess one or more triple helix domains with various other structures, and functions. The fibrillary collagens are a sub-family that contain the most common collagens, and whose structure is dominated by a single triple helix and that all assemble into fibrils with a characteristic banded pattern. The fibrillary collagens are the most common collagens, and consist of collagen type I, II, III, V, XI, XXIV, and XXVII. Collagen types I, II, and III, are often referred to as the major fibrillary collagens due to their abundance, whereas collagens type XXIV and XXVII are outliers within the fibrillary collagens because they have a much shorter collagenous domain.

The characteristic banded pattern of the fibrillary collagens (light and dark bands with a periodicity of ~67 nm seen in Figure 3) arises from the packing order of the collagen molecules into fibrils. The packing order is derived from fibrillogenesis (fibril assembly) which is a well-studied phenomenon where collagen molecules self-assemble, aligning side-by-side to each other with a slight offset, about a quarter of their length, and end-to-end with other similarly stacked collagen molecules. The overlapping of the gaps between collagen molecules gives rise to the light bands which account for ~0.46 of the d-period (67nm), with the dark band, the overlapping of the ends of collagen molecules, accounting for the other 0.54.

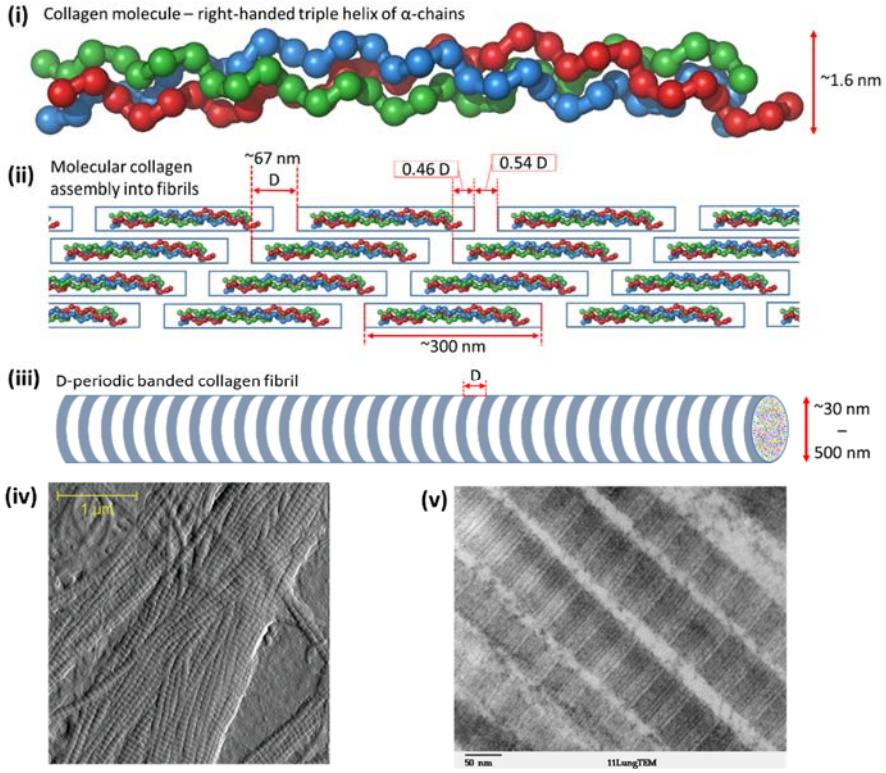


Fig. 3. The structural hierarchy of the major fibrillar collagens, (i) illustrates the assembly of three left-handed helical α-chains into a right handed triple helix of around 300nm in length, (ii) a schematic of collagen fibrillogenesis showing the stacking of collagen molecules side-by-side and end-to-end, (iii) the resultant a D-banded collagen fibril from fibrillogenesis. Microscopy images of collagen fibrils: (iv) AFM of collagen fibrils made in vitro by osteoblasts [12], and (v) TEM of collagen fibrils in lung tissue banded collagen (scale bar: 50 nm).

Fibrillogenesis is a thermodynamically driven process; molecular collagen is unstable under physiological conditions, however, it becomes stable by self-assembling into fibrils. In vitro, fibrillogenesis is mainly controlled by pH, ionic strength, and temperature. Other influencing factors include the ion species, surfactants, and saccharides present, and the presence/absence of telopeptides [13]. These factors are exploited to

make the collagen hydrogels commonly used in TE. Different fibrillar architectures can be exploited by using different conditions, for instance, incubating the neutralised 1X PBS collagen solution at low temperatures (~4°C) slows down fibrillogenesis and allows the fibrils to grow far thicker, but also shorter.

Collagen is one of the most commonly exploited biopolymers, due to its ubiquitous nature. The gelatine that makes jelly is derived from collagen-rich tissue such as skin and bone. Gelatine is collagen that has been degraded (broken down on a molecular level) and denatured (protein losing its shape) through extraction processes involving high heat and extreme pH values to the point where it possesses very different characteristics and a fraction of its CAMs. Less severe processes at low temperatures (~ 4°C) can be used to extract intact collagen, such as acid-soluble collagen (telocollagen), and pepsin-soluble collagen (atelocollagen). Telocollagen and atelocollagen are distinguished by the presence or absence of telopeptides. Different tissues can be excised for the extraction of different collagens, for instance, the Achilles tendon is very rich in collagen type I, whereas cartilage tissue is very rich in collagen type II. These tissues are of waste of the meat industry, hence much collagen is derived from bovine, ovine, and porcine sources. Rat-tails are also a rich source of collagen type I that is frequently extracted in TE research labs.

The possibility for interspecies disease transmission and allergic or immune reactions can make animal derived collagens less desirable, hence collagen is also sourced from humans, the most common source being the human placenta/amniotic membrane. However, extracted human collagen may still illicit an immune response. Recombinant collagens have received growing interest due to the elimination of the possible risks associated with the use of animal-derived collagens in humans (disease transmission, allergens, pathogens, immunogenicity). Recombinant proteins are produced by inserting the relevant gene for a protein into a host organism, which is often yeast since it is a well-researched organism that can be cultured in large volumes (in bioreactors rather than 2D culture flasks). This results in a purer product in contrast to extracted proteins and allows for alterations in the amino acid sequence through gene modification. However, the simple host organisms used often lack the cellular tools to produce collagen in the exact same way as a mammalian cell. For instance, the amino acid, hydroxyproline, is not made in ribosomes; it begins its life as proline, and is turned into hydroxyproline by an enzyme in the cytoplasm of mammalian cells. This enzyme has to be supplemented to the yeast cells so that hydroxyproline may be formed, which is important to collagen's structure.

2 Bio printing

Bio printing is a branch of bio fabrication defined as the automated deposition of living and non-living materials in a prescribed geometry for the production of bioengineered structures [14]. The way in which cells are incorporated into a bio printed structure varies, from simply seeding cells onto a scaffold, to bio printing cells or cell spheroids of a single cell type, to the precise deposition of specific cell types in specific patterns to illicit the maturation of a particular tissue. Should an ink contain cells, it is referred

to as a bio-ink, whereas an ink containing a hydrogel/hydrogel-precursor intended to be printed alongside a bio-ink, or for cell seeding is referred to as a biomaterial ink [15]. Conventional inkjet printers were used as the first bio-printers. Robert J. Klebe performed the first case of bio-printing, or 'cytoscribing' in 1988 by filling the inkjet cartridges of two HP printers with a fibronectin solution, and a solution containing monoclonal antibodies, thereby creating a substrate to which cells could specifically attach in subsequent culture [16]. The first cell printing was performed by Boland & Wilson in 2003 using the firmware and parts from a HP660C inkjet printer but with a modified dispensing system to allow the use of wider nozzles through which large mammalian cells could pass [17]. Since then the field of bio-printing has expanded at a remarkable pace with a number of printing technologies, each of which have their own particular advantages and drawbacks. Bio-printing technologies are usually classified as one of the following: extrusion-based bio-printing(EBB), droplet-based bio-printing(DBB), and laser-based bio-printing(LBB) [2]. More recently, there has been the development of multi-modal printers which integrate different types of print-heads and printing technologies into one system to fulfil multiple needs.

Similar to regular 3D printing, bio-printing frequently makes use of different inks for different purposes within the same print. A sacrificial ink is an ink that may be used to support in voids within a scaffold and prevent the surrounding material from collapsing into the void. The sacrificial ink can then be removed afterwards usually using a solvent or conditions, to which only the sacrificial ink is vulnerable. A cell carrier ink, is a bio-ink that contains cells, but may not have sufficient matrix materials to support itself or an entire scaffold. These are often used where the biomaterial used as the main structural component of the scaffold cannot be printed with cells; thermoplastics such as polycaprolactone. Printing methods have also been developed where the whole scaffold is printed within a support bath made of a thixotropic hydrogel/hydrogel suspension.

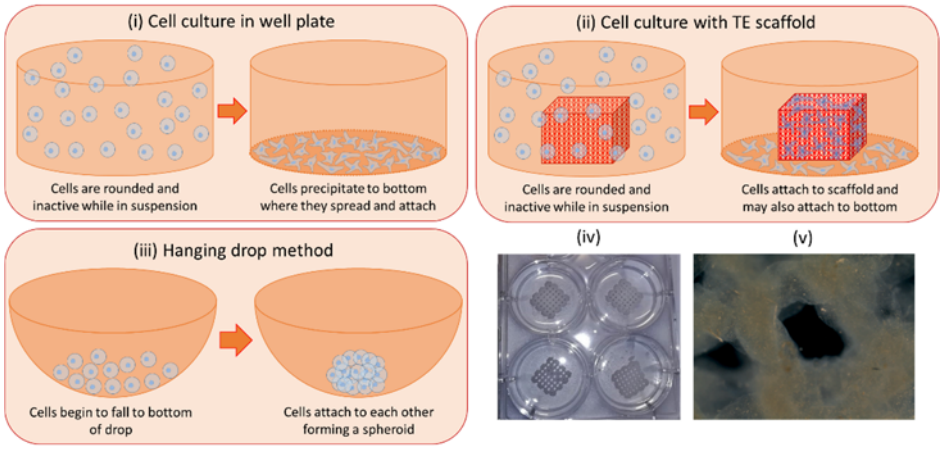


Fig. 4. illustration of some concepts of cell culture and bio-printing. (i) cells precipitate out of suspension when left in a well plate and attach to the bottom, (ii) when cells are incubated with a suitable TE scaffold the cells attach to the scaffold, (iii) the hanging drop method for forming cell spheroids. (iv) An example of bio-printed scaffolds for TE, and (v) a microscope image of an acellular TE scaffold.© R. Gibney 2020

Cell aggregates are structures composed of cells loosely bound to each other, typically by ECM biopolymers secreted by the cells. A monolayer of cells grown in a tissue culture vessel can be broken up into aggregates by harvesting them without the using of ECM degrading enzymes, like trypsin. However, using this method the aggregates are randomly sized. A popular method for producing more consistent spheroid-shaped cell aggregates (cell spheroids) is the hanging drop method. The cell spheroids are formed by pipetting a cell suspension, in small amounts with an estimated number of cells, on to a petri dish which is then inverted. Gravity allows the cells to settle out of suspension and form a cluster at the bottom of the droplet, away from the TCP. The cells, without a substrate for attachment, attach to each other

2.1 Extrusion-based bio-printing (EBB)

EBB is the most common type of bio-printing due to the simplicity of its design. EBB typically loads an ink, usually a hydrogel/hydrogel-precursor, into a syringe, which is then extruded through a nozzle by applying a force by mechanical or pneumatic means. EBB constructs usually fit into a 6, 12, or 24 well plate. Print accuracy is totally dependent on the user identifying the suitable print parameters for the rheology of their ink; print parameters such as: extrusion pressure/speed, and print speed. The ink must have some means by which it can transition from being a fluid with suitable viscosity for extrusion, to hydrogel with sufficient viscoelasticity to support its shape and provide support for proceeding layers. A quicker transition is also advantageous, as it reduces print time. This transition can cause the geometry of the printed material to change as new bonds being formed, or chemical structural changes can cause contraction of the material. Therefore, print fidelity is just as important as print accuracy. This is relatively

straightforward for some hydrogels like gelatin which exhibits thermogelation, alginate which quickly transitions from a solution to a crosslinked gel in the presence of divalent ions (ionic crosslinking), or cellulose nanocrystal hydrogels which exhibit severe shear-thinning. The combination of two or more of these materials also leads to the development of composite inks, which can achieve improved print quality along with enhanced bioactivity/cell survival or other characteristics (e.g. higher stiffness). For instance, Wu et al. made a bio-ink containing human corneal epithelial cells (HCECs) that exploited the thermogelation of gelatin, alginate's ionic crosslinking for stability, and collagen for bioactivity in order to provide a better mimic of native ECM [18]. There has also been widespread use of composite inks in bone tissue engineering such as the "3-3-3" ink by Cidonio et al. which is comprised of 3% laponite, 3% alginate, and 3% methylcellulose resulting in an ink with high printability and high bioactivity [19]. The results achieved with composite inks has improved EBB's performance when compared to DBB and LBB methods which generally have better print accuracy, and print fidelity. EBB methods are also more frequently used for the deposition cell aggregates compared to their DBB, and LBB counterparts, as the aggregates are often too large for most DBB and LBB methods. The simplicity of the EBB design allows EBB printers to have multiple printheads so that several inks can be used such as a cell-carrier ink, a build ink and a sacrificial ink as seen in Figure 6 (i).

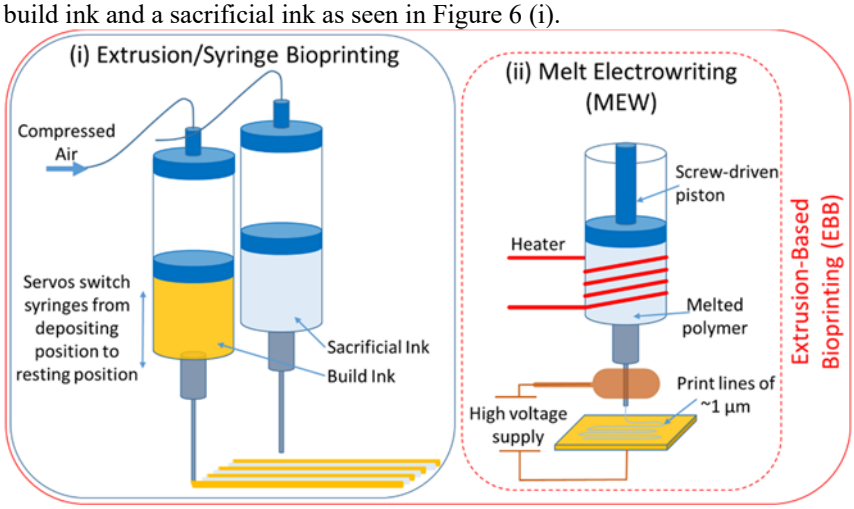


Fig. 5. (i) Extrusion bioprinting usually involves more than one printhead, with a build ink which will form the end-construct, a sacrificial ink, which is removed after printing, and a cell carrier ink. (ii) Melt electrowriting involves melting a polymer and extruding it through a nozzle attached to the anode of a high voltage power supply, and the build platform attached to the cathode, leading to extruded filaments as small as 1 µm in diameter. © R. Gibney 2020

2.2 Droplet-based bio-printing

The inkjet printers used in the first cases of bio-printing are DBB methods. They use thermal, acoustic, or mechanical systems to eject a droplets of an ink on to a substrate.

Inkjet printers have continued their development in cell printing. However, they struggle to produce scalable TE scaffolds due to a number of limitations. Inkjet printers can only print inks in a limited viscosity range (<15mPas), and work properly with cell densities <1× 106 cells/ml. In addition, they cannot be used to print cell aggregates [20]. They are also known to be prone to blockages, particularly with increasing print time due to cells or particulates settling out of suspension within the ink reservoir [21]. However, inkjet printers have been used alongside other printers and have been incorporated into some multi-modal bio-printers. DBB can also be used effectively to deposit a binder solution onto some form of polymer pre-cursor or inorganic material, for example, Inzana et al. used a collagen solution in phosphoric acid to bind calcium phosphate powder in bone tissue engineering, where a collagen solution in phosphoric acid was used to bind calcium phosphate powder [22].Other DBB techniques include microvalve-bio-printing, drop-on-demand bio-printing, and acoustic droplet ejection.

The Bioscaffolder 3.2 utilises an inkjet printhead to aspirate cell suspensions from a 96 well plate and deposit them on the build plate, along with multiple EBB print-heads for the deposition of more structurally relevant materials i.e. hydrogel, polymer (PCL) inks, and sacrificial materials. In this way, an inkjet printhead can print many different types of cells during the same print which is particularly attractive for the manufacturer of biosensors and organ-on-chip microfluidic devices [23].

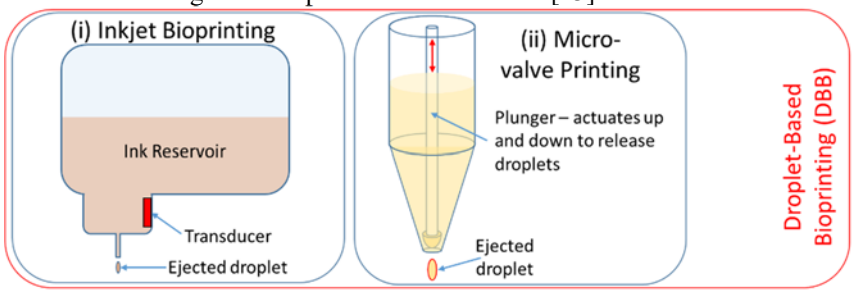


Fig. 6. Droplet based bioprinting methods: (i) inkjet, where a transducer (piezoelectric, thermal) forces a droplet out of the nozzle. (ii) Microvalve printing where an electromagnetic plunger is actuated up and down to force droplets out of the print head. © R. Gibney 2020

2.3 Light-based bio-printing (LBB)

LBB methods employ light to eject droplets of an ink, propel cells in suspension, or to polymerise an ink. Photo-polymerisation techniques typically employ a photo-initiator which absorbs light and releases radicals which are highly reactive unstable molecules that help polymerize monomers in the ink or crosslink molecules in the ink. The photoinitiators used for photo-polymerisation often have low solubility in water, exhibit cytotoxicity in their stable or activated state, or may require excitation by light in the UV range, which has limited the use of photo-polymerisation for cell encapsulation/cell printing, settling instead for cell seeding post print. However, photoinitiators have been developed with improved water-solubility, less cytotoxicity, and higher excitation wavelengths, such as Irgacure I2959 and LAP (lithium phenyl-2,4,6-trimethyl- benzoylphosphinate). Both I2959 and LAP have been used to encapsulate cells with a high

level of cell viability [24, 25], they have also been used as photoinitiators in EBB. LBB photo-polymerisation techniques include 2-photon polymerization (2PP), stereo lithography (STL), and digital light processing (DLP)[26].

Other LBB methods include laser-assisted forward transfer (LAFT), laser-induced forward transfer (LIFT), and laser-guidance direct writing (LGDW). They are typically used for cell patterning and may become more popular as the science of scaffold-free TE evolves, but, for the moment, they are the lesser spotted bio-printing technologies. LIFT and LAFT are used interchangeably to refer to the same method of printing. In brief, a "donor" slide is coated with a thin photo-absorbing layer, often gold, this layer is then coated with a bio-ink, the slide is then placed upside-down above a substrate and a laser is pulsed through the slide and absorbed by the photo-absorbing layer causing extremely local high temperatures. This leads to the formation of a bubble which then collapses causing a droplet in the pico-litre range to be ejected from the bio-ink [27–29]. LGDW is an even rarer LBB method which propels individual cells in a suspension toward a substrate using light [30–32]. LIFT/LAFT and LGDW are both nozzle-free cell printing techniques and are shown to have some of the highest cell survival rates of all bio-printing techniques and have been utilised with many cell types, yet they are more useful for the growing field of scaffold-free tissue engineering, particularly LIFT since it allows the printing of many cell types by simply changing the donor slide.

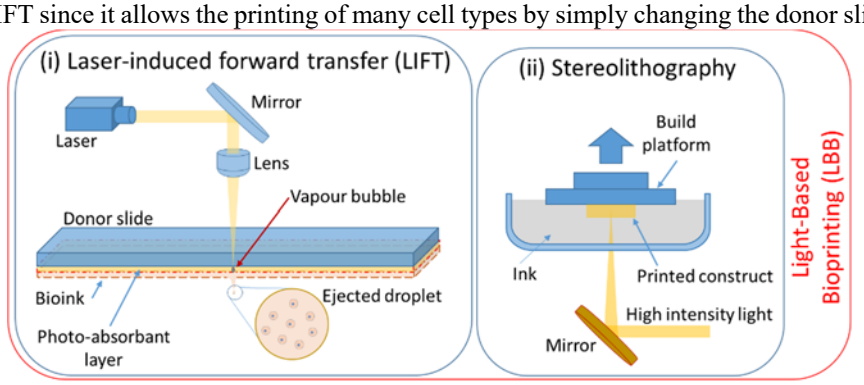


Fig. 7. Examples of light based bio-printing methods: (i) LIFT, where a pulse of light from a laser is focused through a donor slide to a photo-absorbent layer causing it to sublimate and produce a vapour bubble which quickly collapses forcing a droplet to be ejected from the bioink layer. (ii) Stereo lithography where a beam of high intensity light is directed with aid of a mirror into an ink bath where it polymerises the ink only at the focal point of the light. The build is polymerised onto a build platform which is slowly moved out of the ink bath. © R. Gibney 2020

2.4 Multi-modal bio-printers

Each of EBB, DBB, and LBB method has its own advantages and disadvantages, it therefore makes sense that these be incorporated into one bio-printing system to get the best of both, or all worlds. Multiple bio-printing research groups are now working with multi-modal printers, and several commercial multi-modal bio-printers are available, such as the GeSiM Bioscaffolder 3.2, and the RegenHU. Multi-modal bio-printers typ-

ically incorporate different types of EBB methods with DBB methods, such as a pneumatic EBB printhead and a microvalve, or inkjet DBB printhead. Several pneumatic printheads maybe in the same system so that the printer can quickly change between different types of inks. This offers many advantages as it allows the deposition of many different cell types with control over cell count distribution, and allows for the use of inks suited to specific roles, such as a sacrificial ink, cell-carrier ink, and a stiff polymer for mechanical strength. Some multi-modal printers can be procured with a melt-electrowriting (MEW) printhead. MEW is a printing technique pioneered by Paul Dalton at the University of Würzburg which is similar to electrospinning, where electrostatic forces build up in a polymer droplet at the tip of a syringe due to an applied high voltage electric field. The electrostatic forces build up until they overcome the surface tension of the droplet and a filament leaves the droplet and whips around due to the electric field, stretching the filament into nanofibers, before reaching the cathode. MEW prevents the whipping action, which is difficult to control, through a number of processes, including by having the syringe tip and substrate (the cathode) closer together. MEW creates thicker filaments (>1 μm) than electrospinning but in a controlled way that can be incorporated into printer. MEW printheads in multi-modal printers has led MEW lattices being used to encapsulate cell spheroids with some very promising results at published by Daniel J. Kelly at Trinity College Dublin.

3 Discussion & Conclusion

TE first captured public imagination when the image, “Vacanti mouse”, went viral. It inspired much discussion in the media, and memes, and even parodies. This was an ear-shaped polymeric scaffold populated with bovine cartilage cells that was implanted under the skin of a nude mouse (an immunocompromised mouse that will not experience implant rejection). The mouse in this case acted as a bioreactor for the scaffold, allowing the cartilage cells to develop their own matrix as the polymer degraded, and blood vessels to form. This was a proof-of-concept for a xeno-graft model, with a stated intention to proceed to an autologous model where a similar scaffold populated with the patient’s cartilage cells would be implanted under the skin of the patient to allow it to mature into a functional auricular cartilage [33]. Unfortunately at times the image was misinterpreted, and it was thought by some that the mouse was genetically-engineered to grow a human ear on its back, and became a tool in arguments against genetic engineering.

Bio-printing has been victim to similar misinterpretations. Many bio-printer manufacturers, and academics, focus on the more distant possibilities of bio-printing order to stir up interest and attain funding. An image search of “bio-printing” will quickly return some computer-generated images of whole hearts or ears being produced by a printer, which is almost comical to any academics in the bio-printing field until patients on donor waiting lists and their families are considered. Bio-printing has a number of challenges ahead particularly regarding scalability before it could impact medicine; bio-printing whole organs remains in the realm of science-fiction. Bio-printing is likely to be applied far more quickly in drug screening using organ-on-chips, which are bio-

printed miniature versions of functional human organs in microfluidic chips that have become a sub-section of bio-printing. Organ-on-chips are likely to become invaluable tools to the pharmaceutical industry. TE and bio-printing have major manufacturing and regulatory challenges ahead. The typical regulatory framework applied to medical devices is not applicable to a living TE tissue. A TE manufacturer has to be placed close to a hospital (its source of cells), with the patient being treated in that hospital, which makes international distribution difficult to comprehend. However, the possible benefits of implantable functional TE tissues far outweigh the manufacturing and regulatory challenges ahead.

4 References

1. Burdick J.A., Mauck R.L.: Biomaterials for Tissue Engineering Applications (2011).
2. Sun W., Starly B., Daly A.C., Burdick J.A., Groll J., Skeldon G., Shu W., Sakai Y., Shinohara M., Nishikawa M., Jang J., Cho D.-W., Nie M., Takeuchi S., Ostrovidov S., Khademhosseini A., Kamm R.D., Mironov V.A., Moroni L., Ozbolat I.T.: The bioprinting roadmap. Biofabrication 12 (2020).
3. Langer R., Vacanti J.P.: Tissue Engineering. Science (80-) 260(5110), 920–926 (1993).
4. George J.H.S.: Engineering of Fibrous Scaffolds for use in Regenerative Medicine. Imperial College London (2009).
5. Rye C., Wise R., Jurukovski V., Jean D., Choi J., Avissar Y.: Biology (2017).
6. Ferrara N., Gerber H.P., LeCouter J.: The biology of VEGF and its receptors. Nat Med 9(6), 669–676 (2003).
7. Yamada M., Tanemura K., Okada S., Iwanami A., Nakamura M., Mizuno H., Ozawa M., Ohyama-Goto R., Kitamura N., Kawano M., Tan-Takeuchi K., Ohtsuka C., Miyawaki A., Takashima A., Ogawa M., Toyama Y., Okano H., Kondo T.: Electrical Stimulation Modulates Fate Determination of Differentiating Embryonic Stem Cells. Stem Cells 25, 562–570 (2007).
8. Chan Y.C., Ting S., Lee Y.K., Ng K.M., Zhang J., Chen Z., Siu C.W., Oh S.K.W., Tse H.F.: Electrical stimulation promotes maturation of cardiomyocytes derived from human embryonic stem cells. J Cardiovasc Transl Res 6, 989–999 (2013).
9. O'Brien F.J.: Biomaterials & scaffolds for tissue engineering. Mater Today 14(3), 88–95 (2011).
10. Costa P.F., Vaquette C., Baldwin J., Chhaya M., Gomes M.E., Reis R.L., Theodoropoulos C., Hutmacher D.W.: Biofabrication of customized bone grafts by combination of additive manufacturing and bioreactor knowhow. Biofabrication 6(3), 035006 (2014).
11. Williams D.F.: On the nature of biomaterials. Biomaterials 30(30), 5897–5909 (2009).
12. Canelón S.P., Wallace J.M.: β-Aminopropionitrile-induced reduction in enzymatic crosslinking causes in vitro changes in collagen morphology and molecular composition. PLoS One 11(11), 1–13 (2016).
13. Li Y., Asadi A., Monroe M.R., Douglas E.P.: pH effects on collagen fibrillogenesis in vitro: Electrostatic interactions and phosphate binding. Mater Sci Eng C 29(5), 1643–1649 (2009).

14. Moroni L., Boland T., Burdick J.A., De Maria C., Derby B., Forgacs G., Groll J., Li Q., Malda J., Mironov V.A., Mota C., Nakamura M., Shu W., Takeuchi S., Woodfield T.B.F., Xu T., Yoo J.J., Vozzi G.: Biofabrication: A Guide to Technology and Terminology. Trends Biotechnol 36(4), 384–402 (2018).
15. Groll J., Burdick J.A., Cho D., Derby B., Gelinsky M., Heilshorn S.C., Jüngst T., Malda J., Mironov V.A., Nakayama K., Ovsianikov A., Sun W., Takeuchi S., Yoo J.J., Woodfield T.B.F.: A definition of bioinks and their distinction from biomaterial inks. Biofabrication 11 (2019).
16. Klebe R.J.: Cytoscribing: A method for micropositioning cells and the construction of two- and three-dimensional synthetic tissues. Exp Cell Res 179(2), 362–373 (1988).
17. Wilson W.C., Boland T.: Cell and organ printing 1: Protein and cell printers. Anat Rec - Part A Discov Mol Cell Evol Biol 272(2), 491–496 (2003).
18. Wu Z., Su X., Xu Y., Kong B., Sun W., Mi S.: Bioprinting three-dimensional cell-laden tissue constructs with controllable degradation. Sci Rep 6(24474) (2016).
19. Ahlfeld T., Cidonio G., Kilian D., Duin S., Akkineni A.R., Dawson J.I., Yang S., Lode A., Oreffo R.O.C., Gelinsky M.: Development of a clay based bioink for 3D cell printing for skeletal application. Biofabrication 9(3) (2017).
20. Mandrycky C., Wang Z., Kim K., Kim D.H.: 3D bioprinting for engineering complex tissues. Biotechnol Adv 34(4), 422–434 (2016).
21. Pepper M.E., Seshadri V., Burg T.C., Burg K.J.L., Groff R.E.: Characterizing the effects of cell settling on bioprinter output. Biofabrication 4(1) (2012).
22. Inzana J.A., Olvera D., Fuller S.M., Kelly J.P., Graeve O.A., Schwarz E.M., Kates S.L., Awad H.A.: 3D printing of composite calcium phosphate and collagen scaffolds for bone regeneration. Biomaterials 35(13), 4026–4034 (2014).
23. Tse C.C.W., Smith P.J.: Inkjet Printing for Biomedical Applications. In: Ertl P, Rothbauer M (eds) Cell-Based Microarrays: Methods and Protocols. Springer New York, New York, NY, pp 107–117 (2018).
24. Fairbanks B.D., Schwartz M.P., Bowman C.N., Anseth K.S.: Photoinitiated polymerization of PEG-diacrylate with lithium phenyl-2,4,6-trimethylbenzoylphosphinate: polymerization rate and cytocompatibility. Biomaterials 30(35), 6702–6707 (2009).
25. Lin H., Zhang D., Alexander P.G., Yang G., Tan J., Cheng A.W.M., Tuan R.S.: Application of visible light-based projection stereolithography for live cell-scaffold fabrication with designed architecture. Biomaterials 34(2), 331–339 (2013).
26. Kim S.H., Yeon Y.K., Lee J.M., Chao J.R., Lee Y.J., Seo Y.B., Sultan M.T., Lee O.J., Lee J.S., Yoon S. Il, Hong I.S., Khang G., Lee S.J., Yoo J.J., Park C.H.: Precisely printable and biocompatible silk fibroin bioink for digital light processing 3D printing. Nat Commun 9(1) (2018).
27. Koch L., Deiwick A., Franke A., Schwanke K., Haverich A.: Laser bioprinting of human induced pluripotent stem cells — the effect of printing and biomaterials on cell survival , pluripotency , and differentiation Laser bioprinting of human induced pluripotent stem cells — the effect of printing and biomaterials o
28. Koch L., Deiwick A., Schlie S., Michael S., Gruene M., Coger V., Zychlinski D., Schambach A., Reimers K., Vogt P.M., Chichkov B.: Skin tissue generation by laser cell printing. Biotechnol Bioeng 109(7), 1855–1863 (2012).

29. Sorkio A., Koch L., Koivusalo L., Deiwick A., Miettinen S., Chichkov B., Skottman H.: Human stem cell based corneal tissue mimicking structures using laser-assisted 3D bioprinting and functional bioinks. Biomaterials 171, 57–71 (2018).
30. Odde D.J., Renn M.J.: Laser-guided direct writing for applications in biotechnology. Trends Biotechnol 17(10), 385–389 (1999).
31. Odde D.J., Renn M.J.: Laser-guided direct writing of living cells. Biotechnol Bioeng 67(3), 312–318 (2000).
32. Nahmias Y., Odde D.J.: Micropatterning of living cells by laser-guided direct writing: Application to fabrication of hepatic-endothelial sinusoid-like structures. Nat Protoc 1(5), 2288–2296 (2006).
33. Cao Y., Vacanti J.P., Paige K.T., Upton J., Vacanti C.A.: Transplantation of Chondrocytes Utilizing a Polymer-Cell Construct to Produce Tissue-Engineered Cartilage in the Shape of a Human Ear. Plast Reconstr Surg 100(2), 297–302 (1997).

Bio-Fabrication of Corneal Substitutes

Rory Gibney [0000-0003-1670-155X] and Eleonora Ferraris [0000-0003-1618-365X]

Department of Mechanical Engineering, KU Leuven, Leuven, Belgium

Abstract

The cornea, in simplest terms, is the transparent part of the outer eye. It is surrounded by transitional ring of tissue called the limbus, and the sclera, or white of the eye, beyond. It serves two main functions: protecting the inner eye and focussing light into the eye. The transparency of the cornea is due to a number unique features and dynamic processes, and disruption to any one of them can result in blindness, corneal opacities or visual impairment. In some cases, where corrective lenses, topical medication or other treatments are insufficient, a corneal transplant is performed, which despite being the most performed transplant, globally, suffers from a shortage of donors. In this chapter, we wish to build on concepts introduced in the previous chapter and explore the cornea, the occurrence of blindness and other visual impairments, corneal transplants, and the potential for bioprinting to generate a biofabricated cornea that could be used as an alternative to transplant.

Keywords: Tissue engineering, biofabrication, corneal tissues, collagen

1 Anatomy of the cornea

The cornea is distinguished from other parts of the outer eye by its transparency and curvature. As light hits the cornea, it is refracted (the cornea is responsible for 75% of refractive power of the vision system of the eye) through the pupil into the lens where the light is further focused on to the cone cells of the retina. The cornea is a slightly oblate spheroid cap shell, with a horizontal diameter around 11.5 mm, and a vertical diameter of around 10.6 mm white-to-white [1], and a central corneal thickness of around 550 μm which increases with the distance from the centre up to 700-800 μm near the limbus [1, 2]. The radius of curvature of the anterior corneal surface is often reported to be around 7.7mm [3]. The cornea is composed of 5 distinct layers, as seen in Figure 1: i) the anterior surface is the epithelium, followed by ii) the acellular Bowman's Membrane, iii) the stroma which accounts for the bulk of the cornea, iv) the Descemet's membrane, another acellular layer, and v) the endothelium.

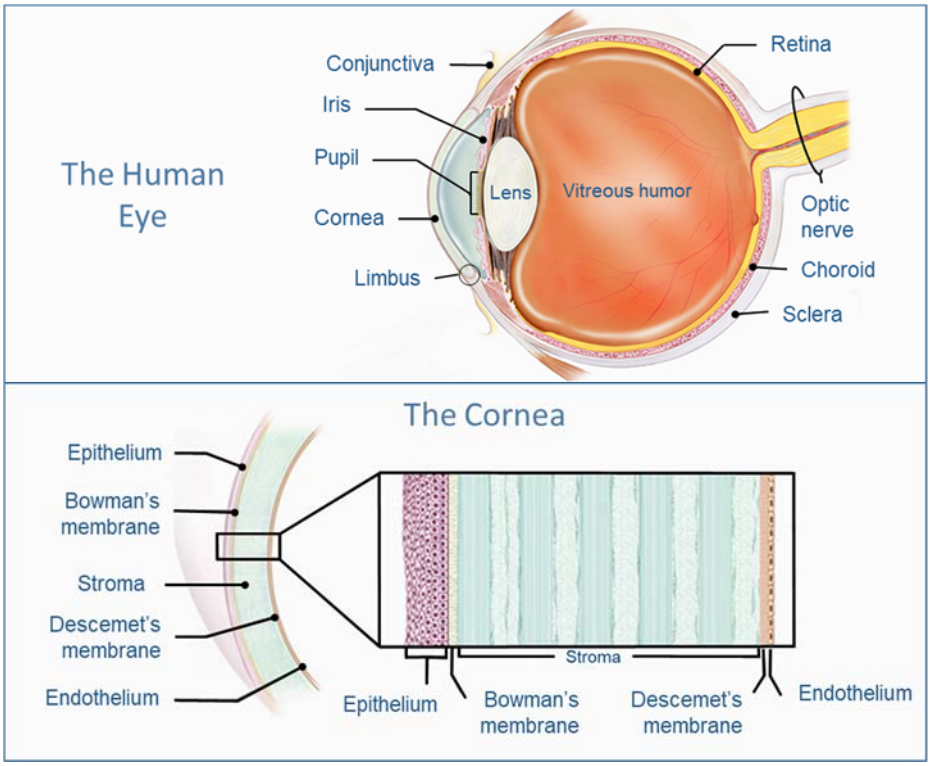


Fig. 1. Diagram of the human eye and the cornea. The cornea has an increased curvature compared to sclera. Contrasting layers in the corneal stroma depict layers of orthogonally oriented fibrils. Adapted from nei.nih.gov (National Eye Institute, National Institutes of Health)

The epithelium is composed of 4-6 layers of corneal epithelial cells which have a lifespan of 7 to 10 days before being shed with replacement cells, located underneath and provided by stem cells in the limbus [4]. The epithelium should always be covered by the tear film, which not only provides lubrication to the cornea-eyelid tribo-system but also protects the cornea from pathogens, and provides oxygen and nutrients (e.g. glucose) to the anterior layers of the cornea. Bowman's membrane is a thin acellular layer with a dense collagen meshwork and serves as an anchoring point for many of the anterior-most collagen fibrils of the stroma.

The corneal stroma accounts for up to 90% of the cornea's thickness and is mostly composed of collagen type I, with significant amounts of collagen type V, proteoglycans, and cells (mostly keratocytes). The collagen of the corneal stroma is assembled into fibrils that are much narrower than in other connective tissues in order to maintain transparency, which has been shown to be a function of fibril diameter [5]. The fibrils are heterotypic with collagen type I interweaved with small amounts of collagen type V (around 10%). This interweaving is believed to regulate the fibril diameter along with other factors, such as the electrostatic repulsion between fibrils, which is generated, in part, by chloride ions forming complexes with fibrils [6]. The corneal fibrils are aligned

in parallel into lamellae which are around 200 μm wide and 2 μm thick [7]. The lamellae can span the entire cornea, and overlap orthogonally in the central area, as illustrated in Figure 1, with lines depicting the fibrils running parallel to the cross-section and dots representing the fibrils within the lamellae running perpendicular to the cross-section. In the peripheral cornea, the lamellae are arranged annularly circum-cornea, which is believed to support the change in curvature of the cornea from the sclera [8]. At several specific regions of the corneal stroma, the lamellae are interweaved, particularly in the anterior stroma which is denser than the posterior stroma and responsible for much of the cornea's strength. Keratocytes are the main cell type present in the stroma, and are responsible for maintaining the stromal ECM. They are mostly transparent, apart from their nucleus, due to the presence of "crystallins", a type of protein that is believed to reduce the refractive index of the cell to a value closer to that of the surrounding ECM [4]. Upon injury, keratocytes can differentiate to myofibroblasts and migrate to the site of injury in order to repair the damaged tissue, however myofibroblasts contain less crystallins and hence are less transparent. At the site of injury, myofibroblasts will begin producing collagen type I and other biomolecules to heal the injury, however the collagen produced does not possess the same ordered structure of the corneal stroma and hence is less transparent. Eventually scar tissue can develop, causing a corneal opacity. Other cell types within the corneal stroma include corneal stromal stem cells, and corneal fibroblast.

Descemet's membrane is a thin acellular dense collagenous layer that supports the endothelium, which is a monolayer of corneal endothelial cells that are tightly packed into a characteristic hexagonal pattern and have a limited ability to regenerate. The corneal endothelium regulates the ionic balance and hydration of the stroma; it is therefore responsible for the electrostatic repulsion between fibrils, and prevents corneal swelling which can cause corneal opacities. Endothelial cells do not divide and multiply in vivo, hence humans are born with around 3500 corneal endothelial cells per square millimetre which gradually decreases by 0.6% per year with cells spreading to fill gaps left by dead cells [4]. Many degenerative corneal diseases, such as Fuch's dystrophy, are attributed to a poorly functioning corneal endothelium which may require surgical intervention with a corneal transplant. The cornea also has a network of nerves that extend from the sclera into the corneal epithelium and the anterior layers of the corneal stroma, and that are fundamental to the corneal blink reflex. The blink reflex serves numerous functions. It replenishes the corneal tear film which lubricates cornea-eyelid tribosystem, and sustains cells in the anterior layers of the corneal stroma and epithelium. It also clears debris and pathogens, and prevents injury. Damage to the corneal nerve network caused by laser-eye surgery, or herpetic infections can lead to lower sensitivity in the cornea resulting in a number of degenerative conditions such as dry eye, and keratitis.

Being an avascular tissue with low metabolic demand, and mostly composed of collagen (the dry weight of the corneal stroma is about 70% collagen [9]; mostly collagen type I), the cornea is a very attractive tissue for tissue engineering and regenerative medicine (TERM) treatments. The cornea has already had clinical trials for cell therapies [10] and biosynthetic corneas[11]. This is possible due to the cornea's immune-

privileged nature, which arises from the culmination of a number of factors and mechanisms such as, its avascularity, cellular and molecular barriers present in the cornea, and an immunosuppressive intraocular environment [12]. This immune-privilege means that donor corneas do not need to be matched to the recipient, and topical immunosuppressive drug treatment can be sufficient for recipients.

2 Corneal visual impairment, treatments, and surgical techniques

2.1 Blindness and visual impairment

The WHO's International Classification of Diseases (ICD-11, 2018) defines a person as being blind if they have a visual acuity below 3/60, above this, a person still presenting (with corrective lenses) a visual acuity of 6/12 or less is considered visually impaired. Visual impairment is defined in 3 levels: mild, moderate, and severe. In 2017, it was estimated by the Vision Loss Expert Group that 36 million people are blind and 216.6 million people suffer from moderate to severe visual impairment [13]. Poverty remains the most prevalent cause of blindness and visual impairment due to poor access to clean water, hygiene facilities, and primary healthcare facilities, which the WHO is addressing through the SAFE program for trachoma [14]. The WHO estimated in 2010 that 4% of visual impairments are due to corneal opacification and scarring [15]. The standard treatment for corneal opacity is a corneal transplant, also known as a corneal graft, typically performed via a penetrating keratoplasty (PK) (see sect. 2.3).

2.2 Causes of blindness and visual impairments

A large number of diseases can result in blindness and visual impairments, along with trauma. Keratoconous is one of the most prevalent corneal diseases that lead to the need for corneal transplant. It is a gradual thinning and weakening of the cornea which results in distortion of the cornea into a cone shape, and hence, severely blurred vision. Fuch's dystrophy, as mentioned earlier, is associated with a deterioration of the corneal endothelium, leading to loss its regulation of the ionic and hydration balance in the cornea. This disrupts the electrostatic repulsions between fibrils, and they become less ordered and absorb more light. Other diseases that can lead to the need for transplants include chronic inflammation of the cornea (chronic keratitis), and bullous keratopathy. Later stages of these diseases can lead to blindness due to loss of refractive power, loss of corneal transparency (corneal opacity), and/or corneal neovascularization, where blood vessels grow over the cornea.

2.3 Corneal transplants

Corneal transplants are the most common and the oldest transplant performed. Unlike other transplants, donor corneas do not need to be matched to a patient in any way

due to its immune-privilege, and they can be stored for a long time after they are explanted from a donor, typically in a national eye bank for up to 14 days. A survey carried out between 2012 and 2013 estimated that approximately 185,000 corneal transplants are performed each year in 116 different countries, and that there are 12.7 million people on corneal transplant waiting lists, but only the needs of 1 in 70 are covered [16]. Countries vary in their cornea donation rates with obvious legal, cultural, and ethical barriers impeding them. Corneal grafts are internationally distributed, with the USA, Sri Lanka and Italy being net exporters. Campaigns have been implemented in some countries to increase the number of cornea donors. One such campaign in the Brazilian state of Espirito Santo managed to eliminate the corneal transplant waiting list soon after implementation in 2011, but was soon followed by an increase in the waiting list showing the need for continued efforts to increase public awareness of the need for organ donors [17]. However, awareness campaigns alone are unlikely to solve the global donor shortage.

Penetrating Keratoplasty (PK) involves a resection of full thickness of the patient's cornea; a donor cornea is then implanted with sutures around the periphery of the graft, as seen in figure 2. However, donor shortages lead to long waiting lists and a decreased quality of life for patients. The initial success rate of corneal transplants is relatively high in developed nations, being at 85% after 2 years, but lower in developing countries like India, where initial success rates are 69% [18]. The long term success at 10 and 15 years after transplantation is instead very similar being at 62% and 55%, respectively [19], as reported in 2006 from a census of corneal transplant recipients in Australia. A more recent study performed in Spain reports data at 1, 5 and 10 years, which are very similar and specifically a 10-year success rate of 65%. The leading cause of graft failure is graft rejection; where the graft is recognised as foreign tissue and the patient's immune system responds to the graft. Immunosuppressive drugs and similar treatments are administered to allo-/xeno-graft recipients to prevent graft rejection, though in the case of a corneal graft these treatments can be tailored to the risks that the patient presents, due to the severe side effects and high mortality of patients on immunosuppressive drug treatment. The second leading cause of corneal graft failure via PK is corneal endothelial cell failure [19]. New surgical methods such as the deep anterior lamellar keratoplasty (DALK) aim to eliminate this mode of failure by leaving a patient's healthy endothelium and Descemet's Membrane intact.

Lamellar keratoplasties (LK) are a range of corneal surgical techniques that have become increasingly popular and allow for the selective transplant of parts of the cornea, so that only the pathological layers are resected [20]. LK's have facilitated new treatments that have the potential to address the donor shortage. The split-cornea transplantation has increased the number of patients that may be treated with a single donor cornea, by, as the name suggests, splitting the donor cornea so that one part may be used in a Descemet's membrane endothelial keratoplasty (DMEK) to replace a defective posterior corneal layers, and the other part may be used in a deep anterior lamellar

keratoplasty (DALK) to replace a failing anterior layers [21]. The same wealth of statistics are not available for LK surgeries due to their relatively recent invention, however, several small studies have concluded that particularly DALKs have a reduced risk of graft failure.

Keratoprotheses, or artificial corneas, have been developed but, rather than addressing the donor shortage, they are only used in patients who have experienced the rejection of a transplanted cornea or who are deemed unsuitable for a transplant. The AlphaCor was one of such keratoprosthesis made of a synthetic hydrogel which achieved clinical approval but was discontinued after severe complications, such as corneal melt and corneal neovascularization [22]. The Boston KPro is a widely used keratoprosthesis; since receiving FDA approval in 1992, it has been continually improved and it has increased the quality of life of many patients. However, it still usually requires a donor cornea, out of which a central 3mm hole is trephined to make room for a central PMMA button which is then held in place by a rear plate [23]. Hence, keratoprosthesis are not intended to address the donor shortage.

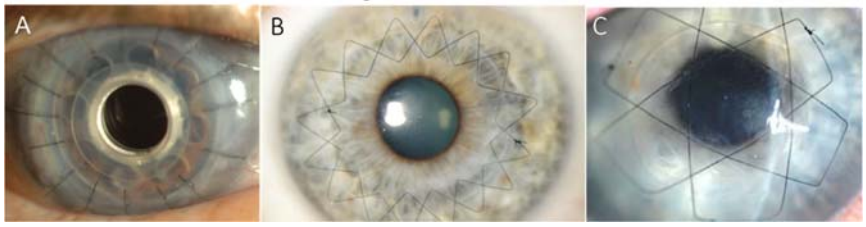


Fig. 2. Treatments for corneal dystrophies from left to right are a Boston Keratoprosthesis, a PK corneal transplant exhibiting peripheral sutures typically used in corneal transplants, and the biosynthetic stromal substitute implanted using mattress sutures, [24]© BMJ Publishing Group Ltd. , [25] © 2013 Kang et al., and [26] © ARVO Journals 2016. All reprinted with permission respectively.

2.4 Biosynthetic corneal substitute

The most promising potential treatment for corneal stroma replacement has been the biosynthetic stromal substitute (BSS) from the research group of May Griffith. The BSS is composed of recombinant human collagen type III (RHCIII) produced by FibroGen using yeast (Pichia pastoris). The BSS are produced, on ice, using a double syringe mixing system, where one syringe is loaded with highly concentrated RHCIII 10-18% and the other syringe is loaded with a concentrated MES buffer solution containing a cross-linker EDC. The low temperature, high concentration of buffer, and relatively low pH (~pH 5), is intended to form very small fibrils which absorb less light. The syringe contents are mixed in a closed system in order to prevent bubble formation and remain sterile, and is then deposited into a contact lens mould. The result is an optically clear, dense collagenous construct similar in appearance to a contact lens. Although it is not bioengineered, in that it is never cultured before implantation, the BSS is com-

posed of human collagen on which corneal epithelial cells have been shown to proliferate and keratocytes can remodel. The outcome of a 4-year clinical trial of 10 patients suffering from keratoconus or corneal scarring and treated with biosynthetic corneas via DALK, was published in 2014. It showed that no recipients of the BSS experienced rejection. Immunosuppressive drugs were administered to BSS recipients for the first 6-7 weeks while the sutures remained in place, whereas donor cornea recipients underwent much longer periods of immunosuppressive drug treatment. The company FibroGen is currently running preclinical testing of the BSS (FG-5200) in order to use them in clinical trials in China.

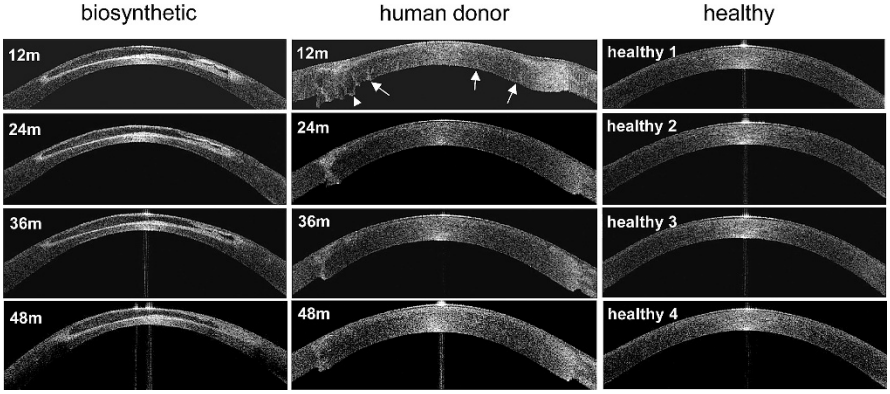


Fig. 3. Optical Coherence Tomography (OCT) of the anterior cornea at 12, 24, 36, and 48 months post implantation for a BSS implant, a PK-type human donor transplant, and a healthy cornea. The BSS can be seen to as the anterior-most clear layer with some changes in appearance between measurements due to different angles of measurement. The human donor transplant appears thicker than the other cases, which is common. Arrows indicate the locations of wound compression due to the presence of sutures [11]. © Elsevier 2014

Given that the BSS is acellular, it must rely on the capability of the cells remaining in the host cornea to populate the implant and remodel it. Bearing in mind that patients eligible for keratoplasty do not have healthy corneas, the capability of cells within their cornea to populate and remodel a biosynthetic implant may be limited. Therefore, there is a need for a bioengineered cornea with a healthy population of patient keratocytes or keratocyte-progenitors present at the time of implant.

Due to the comparatively lower mechanical properties of the BSS, these substitutes are implanted with mattress sutures, rather than peripheral sutures, as is typically done with donor corneas and in DALK surgeries. This lead to slower epithelium regeneration and some corneal thinning [27]. BSS recipients were observed to have a significantly thinner central cornea, a steeper corneal curvature and more irregularities than healthy corneas [11][26]. Therefore, a bioengineered cornea with superior mechanical strength could improve the outcome. Further work by some of the authors involved in the BSS stated the importance of shape customisation to minimise post-operative ametropia (refractive errors, long and short sightedness) in reference to biosynthetic corneas [28]. A bioengineered cornea with patient-specific shape customisation could be achieved using bioprinting technologies.

3 Bioprinting strategies for the cornea

The most likely strategy for the application of bioprinting in the treatment of corneal visual impairments is the fabrication of a stromal substitute, similar to the BSS. Such a stromal substitute should have suitable optical, geometrical and mechanical properties to be considered. It should also have sufficient glucose and oxygen permeability to sustain its population of cells and prevent neovascularisation. A suitable stromal substitute should allow nerve growth from the limbus over the cornea upon implantation to restore the blink reflex. A healthy corneal epithelium is essential to the function cornea-eyelid tribosystem, hence a bioprinted stromal substitute should exhibit full epithelial cell coverage.

The most important optical properties are transmission (the opposite of absorbance) and refractive index. As with all materials, transmission of the cornea varies by wavelength. It absorbs particularly in the lower wavelengths of visible light at 400-450 nm, where it reaches a minimum transmission of ~85%. The refractive index of the cornea is reported as 1.3375 [29]. The cornea's *in vivo* mechanical properties are normally reported in terms of corneal hysteresis (CH) and corneal resistance factor (CRH), however these are difficult to relate to an elastic modulus for material selection. *In vitro* measurements of the cornea have yielded a tensile strength of 3.8 MPa, and elastic moduli of 3-13 MPa [30]. In terms of geometry, as mentioned earlier the cornea is a, slightly oblate, hollow spheroid cap with a central cornea thickness of 550µm and anterior surface radius of curvature of 7.7 mm. Corneal transplants via DALK usually measure 10 mm in diameter and 500-550 µm in thickness, though this varies between donor and patient requirements. Therefore, a bioprinted corneal substitute, in terms of pure geometry, would be a shelled spherical cap with an anterior spherical radius of 7.7mm, a width of 10mm, a height of 1.8mm, and a shell thickness of around 500 µm (**Fout! Verwijzingsbron niet gevonden.**).

A major engineering challenge of printing a corneal substitute is the surface roughness of the printed construct. The typical method for producing the printer toolpaths, "slicing", is likely to exacerbate these problems and much of the mechanics of the structure would be dependent on the bonds between layers, which are typically the weakest part of a 3D printed structure. Laser-ablated porcine corneas, which are rougher than the untreated corneas due to ablation, have been reported to have a surface roughness of 150-250 nm RMS [31]. This level of roughness would be very difficult to achieve with any printing method without some form of post-print processing. However, the epithelial cell layer, which would likely be expected to proliferate on the anterior surface of a bioprinted corneal substitute after implantation, is capable of compensating for some level of roughness. A rougher anterior surface can also be compensated for, after implantation, using contact lenses. This method was applied to patients who received May Griffith's RHCIII biosynthetic corneas in their 4-year clinical trial [11].

3.1 Examples of bioprinted corneas

Corneal stroma substitutes have been bioprinted using extrusion-based bioprinting (EBB, explained in the An Introduction to Tissue Engineering & Bio printing chapter)

methods with limited success. In one study the construct only supported itself for one day when keratocytes were present [32].Another attempt printed a scaled-up version of the cornea with a larger diameter using a micro-extrusion bioprinter programmed to deposit material more like a drop-on-demand printer with no continuous lines, and survived for up to 7 days in culture when the experiment was ended [33]. In both cases keratocytes were deposited with high cell viability, however, the materials applied at their respective concentrations in the both inks could not be expected to reach the desired mechanical and optical properties of a stromal substitute. The print resolution in both cases was >= 200μm which inevitably led to the anterior surface of the construct appearing quite rough. Most EBB methods struggle to achieve print-lines below 200μm, which is problematic for print resolution and surface roughness. Higher print resolution is possible, though it usually depends on the EBB system being used, the material being extruded, and whether cells are present in the ink. Silk fibroin, which has similar properties to collagen, has been printed with print lines of 5μm in diameter at high concentrations 28-30% w/v, though it is not clear if this would have sufficient optical properties for a bioprinted cornea [34]. Different EBB systems also offer different levels of control in terms of extrusion parameters (e.g. extrusion pressure, print speed, nozzle diameter and shape, etc.) and environmental factors (active cooling and/or heating of nozzle and of substrate), which have a major influence over ink rheology. For example, nozzles can clog due to inconsistencies in the ink (a folded protein or undissolved particle), due to the build-up of dehydrated material at the nozzle, or, when working with thermo-responsive materials, due to different temperatures within the nozzle. Generally, EBB cell printing requires much larger nozzles to avoid clogging, and to reduce the shear forces imparted on the cells as they flow through the nozzle. However, cell seeding can be performed in vitro post print, and the work of May Griffith has also shown how an acellular scaffold can recruit keratocytes upon implantation.

To date there is one example of a light-based bioprinting (LBB, also explained in the “An Introduction to Tissue Engineering & Bio printing” chapter) corneal substitute; a stromal substitute bioprinted using digital light processing, DLP. Mahdavi et al. used methacrylated gelatin and the photoinitiator, Eosin Y, to encapsulate keratocytes. The constructs appear cornea-shaped and were robust enough to be held with a tweezers out of solution. The constructs were stable for up to 28 days in culture, however, unfortunately, the constructs did not exhibit high transmissibility [35]. DLP has been used to print methacrylated silk fibroin, which could potentially be applied to corneal bioprinting but the mechanical properties of the resultant gel would need to be improved [36]. At the time of writing this piece there were no example of 2-photon polymerisation (2PP) corneal substitutes, however given its resolution, it has potential for great impact in bioprinted corneal substitutes. The LBB methods, stereolithography, DLP, and 2PP, all have potential for bioprinting corneal substitutes and could potentially achieve higher print resolution and lower surface roughness than EBB methods

3.2 Novel bioprinting strategies for the cornea

A novel method of bioprinting, aerosol jet printing (AJP®, Optomec Inc.), is currently being investigated by the authors of this work for application in corneal substitutes. AJP was originally developed for printing electrically conductive inks for printed electronics, and despite being a complicated process that can be difficult to control, it has some advantages over the more common inkjet printing in terms of accuracy and its range of suitable inks and substrates. AJP has been investigated for the deposition of biomaterials several times in the context of functionalised surfaces and biosensors, but AJP has not been investigated for the purposes of scaffold fabrication. However, there are aspects of the AJP process and the nature of the deposited material that make it suited to some areas of biofabrication.

AJP forms an aerosol from an ink using either ultrasonic agitation or high velocity pneumatic flow. The aerosolised ink then becomes entrained in a flow of a carrier gas, usually an inert gas like nitrogen, which brings the aerosol to a nozzle where a coaxial gas flow (sheath gas) collimates the aerosol flow and prevents the aerosol from making contact with the walls of the nozzle. As the nozzle diameter decreases towards the end of the nozzle (orifice diameter of 100-300 μm), the aerosol accelerates exiting the nozzle at speed until it reaches the print substrate where the aerosol droplets coalesce. Since the aerosol is separated from the nozzle walls by the sheath gas, the resultant print lines can be as little as 10% of the nozzle diameter, i.e. 10 μm lines can be printed with a 100 μm nozzle. The flow of the carrier gas, and sheath gas are precisely controlled using mass flow controllers (MFCs,) which along with the print speed, allow the user to control the diameter of the print lines, though this often must be calibrated for each ink. Due to the inert gas in the carrier and sheath gas flows being dry, and the high surface area to volume ratio of the aerosol droplets, much of the solvent in the ink evaporates from the aerosol droplets very quickly. Solvent evaporation continues at a similar rate on the print substrate due to the small size of the print lines typically achieved with AJP, and the print bed can be heated to enhance evaporation. AJP also has the advantage that the distance between the print nozzle and the substrate (stand-off distance) can be much higher compared to other printing methods, due to the fact that the ink is accelerated toward the substrate and the sheath gas continues to collimate the aerosol for around 5 mm after exiting the nozzle. The recommended standoff distance is then 1-5 mm, which allows for printing on rough and curved surfaces.

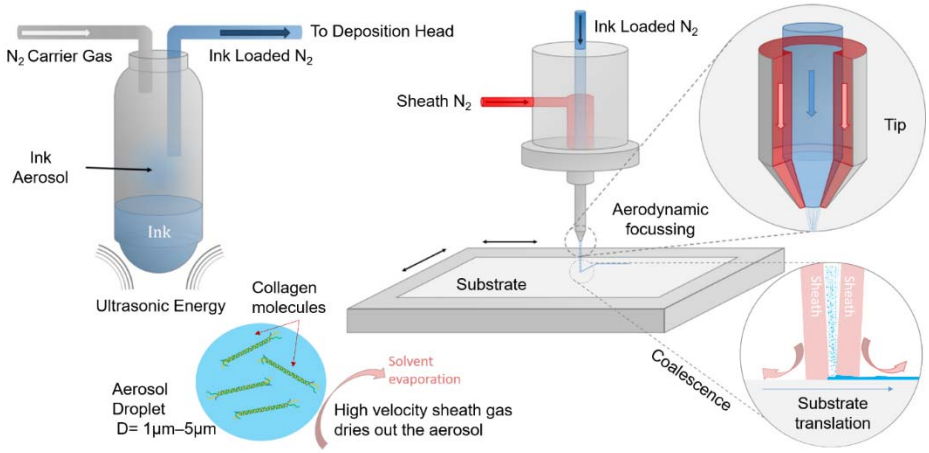


Fig. 4. Schematic of an ultrasonic aerosol jet printing system. Aerodynamic focusing collimates the aerosol into a dense stream of aerosol. Aerosol droplets of 1-5μm in diameter coalesce into a film on the substrate. Adapted with permission from [37] ©2017 IOP Publishing Ltd

The deposited material is quite different from other bioprinting methods as the print lines have a very low aspect ratio with a width of around 20 μm but a thickness of ~0.2 μm. The thickness of the print lines can change with the loading of the ink, the print speed, and by using relative gas flow rates to dry out the ink quicker so that the ink does not get the chance to flow/spread on the substrate. However, these techniques may not be possible for the ink, or the application, and may only deliver a maximum thickness of around 2 μm. Therefore, AJP is a slow process for building thick clinically relevant scaffolds for biofabrication, however, it possesses a similar z-resolution to the lamellar structure of the cornea, discussed in Section 1. The deposited collagen, although dry, has an inherent acidity from the acidic collagen solution. The printed material is stabilised in a neutralizing buffer developed to slowly neutralize the collagen and to prevent re-solubilisation and the subsequent formation of fibrils. EDC/NHS crosslinking is often performed during the neutralization (EDC, N-(3-Dimethylaminopropyl)-N′-ethyl-carbodiimide, along with NHS, N-hydroxysuccinimide). The printed scaffolds are then washed several times in phosphate buffered saline (PBS).

Cells are highly unlikely to survive this printing process, as even if they survive the mechanical stresses of atomization they would also have to survive likely dehydration, therefore AJP would only be considered as a method for bioprinting scaffolds for cell seeding. Due to the evaporation of much of the ink's solvent, the deposited material is much denser and stiffer than other bioprinting methods. This may lead to a construct that is too dense and stiff for cell culture, however, the high print resolution of AJP allows for the creation of very small pores divided by dense stiff hydrogel strands that can be designed not to exceed the oxygen diffusion limit.

We have previously reported with our collaborators at the University of Antwerp, on the AJP of collagen for corneal TE, and the use of AJP collagen films, ~87μm thick, as a substrate for keratocytes [38]. AJP has since been used to print recombinant human

collagen type III scaffolds with lowest average transparency of 89% at 405 nm, an average refractive index of 1.34625, an average elastic modulus of 0.5 MPa, and a thickness of ~500 µm. Curved collagen scaffolds may yet be produced using AJP. However, there remains much work to be done to characterise any adverse effects imparted on the collagen due to the mechanical processes involved in aerosolisation of the collagen ink.

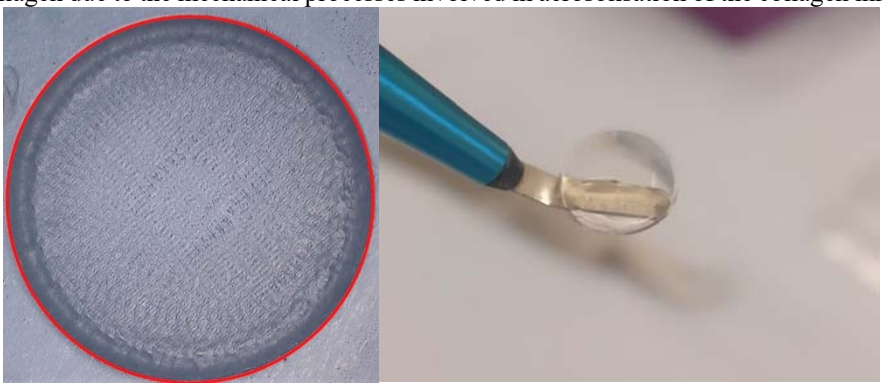


Fig. 5. A 4.5mm in diameter, and ~800 layers thick, AJP RHCIII sample imaged under a microscope immediately after printing, when dry, with print-lines of 60 µm, and a sample from the same set after the scaffold was washed in neutralizing buffer and detached from its substrate using a corneal scalpel.

4 Discussion & Conclusion

The cornea is an ideal tissue for the application of bioprinting, however the materials used in existing attempts at bioprinted corneas are mostly selected because of their printability rather than their suitability for the application. Optical and mechanical properties are unlikely to be met using the popular bioprinting hydrogels, gelatin and alginate, and collagen must gelate under specific conditions to prevent the formation of large light-absorbent fibrils. However, these existing bioprinted cornea attempt did show that keratocytes can be bioprinted and may yet find an application as a high throughput cornea model. The low resolution of AJP appears to allow collagen to be printed with high transparency, but a curved, cornea-shaped construct has yet to be produced. The BSS is an exciting prospect as biosynthetic alternative to corneal transplant, and whether successful or not the need remains for a bioengineered cornea with a healthy population of keratocytes, corneal epithelial cells, or other corneal cells. For the immediate future, the results of the awareness campaign mentioned in Espirito Santo, Brazil should inspire national health organisations to increase their efforts to spread public awareness of the need for donor corneas and how they can improve the quality of life of the recipients.

5 References

1. Gonzalez-Andrades M., Argüeso P., Gipson I.: Corneal Anatomy. , 3–12 (2019).
2. Doughty M.J., Jonuscheit S.: An assessment of regional differences in corneal thickness in normal human eyes, using the Orbscan II or ultrasound pachymetry. Optometry 78(4), 181–190 (2007).
3. Kohlhaas M., Boehm A.G., Spoerl E., Pürsten A., Grein H.J., Pillunat L.E.: Effect of central corneal thickness, corneal curvature, and axial length on applanation tonometry. Arch Ophthalmol 124(4), 471–476 (2006).
4. Delmonte D.W., Kim T.: Anatomy and physiology of the cornea. J Cataract Refract Surg 37, 588–598 (2011).
5. Hart R.W., Farrell R.A.: Light scattering in the cornea. J Opt Soc Am 59(6), 766–774 (1969).
6. Meek K.M.: Corneal collagen-its role in maintaining corneal shape and transparency. Biophys Rev 1(2), 83–93 (2009).
7. Meek K.M., Knupp C.: Corneal structure and transparency. Prog Retin Eye Res 49, 1–16 (2015).
8. Boote C., Hayes S., Young R.D., Kamma-Lorger C.S., Hocking P.M., Elsheikh A., Inglehearn C.F., Ali M., Meek K.M.: Ultrastructural changes in the retinopathy, globe enlarged (rge) chick cornea. J Struct Biol 166(2), 195–204 (2009).
9. Tseng S.C.G., Smuckler D., Stern R.: Comparison of collagen types in adult and fetal bovine corneas. J Biol Chem 257(5), 2627–2633 (1982).
10. Kinoshita S., Koizumi N., Ueno M., Okumura N., Imai K., Tanaka H., Yamamoto Y., Nakamura T., Inatomi T., Bush J., Toda M., Hagiya M., Yokota I., Teramukai S., Sotozono C., Hamuro J.: Injection of Cultured Cells with a ROCK Inhibitor for Bullous Keratopathy. N Engl J Med 378(11), 995–1003 (2018).
11. Fagerholm P., Lagali N.S., Ong J.A., Merrett K., Jackson W.B., Polarek J.W., Suuronen E.J., Liu Y., Brunette I., Griffith M.: Stable corneal regeneration four years after implantation of a cell-free recombinant human collagen scaffold. Biomaterials 35(8), 2420–2427 (2014).
12. Hori J., Yamaguchi T., Keino H., Hamrah P., Maruyama K.: Immune privilege in corneal transplantation. Prog Retin Eye Res 71 (2019).
13. Bourne R.R.A., Flaxman S.R., Braithwaite T., Cicinelli M. V, Das A., Jonas J.B., Keeffe J., Kempen J.H., Leasher J., Limburg H., Naidoo K., Pesudovs K., Resnikoff S., Silvester A., Stevens G.A., Tahhan N., Wong T.Y., Taylor H.R.: Magnitude, temporal trends, and projections of the global prevalence of blindness and distance and near vision impairment: a systematic review and meta-analysis. Lancet Glob Heal 5, 888–897 (2017).
14. Whitcher J.P., Srinivasan M., Upadhyay M.P.: Corneal blindness : a global perspective. 79(3), 214–221 (2001).
15. WHO: Global Data on Visual Impairment. Glob Data Vis Impair 2010 , 17 (2010).
16. Gain P., Jullienne R., He Z., Aldossary M., Acquart S., Cognasse F., Thuret G.: Global survey of corneal transplantation and eye banking. JAMA Ophthalmol 134(2), 167–173 (2016).
17. Freitas L., Rocon P., de Almeida A., Erlacher R., Paro F., Santo E., Pró-Vidas

Transplantes A.: Corneal donor profile and evolution of corneal donation in a Brazilian state where the number of individuals on the waiting list reached zero but increased again. Transplant Proc 50, 509–512 (2018).

18. Wright B., Griffith M., Merrett K., Connon Editors C.J.: Corneal Regenerative Medicine Methods and Protocols Methods in Molecular Biology 1014
19. Williams K.A., Esterman A.J., Bartlett C., Holland H., Hornsby N.B., Coster D.J.: How effective is penetrating corneal transplantation? Factors influencing long-term outcome in multivariate analysis. Transplantation 81(6), 896–901 (2006).
20. Hos D., Matthaei M., Bock F., Maruyama K., Notara M., Clahsen T., Hou Y., Le V.N.H., Salabarria A.C., Horstmann J., Bachmann B.O., Cursiefen C.: Immune reactions after modern lamellar (DALK, DSAEK, DMEK) versus conventional penetrating corneal transplantation. Prog. Retin. Eye Res. 73:100768 (2019).
21. Heindl L.M., Riss S., Adler W., Bucher F., Hos D., Cursiefen C.: Split cornea transplantation: relationship between storage time of split donor tissue and outcome. Ophthalmology 120(5), 899–907 (2013).
22. Salvador-Culla B., Kolovou P.: Keratoprosthesis: a review of recent advances in the field. J Funct Biomater 7(2), 13 (2016).
23. Zerbe B.L., Belin M.W., Ciolino J.B.: Results from the multicenter Boston Type 1 keratoprosthesis study. Ophthalmology 113(10), 1779–1785 (2006).
24. Wessel J.M., Bachmann B.O., Meiller R., Kruse F.E.: Fungal interface keratitis by Candida orthopsilosis following deep anterior lamellar keratoplasty. BMJ Case Rep , 1–4 (2013).
25. Kang J.J., Allemann N., Vajaranant T., de la Cruz J., Cortina M.S.: Anterior segment optical coherence tomography for the quantitative evaluation of the anterior segment following Boston keratoprosthesis. PLoS One 8(8) (2013).
26. Ong J.A., Auvinet E., Forget K.J., Lagali N., Fagerholm P., Griffith M., Meunier J., Brunette I.: 3D corneal shape after implantation of a biosynthetic corneal stromal substitute. Investig Ophthalmol Vis Sci 57(6), 2355–2365 (2016).
27. Fagerholm P., Lagali N.S., Merrett K., Jackson W.B., Munger R., Liu Y., Polarek J.W., Söderqvist M., Griffith M.: A biosynthetic alternative to human donor tissue for inducing corneal regeneration: 24-Month follow-up of a phase 1 clinical study. Sci Transl Med 2(46), 46–61 (2010).
28. Durr G.M., Auvinet E., Ong J., Meunier J., Brunette I.: Corneal shape, volume, and interocular symmetry: Parameters to optimize the design of biosynthetic corneal substitutes. Investig Ophthalmol Vis Sci 56(8), 4275–4282 (2015).
29. Merrett K., Fagerholm P., McLaughlin C.R., Dravida S., Lagali N., Shinozaki N., Watsky M.A., Munger R., Kato Y., Li F., Marmo C.J., Griffith M.: Tissue-engineered recombinant human collagen-based corneal substitutes for implantation: Performance of type I versus type III collagen. Investig Ophthalmol Vis Sci 49(9), 3887–3894 (2008).
30. Crabb R.A.B., Chau E.P., Evans M.C., Barocas V.H., Hubel A.: Biomechanical and microstructural characteristics of a collagen film-based corneal stroma equivalent. Tissue Eng 12(6), 1565–1575 (2006).
31. Lombardo M., De Santo M.P., Lombardo G., Barberi R., Serrao S.: Roughness of excimer laser ablated corneas with and without smoothing measured with atomic force microscopy. J Refract Surg 21(5), 469–475 (2005).

32. Isaacson A., Swioklo S., Connon C.J.: 3D bioprinting of a corneal stroma equivalent. Exp Eye Res 173(April), 188–193 (2018).
33. Duarte Campos D.F., Rohde M., Ross M., Anvari P., Blaeser A., Vogt M., Panfil C., Yam G.H.F., Mehta J.S., Fischer H., Walter P., Fuest M.: Corneal bioprinting utilizing collagen-based bioinks and primary human keratocytes. J Biomed Mater Res - Part A (January), 1–9 (2019).
34. Ghosh S., Parker S.T., Wang X., Kaplan D.L., Lewis J.A.: Direct-write assembly of microperiodic silk fibroin scaffolds for tissue engineering applications. Adv Funct Mater 18(13), 1883–1889 (2008).
35. Mahdavi S.S., Abdekhodaie M.J., Kumar H., Mashayekhan S., Baradaran-Rafii A., Kim K.: Stereolithography 3D Bioprinting Method for Fabrication of Human Corneal Stroma Equivalent. Ann Biomed Eng 48(7), 1955–1970 (2020).
36. Kim S.H., Yeon Y.K., Lee J.M., Chao J.R., Lee Y.J., Seo Y.B., Sultan M.T., Lee O.J., Lee J.S., Yoon S. Il, Hong I.S., Khang G., Lee S.J., Yoo J.J., Park C.H.: Precisely printable and biocompatible silk fibroin bioink for digital light processing 3D printing. Nat Commun 9(1) (2018).
37. Smith M., Choi Y.S., Boughey C., Kar-narayan S.: Controlling and assessing the quality of aerosol jet printed features for large area and flexible electronics. Flex Print Electron 2, 015004 (2017).
38. Gibney R., Matthyssen S., Patterson J., Ferraris E., Zakaria N.: The Human Cornea as a Model Tissue for Additive Biomanufacturing: A Review. Procedia CIRP 65, 56–63 (2017).

Mechanical tests and properties of living tissues and biomaterials in a biomechanics course

Sergii Podlesnij [1] [0000-0001-8271-4004], Oleksandr Hrushko [2] [0000-0001-5551-375X],

Andrii Kovalenko [3] [0000-0003-3379-2000] , Ihor Stashkevych [4] [0000-0002-7411-9835]

[1, 3, 4] Donbass State Engineering Academy, Akademichna str.,72, Kramatorsk, 84313, Ukraine
[2]Vinnytsia National Technical University, Khmelnytske shose, 95, Vinnytsia, 21021, Ukraine
[1]sergeypodlesny@gmail.com,
[2]grushko1alex@gmail.com
[3]dsea.kovalenko.andrey@gmail.com
[4]stashkevich_dgma@ukr.net

Abstract. For the development of biomechanics, biology, bioengineering, and medicine, including for the purpose of creating substitutes for tissues and organs (implants and prostheses), it is important to know the mechanical properties of living tissues and biomaterials and methods for their determination. This chapter briefly lists the types of living tissues, biomaterials and reviewed their mechanical properties and methods for their determination. Such types of biomaterials as Metallic biomaterials, Ceramics, Polymers as biomaterials, Composites, Nanomaterials and Fabric Engineering are considered.

The main types of deformations that human living tissues undergo are considered, these are stretching-compression, shear, torsion and bending. The basic concepts of normal and shear stress are revealed. Hooke's law is given and the elastic constants of an isotropic material are described – Young's modulus, shear modulus, Poisson's ratio. The classical methods of mechanical testing and determination of mechanical properties are described: tensile and compressive strength, stiffness, stability, hardness. The tensile diagram in the coordinates of deformation-stress is considered and the concepts of characteristics strength are given. The values of the ultimate compressive strength, Young's modulus and Brinell hardness of a number of living tissues of the human body are given. Some questions of tribology of joints are stated, in particular, the concepts of friction in joints, the phenomenon of metallosis and features of tribological tests of endoprostheses are disclosed. Examples of test equipment are provided.

Keywords : living tissues, biomaterials, implants, mechanical tests, material properties, types of deformations, testing machines

Introduction

The main feature of modern society is the growing interest in increasing the quality and duration of human life. Achieving this goal involves, in particular, the creation of materials for artificial organs and tissues. Over the past 30 years, more than 40 different

materials (ceramics, metals, polymers, composites, etc.) have been used to treat, restore, and replace more than 40 different parts of the human body, including skin, muscle tissue, blood vessels, nerve fibers, and bone tissue. The field of biomaterials is a pronounced interdisciplinary field in which the achievements of chemistry, physics, medicine, biotechnology, metallurgy, and electronics are used.

According to the decision of the congress of the European Society for Biomaterials (European Society for Biomaterials), adopted back in 1987, biomaterials should be called "non-living materials used in medical devices and designed to interact with living systems."

In 2012, the authoritative IUPAC gave a slightly broader interpretation of the term – the International Union of Chemists recommended calling them “materials used in contact with living tissues”.

When creating and using biomaterials, it is important to know their properties in comparison with the properties of replaceable living tissues. The topic itself is very extensive and we will not consider all its aspects, but restrict ourselves to considering the mechanical properties and methods for their determination.

1 Types of living tissues

To begin with, very briefly recall the main types of living human tissues. In total, four types of tissues are distinguished [1]: epithelial, connective (including blood and lymph), muscle and nervous.

Epithelial (integumentary) tissue, or epithelium, is a boundary layer of cells that lines the integument of the body, mucous membranes of all internal organs and cavities, and also forms the basis of many glands. Epithelium includes the surface of the skin, mouth, oesophagus, alveoli, respiratory tract, etc.

The most interesting for us is the connective tissue, which consists of cells, intercellular substance, and connective tissue fibers. It consists of bones, cartilage, tendons, ligaments, blood, fat, it is in all organs (loose connective tissue) in the form of the so-called stroma (frame) of organs.

In contrast to epithelial tissue, in all types of connective tissue (except adipose), the intercellular substance predominates over the cells in volume, i.e., the intercellular substance is very well expressed. The chemical composition and physical properties of the intercellular substance are very diverse in various types of connective tissue.

In dense fibrous connective tissue (tendons of muscles, ligaments of joints) fibrous structures predominate, it experiences significant mechanical stress.

The bone tissue that forms the bones of the skeleton is very durable. It maintains the body shape (constitution) and protects organs located in the cranium, chest and pelvic cavities, participates in mineral metabolism. The tissue consists of cells (osteocytes) and the intercellular substance, in which the nutrient channels with blood vessels are located. The intercellular substance contains up to 70% of mineral salts (calcium, phosphorus, and magnesium).

Cartilage tissue consists of cells (chondrocytes) and intercellular substance (cartilage matrix), characterized by increased elasticity. It performs a supporting function, as it forms the bulk of the cartilage.

Nerve tissue consists of two types of cells: nerve (neurons) and glial.

Two types of muscle tissue are distinguished: smooth (unstriated) and striated (striated).

2 The main types of biomaterials.

2.1 Key Terms

A biomaterial is a non-viable material designed to come into contact with living tissue to perform biomedical functions. The biomaterial must be biocompatible and may be biodegradable.

An implant is a biomaterial product that performs (replaces) a certain structural and functional load of living tissue or organ.

Biodegradation is the process of decomposition of non-viable materials in contact with living tissues, cells and biological fluids.

Biocompatible materials and devices – act or function harmoniously and in concert when in the body or contact with body fluids, without causing diseases or painful reactions.

Bioactive materials (BAM) – materials designed to bind them to biological systems to increase the effectiveness of treatment, formation or replacement of any tissue, organ when performing certain body functions. Currently, there are 5 main categories among the BAM family: 1) calcium phosphate ceramics; 2) glass and glass-ceramics; 3) bioactive polymers; 4) bioactive gels; 5) composites.

Bioinert materials – the conditional release of inanimate materials that interact weakly with biological structures and liquids.

Medical materials include:

- non-viable polymer, ceramic, metal, carbon, textile, glass and other materials and their composites;
- biological materials (transplants);
- rubbers, latexes, papers, dyes, varnishes, enamels, adhesives and other materials used for the manufacturing of medical devices.

2.2 Types of Biomaterials [2]

There are several important areas for the use of biomaterials:

- materials used to create implants and endoprostheses used in cardiovascular, bone surgery, ophthalmology, dental technology, when replacing soft tissues, etc.;
- materials used to create systems with biological activity;
- materials used in bioengineering technologies (cellular, tissue and gene);
- materials used to create separation systems;

- materials for biochemical methods of analysis and synthesis;
- materials used to obtain systems with enzymatic activity;
- materials used to create products that are not in direct contact with blood and lymph, such as contact lenses and devices for external osteosynthesis;
- biodegradable materials.

Metallic biomaterials

Currently, over 500 alloys are used in dentistry and surgery. According to International standards (ISO, 1989) all metal alloys are divided into the following groups:

- alloys of precious metals based on gold;
- alloys of precious metals containing 25-50% of gold or platinum, or other precious metals;
- alloys of base metals;
- alloys for cermet structures:
  - with a high gold content (> 75%);
  - with a high content of noble metals (gold and platinum or gold and palladium >75%);
  - based on palladium (more than 50%);
  - based on base metals: based on cobalt (+ chromium > 25%, molybdenum > 2%);
  - nickel-based (+ chromium > 11%, molybdenum > 2%).

Base metal alloys include: chromium-nickel (stainless) steel; cobalt-chrome alloy; nickel-chromium alloy; cobalt-chromium-molybdenum alloy; titanium alloys; alloys of aluminium and bronze.

Ceramics

Bioceramics should have certain chemical properties (the absence of undesirable chemical reactions with tissues and interstitial fluids, the absence of corrosion), mechanical characteristics (strength, crack resistance, delayed fracture resistance, wear resistance), biological properties (lack of reactions from the immune system, fusion with bone tissue, stimulation of osteosynthesis).

By the nature of the body's response to the implant, biomaterials are classified as follows: 1) toxic (if the surrounding tissue dies upon contact) - most metals; 2) bioinert (non-toxic, but biologically inactive) - ceramics based on Al_2O_3, ZrO_2; 3) bioactive (non-toxic, biologically active, fused with bone tissue) - composite materials such as biopolymer / calcium phosphate, ceramics based on calcium phosphates, bio-glass.

Polymers as biomaterials

Polymeric biomaterials are often obtained by targeted modification of well-known polymers based on silicone rubber, polyacrylamide, polysiloxanes, polyester urethane, polymer-ceramic materials, etc.

Composites, Nanomaterials and Fabric Engineering

Composites are construction materials, they are constructed from two, three or more materials with different properties and reliably connected to each other to ensure the operation of the resulting system as a whole.

Nanomaterials can be divided into three main groups: raw materials; nanostructured materials; nanotubes and fullerene.

Biocomposites consist of a biodegradable polymer as a matrix and, usually, biofibers as reinforcing elements. Natural fibers are divided into plants, animals, and minerals. Typically, plant fibers are lignocellulosic and are composed of cellulose, hemicellulose, and lignin (flax, jute, sisal, and kenaf). Natural fibers of animal origin are made up of protein, such as wool, cobwebs, and silkworm silk.

Tissue engineering is an area of research whose goal is to replace or restore the anatomical structure and functions of damaged or missing tissue. One of the most widely studied aspects of tissue engineering is the design of a polymer framework (matrix) of elements with special mechanical and biological properties similar to those of an extracellular matrix. The success of the use of matrices in tissue engineering is determined by the properties of the material created by solving the problems of optimizing mechanical, chemical and geometric properties for effective functioning at macroscopic and microscopic levels.

3 Mechanical tests and material properties

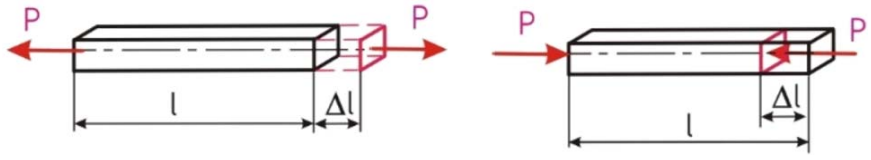
3.1 The main types of deformations.

Under the action of the loads, the bodies are deformed, i.e., their shape and size can vary [3]. Deformations are elastic, that is, those that disappear after the cessation of the forces that caused them, and plastic, or residual, that do not disappear.

Deformations can be very complex, but these complex deformations can always be represented as consisting of a small number of basic types of deformations.

The main types of deformations are:

The extension (Fig. 1, a) or compression (Fig. 1, b). Stretching or compression occurs, for example, in the case when oppositely directed forces are applied to the rod along its axis.



a b

Fig. 1. Extension (a) or compression (b).

The change Δl in the initial length l of the rod is called absolute elongation in tension and absolute shortening in compression. The ratio of absolute elongation (shortening)

Δl to the initial length l of the rod is called relative elongation along the length l and denote ε

$$\varepsilon = \frac{\Delta l}{l} \tag{1}$$

Shear (Fig. 2). Shear occurs when external forces displace two parallel planar sections of the rod relative to one another at a constant distance between them.

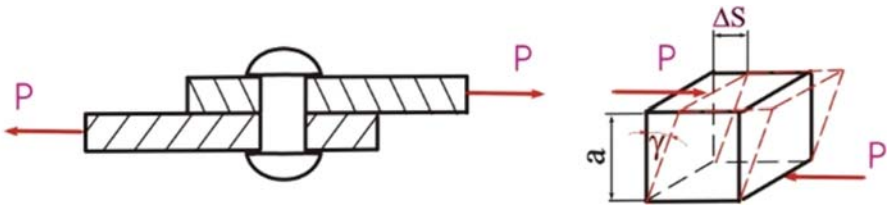


Fig. 2. Shear

The magnitude of the offset Δs is called the absolute offset. The ratio of the absolute shift to the distance a between the shifting planes is called the relative shift. Due to the smallness of the angle γ during elastic deformations, its tangent is taken equal to the skew angle of the element under consideration. Therefore, the relative shift

$$\gamma = \frac{\Delta s}{a} \tag{2}$$

Torsion (Fig. 3). Torsion occurs when external forces act on the rod, forming a moment relative to the axis of the rod.

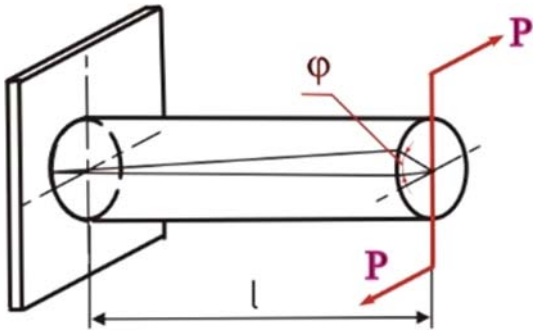


Fig. 3. Torsion

Torsion deformation is accompanied by a rotation of the cross-sections of the rod relative to each other around its axis. The angle of rotation of one section of the rod relative

to another, located at a distance l , is called the angle of twist on the length l . The ratio of the twist angle φ to the length l is called the relative twist angle:

$$\theta = \frac{\varphi}{l} \tag{3}$$

Bending (Fig. 4). Bending deformation consists in curving the axis of a straight bar or in changing the curvature of a curved bar.

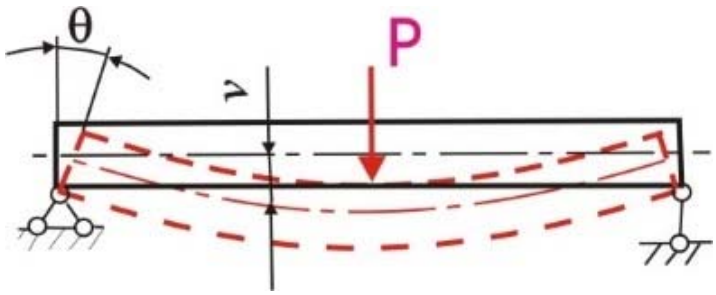


Fig. 4. Bending

3.2 Stresses. Hooke's Law.

The value of σ, which characterizes the intensity of the distribution of internal forces over the cross-section, is called normal stress:

$$\sigma = \frac{N}{S} \tag{4}$$

where N is the force acting along the axis, S is the cross-sectional area.

The force Q perpendicular to the axis leads to shear stresses:

$$\tau = \frac{Q}{S} \tag{5}$$

The stress according to the International System of Units is measured in pascals (Pa; 1Pa=N/m2), in practice, megapascals are often used (1 MPa=1 N/mm2).

Hooke's Law. In 1660, the English scientist R. Hook formulated a law that defines a linear relationship between stress and strain:

for stretching (compression)

$$\sigma = E\varepsilon \tag{6}$$

for shift

$$\tau = G\gamma \tag{7}$$

where E is the elastic modulus of the first kind (Young's modulus (Table 1) or longitudinal modulus), G is the shear modulus [4]. For most materials, G = 0.4E.

The ratio of the relative transverse compression δ to the relative longitudinal tension ε is determined through the Poisson's ratio ν by the dependence:

$$\delta = -\nu\,\varepsilon. \tag{8}$$

Table 1. Elastic modulus (Young's modulus) of some materials

Material	Young's modulus E, Pa
Elastin	10^5-10^6
Collagen	10^7-10^8
Erythrocyte membrane	$4\cdot10^7$
Smooth muscle cells	10^4
Muscle alone	$9\cdot10^5$
Bone	$2\cdot10^9$
Tendon	$1.6\cdot10^8$
Nerve	$18.5\cdot10^6$
Vein	$8.5\cdot10^5$
Artery	$5\cdot10^4$
Wood	$12\cdot10^9$
Rubber	$5\cdot10^6$
Steel	$2\cdot10^{11}$
Glass	$7\cdot10^{10}$

However, many biomaterials react as hyperelastic materials, so Hooke's law does not comply. The full version of Hooke's law should be used. Hooke's Law can be applied to isotropic material undergoing three dimensional stress (Fig. 5.).

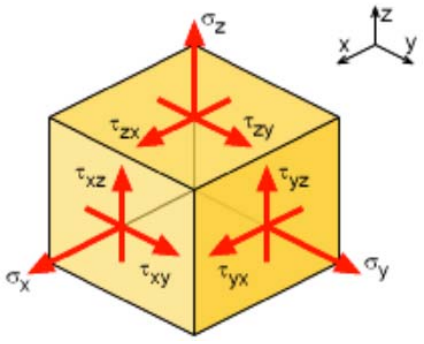


Fig. 5. Stress Directions in 3D ($\tau_{xy} = \tau_{yx}$, $\tau_{yz} = \tau_{zy}$, $\tau_{xz} = \tau_{zx}$)

The final equations are summarized in the table 2 below.

Table 2. Elastic modulus (Young's modulus) of some materials

3D Hooke's Law (Stress-Strain Relationship)	
Compliance Format	**Stiffness Format**
$\varepsilon_x = \frac{1}{E}\left[\sigma_x - \nu\left(\sigma_y + \sigma_z\right)\right]$	$\sigma_x = \frac{E}{(1+\nu)(1-2\nu)}\left[(1-\nu)\varepsilon_x + \nu(\varepsilon_y + \varepsilon_z)\right]$
$\varepsilon_y = \frac{1}{E}\left[\sigma_y - \nu\left(\sigma_z + \sigma_x\right)\right]$	$\sigma_y = \frac{E}{(1+\nu)(1-2\nu)}\left[(1-\nu)\varepsilon_y + \nu(\varepsilon_z + \varepsilon_x)\right]$
$\varepsilon_z = \frac{1}{E}\left[\sigma_z - \nu\left(\sigma_x + \sigma_y\right)\right]$	$\sigma_z = \frac{E}{(1+\nu)(1-2\nu)}\left[(1-\nu)\varepsilon_z + \nu(\varepsilon_x + \varepsilon_y)\right]$
$\gamma_{xy} = \frac{\tau_{xy}}{G};\ \gamma_{yz} = \frac{\tau_{yz}}{G};\ \gamma_{xz} = \frac{\tau_{xz}}{G}$	$\tau_{xy} = G\gamma_{xy};\ \tau_{yz} = G\gamma_{yz};\ \tau_{xz} = G\gamma_{xz}$

In continuum mechanics, the finite strain theory — also called large strain theory, or large deformation theory — deals with deformations in which strains and/or rotations are large enough to invalidate assumptions inherent in infinitesimal strain theory. In this case, the undeformed and deformed configurations of the continuum are significantly different, requiring a clear distinction between them. This is commonly the case with elastomers, plastically-deforming materials and other fluids and biological soft tissue [5].

Ronald Rivlin and Melvin Mooney developed the first hyperelastic models, the Neo-Hookean and Mooney–Rivlin solids. Many other hyperelastic models have since been developed. Other widely used hyperelastic material models include the Ogden model and the Arruda–Boyce model. A hyperelastic or Green elastic material is a type of constitutive model for ideally elastic material for which the stress–strain relationship derives from a strain energy density function. The hyperelastic material is a special case of a Cauchy elastic material [6, 7].

3.3 The study of the mechanical properties of biological tissues and implants

All structures of the human body and implants have the structural design necessary to fulfill the functional purpose and protection from external influences.

The change in the volume of solid implants is taken into account only with volume compression. In other cases, they suggest a constant volume. For example, the elongation of a structural element under tension leads to a decrease in its cross-section [8].

A deformation is considered elastic if it disappears after eliminating the force causing it. Otherwise, the deformation is called residual. Elasticity is the property of the body to exhibit elastic deformation. Plasticity is the property of the body to exhibit permanent deformation. The greater the elongation of the body at the break, the more ductile the material. The opposite of ductility is fragility. Fragility – the ability of a material to fail without the formation of noticeable residual deformations. For brittle

materials, the elongation at break does not exceed 2-5% and is sometimes measured in fractions of a percent.

If the forces acting on a solid body are large enough, then, after passing a certain stage of deformation, the body loses its integrity – breaks up into separate parts – collapses. Destruction can last from fractions of a second to many years in the case of "long-term destruction". The property of a solid to resist fracture is called strength.

Rigidity - the body's ability to resist the formation of deformations.

Stability - its ability to withstand influences seeking to bring it out of its original state of equilibrium.

All solids to one degree or another have the properties of strength and stiffness, that is, they are capable of perceiving, within certain limits, the influence of internal forces without breaking down and without substantially changing their geometric dimensions.

To study the mechanical properties of materials, many different types of tests have been developed. The main and most common tests, in which it is possible to obtain the most important mechanical characteristics of the material, are tensile and compression tests.

A general view of the tensile testing machine Smart Tensor UIT STM 010 is shown in Fig. 6. The machine has a maximum breaking strength of 1 ton and was used for testing.

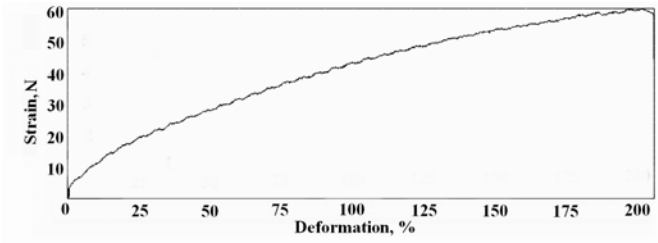
Testing of medical samples is also performed on static and dynamic testing machines. Natural biomaterials are subjected to multiaxial loads, which can be applied and measured using suitable testing machines. Static testing machines are specially designed for tensile, compression, bending, shear and torsion tests, brilliantly solving complex problems in the field of testing materials and finished products. These include: universal testing machines; fatigue testing systems / creep testing machines; torsion testing machines, biaxial testing machines, deep drawing testing machines [9].

Recently, robotic and automated testing systems have been used to efficiently perform many important tests.

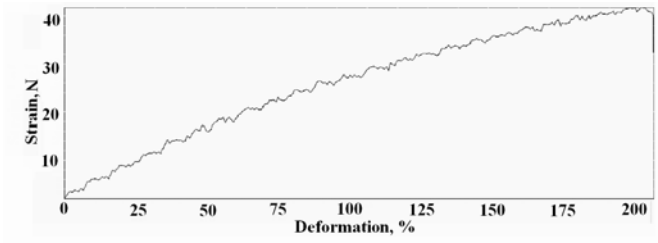
Medical specimens were cut from a blood transfusion bottle, a rubber transfusion tube and a blood storage container.

Fig. 6. Laboratory testing of tissues and biomaterials using Smart Tensor UIT STM 010 and sample from a rubber-plastic tube

A rubber transfusion tube sample with an outer diameter of 3.8 mm and a wall thickness of 0.15 mm had a length of 100 mm, the distance between the jaws of the testing machine was 60 mm. Tensile elongation to rupture was 196.27%. A sample from a blood storage container had dimensions: length 110 mm, width 11.5 mm, thickness 0.5 mm. Tensile elongation to rupture was 184.38%. The tensile diagrams of the samples are shown in Fig. 7.



a



b

Fig. 7. Tensile diagrams of a - samples fragment of a rubber-plastic tube for blood transfusion; b - sample fragment of a container for storing blood

Ultimate tensile strength (UTS), often shortened to tensile strength (TS), ultimate strength, or Ftu within equations, is the maximum stress that a material can withstand while being stretched or pulled before breaking. In brittle materials, the ultimate tensile strength is close to the yield point, whereas in ductile materials the ultimate tensile strength can be higher.

The ultimate tensile strength is usually found by performing a tensile test and recording the engineering stress versus strain. The highest point of the stress-strain curve is the ultimate tensile strength and has units of stress.

Tensile strengths are rarely used in the design of ductile members, but they are important in brittle members. They are tabulated for common materials such as alloys, composite materials, ceramics, plastics, and wood.

In the Table 3 shows the strength characteristics of various tissues [4].

The viscosity of solids is the property of solids to irreversibly absorb energy during plastic deformation.

The behaviour of real bodies can be modelled using various combinations of viscous and elastic elements.

The change in strain overtime at constant load is called creep.

The aftereffect of the material can also manifest itself in the form of stress relaxation. If in the test sample to create some deformation and leave it unchanged in time, then as a result of plastic flow, the stress will decrease. This is due to the fact that plastic deformation increases over time. Since the total strain, consisting of elastic and plastic deformation, is constant, an increase in plastic deformation leads to a decrease in elastic deformation and, consequently, to a decrease in stresses in the sample material. Stress reduction over time with constant deformation - stress relaxation.

Table 3. Strength characteristics of various fabrics

Type of fabric	The limit of compressive strength, MPa
Solid bone	147
Mineral component	44
Protein component	0.1
Enamel	34—45
Dentine	20
Edge	1—4
Vertebra	7
Spongy femur	68
Ligaments of large joints	10—16
Skin (belly)	17—36

Fibrous materials are capable of absorbing greater loads under tension than under compression.

A great influence on the manifestation of the properties of ductility and brittleness has a loading time. With rapid loading, the property of brittleness manifests itself more sharply, and with prolonged exposure to the load, the property of ductility.

During the destruction, external forces lead to micro- and macrodamage in the material. The process of damage accumulation is divided into two stages: structural changes at the micro-level, which are scattered in large quantities throughout the structural element and the development of structural damage to cracks. A crack is a crack with a variable surface.

The crack nucleation and its growth lead to a change in the structural qualities of the deformable body and may destroy the body.

Below, for example, the lesions characteristic of long tubular bones are considered. The destruction of such bones can be considered as the destruction of the rod when exposed to loads in the longitudinal or transverse directions.

Longitudinal loads (compression) occur, for example, when an outstretched arm falls on a hand, on a hand bent at the elbow joint or on a bent knee. In sports practice, bones are often damaged due to their bending under the influence of external influences. The zone of the onset of the destruction of the diaphysis of the long tubular bone during bending is located on the convex side (Fig. 8) of the arc, where the highest tensile stress values are concentrated [10].

Dynamic tests are tests in which the load is applied by impact at high speed. The rate of load application in these tests is significantly higher than in static tests.

To assess the resistance of materials to dynamic loads, shock bending tests are carried out using a pendulum ram.

In the process of fatigue tests, the fatigue limit is determined - the greatest stress at which the sample does not break at a given number of load changes.

Hardness is the property of a material to resist contact deformation or brittle fracture when an indenter is embedded in a surface.

Static hardness testing methods for indenter indentation have received the greatest application in technology: the Brinell method, the Vickers method, and the Rockwell method. In the Brinell hardness test, a ball of diameter *D* is pressed into the surface of the material under the influence of load P, and after the load is removed, the indentation diameter d is measured. The Brinell hardness number *HB* is calculated as the ratio of the load *P* to the surface area of the spherical imprint *M*:

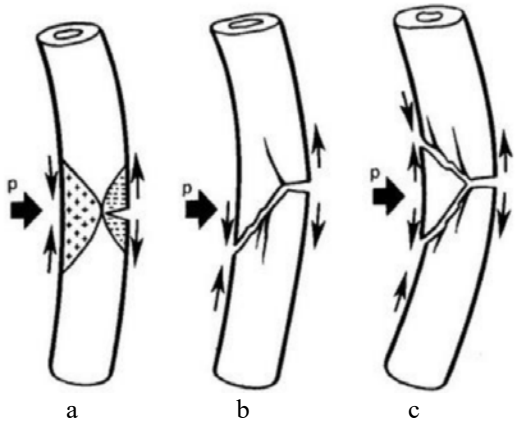


Fig. 8. Mechanisms of fractures of long tubular bones a - power distribution of stresses in the moment of formation of fracture;b - formation of a membrane-free fracture;c - fragmentation fracture.

$$HB = \frac{P}{M} \tag{9}$$

In the Table 4 shows the values of hardness for tissues of the jaw bones and teeth [4].

Table 4. Brinell hardness for tissues of the jaw bones and teeth

Plot	Test tissue	Hardness HB, 10^4 Nm^{-2}
Upper jaw (side section)	Compact substance	444
	Sponge trabeculae	452
Lower jaw	Compact substance	458

(side section)	Sponge trabeculae	457
Enamel	Incisors, fangs, premolars, molars	3776
Dentine	Incisors, fangs	726

Some examples of mechanical tests.
Load test of the human femur with installed strain gauges (DMS). The purpose of the test (Fig. 9) was to determine the extent to which the implanted endoprosthesis strengthens the bone, creating the so-called load protection effect. To do this, the human femur was installed in a 20 kN Zwick AllroundLine test machine. Horizontal mounting with a spherical cushion was supposed to exclude the impact of transverse forces arising. Then the femoral head was subjected to axial loading. DMS were glued onto the surface of the bone to compare the tension of the surface of the femur before and after implantation of the prosthesis [9].

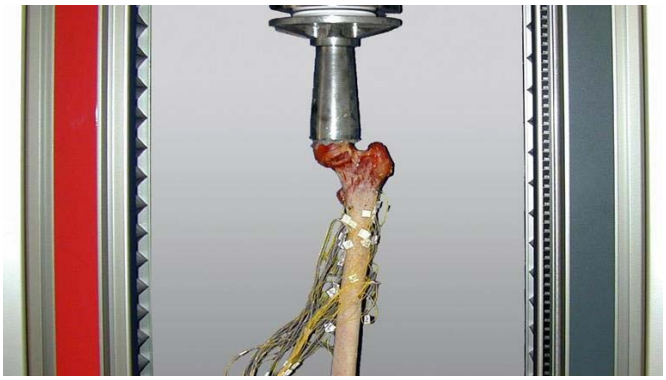


Fig. 9. Test under the load of the human femur [9]

Bent sheep test (Fig. 10). To determine to bend stiffness after fracture treatment, a 3-point bend test of sheep bones was developed. For this, the ends of the bone that have healed after a fracture are fixed in the fixtures (or poured). The load is applied using a 0.5 kN zwicki Line bench test machine. The capture is designed so that it is possible to rotate the bone by a certain angular degree. Due to this, it is possible to determine the stiffness in bending around the entire circumference of the bone. The resulting characteristics are used in FEM modelling of bone fracture healing [9].

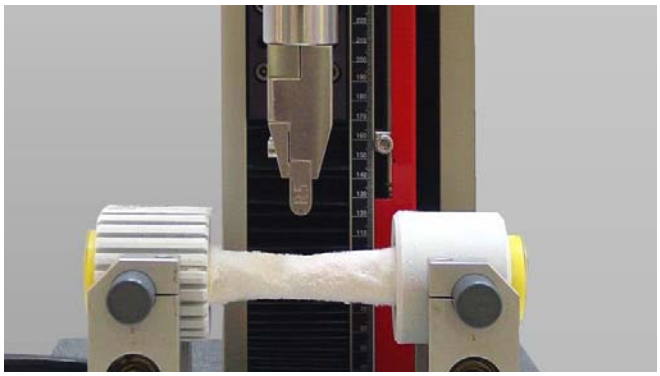


Fig. 10. Sheep bending test [9]

Comprehensive mechanical tests are widely used to create advanced materials. In recent years, new natural and synthetic materials have been discovered for use in tissue engineering. Many researchers have demonstrated the possibility of using nanostructured materials, such as nanoapatite, nanoclay and nanofibers (based on polymers or carbon nanotubes), to increase the mechanical strength and thermal stability of biopolymers from which artificial joints are made. Tissue engineering is extremely actively using the achievements of modern science to recreate and initiate the growth of living tissue [11].

3.4 Friction and lubrication in the joints. Tribology.

Friction is important for joint function. Of great importance are developments related to the creation of artificial human organs, in which the problems of tribology play a paramount role [2].

It should be noted that the skeleton of the human body is formed by separate mobile and motionless bones, the total number of which is more than 200, and 148 of them are mobile. From a mechanical point of view, the joint can be represented (modelled) by some "sliding bearing", which is lubricated with synovial fluid, and the role of contact surfaces in it is played by hyaline cartilage.

When measuring the quantitative parameters of friction in joints, an experimental method is used, called the "physical pendulum method". The bone link connected by the appropriate joint is allowed to perform free vibrations like a pendulum. Then, according to nature and decay time of bone link oscillations, the parameters of friction in the joint are quantified. On average, the coefficient of friction in the joints is a value that takes on a value from 0.003 to 0.02, which primarily depends on the specific connection of the bone link in the joint and on the physiological characteristics of the person. It should be noted that intense movement increases the amount of synovial fluid in the joints, and therefore reduces inter-articular friction and improves mobility.

As for artificial joints, this is a rather complex device, which also contains lubricating fluid, as in natural joints (Fig. 11) [12].

Fig. 8. Knee endoprosthesis [12]

In the era of endoprosthetics, the previously unknown term metallosis appeared, which means the intensive clogging of soft tissues with endoprosthesis wear products. Metallosis and endoprosthesis instability have become synonymous, and aseptic inflammation and metallosis have become inseparable companions. Therefore, the study of friction pairs of endoprostheses is especially relevant. For this, in addition to the pendulum method mentioned above, other methods are also used. Tribological tests of friction pairs of materials can be carried out according to the scheme of rotational friction with contact geometry of the "ball-on-disk" type.

Currently, there is an acute shortage of techniques to objectively assess the quality level of endoprostheses. With a large abundance of various types of endoprostheses, the absence of such information leads to the penetration of poor-quality products into clinical practice and subjectivity in the choice of endoprostheses by surgeons.

The availability of objective technical information on the work of the joint endoprosthesis in the body will allow surgical operators and patients to get an idea about the future fate of the operated joint.

Therefore, the interstate standard “Implants for surgery, replacement of a joint with a total endoprosthesis, was developed and used since 2015 [13]. The standard allows you to quickly and completely objectively determine the quality of the friction unit, interpret the cause of implant failure during the examination of negative cases of endoprosthesis use, and compare the performance of various endoprostheses during product certification. The standard gives specialists a technique for evaluating friction nodes, which allows for the development or selection of hip joint endoprostheses that guarantee their operation for the first 10 years after implantation.

As the equipment used is a testing machine capable of generating rotational movements and detecting the torque arising from friction during rotation of the head in the cup during the test. The force sensor provides a measurement of the axial load of the force in the range from 0 to 5 kN with an accuracy of ± 1%. The torque sensor provides registration of the torque force in the range from 0 to 10 Nm with an accuracy of ± 1%. The recording device provides a record of information obtained as a result of the test with an accuracy of ± 1%.

The layout of the test samples in the test machine is shown in Figure 12.

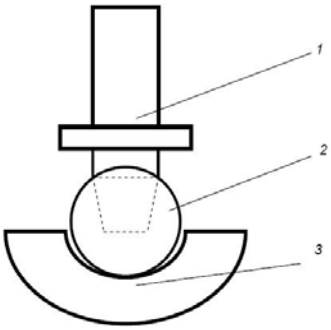


Fig. 9. - Installation diagram of the head and bowl of the endoprosthesis for testing:1 - a device for installing the head; 2 - the head of the endoprosthesis; 3 - a bowl.

The friction assembly is assembled following Figure 12 and tests are performed. Provide loading of the friction unit with a force of 2250 N at a speed of 1 mm/min. To test the friction units made using polymeric materials, it is recommended to preload the friction unit with a force of 3000 N at a speed of 1 mm/min, and then start the torsion test at the moment when the value of the force due to polymer relaxation decreases to 2250 N. Set the rotation speed bowl, equal to 0.5 r/s, and with an axial load of 2250 N carry out a graphic registration of torque for 600 s, during which the bowl makes 300 full revolutions. Friction units of hip joint endoprostheses, in which the torque does not exceed 1.5 Nm and in which wear products are not found, can provide the durability of at least 10 years. Friction units of hip arthroplasty, in which the torque exceeds 1.5 Nm and in which wear products are found, are not allowed for clinical use.

Conclusions

The study of methods for determining the mechanical properties of living tissues and biomaterials is an important part of the biomechanics course. Mechanical tests are carried out using tradional methods of theoretical and applied mechanics. The study aims to determine the fundamental physical characteristics and properties of living tissues, structures and biomaterials. The research results are used for the development of biomechanics, biology, bioengineering and medicine, including to create substitutes for tissues and organs (implants and prostheses). On the other hand, many characteristics of the musculoskeletal system are used in the design of other technical systems, which is the subject of bionics. The knowledge gained is also used to create mathematical models of biomechanical objects.

References

1. Human Tissue Ethical and Legal Issues. Nuffield Council on Bioethics 28 Bedford Square London WC1B 3EG. ISBN 0 9522701 1 0 April 1995. – 182p.

2. Larry L. Hench and Julian R. Jones. Biomaterials, artificial organs and tissue engineering. Woodhead Publishing.Cambridge England. – 2005. – 304 p.
3. Spravochnik po soprotivleniyu materialov/ Pisarenko G.S., Yakovlev A.P., Matveev V.V.; Otv. Red. Pisarenko G.S. – Kiev: Nauk. Dumka, 1988. – 736 s.
4. Biomekhanika: Ucheb. dlya sred, i vy`ssh. ucheb, zavedenij. — M.: Izd-vo VLADOS-PRESS, 2003. — 672 s.
5. Dimitrienko, Yuriy (2011). Nonlinear Continuum Mechanics and Large Inelastic Deformations. Germany: Springer. ISBN 978-94-007-0033-8.
6. Gao, H; Ma, X; Qi, N; Berry, C; Griffith, BE; Luo, X. "A finite strain nonlinear human mitral valve model with fluid-structure interaction".Int J Numer Method Biomed Eng. 2014 Dec; 30(12): 1597–1613. Published online 2014 Nov 26. doi: 10.1002/cnm.2691
7. Naumann, C.; Ihlemann, J. (2015). "On the thermodynamics of pseudo-elastic material models which reproduce the Mullins effect". International Journal of Solids and Structures. 69–70: 360–369. doi:10.1016/j.ijsolstr.2015.05.014.
8. Begun P.I. Modelirovanie v biomekhanike: Uchebnoe posobie/ P.I.Begun, P.N.Afonin. – M.: Vy`ssh. Shk., 2004. – 390 s.
9. Biaxial and Triaxial Testing of Biomaterials, https://www.zwickroell.com/en/medical/biomaterials-clinical-research/biaxial-triaxial-tests, last accessed 2020/09/01
10. Ioannis 0. Skagias and Theodoros B. Grivas. Femoral distal-end fractures treatment using the Ilizarov circular frame. Actaortopedica et traumatologicalhellenica. volume 68 i issue 1&2 i january - june 2017. P. 15-23.
11. Nano- i biokompozity` [E`lektronny`j resurs] /pod red. A. K.-T. Lau, F. Khussejn, Kh. Lafdi; per. s angl. — E`l. izd. — E`lektron. tekstovy`e dan. (1 fajl pdf: 393 s.). — M.: BINOM. Laboratoriyaznanij, 2015.
12. Total Knee Joint Replacement Revision Surgery, https://www.physio-pedia.com/Total_Knee_Joint_Replacement_Revision_Surgery, last accessed 2020/09/01
13. GOST 31621-2012 Mezhgosudarstvenny`j standart. Implantaty` dlya khi-rurgii zameshhenie sustava total`ny`m e`ndoprotezom. Opredelenie dolgovechnosti raboty` uzla treniya e`ndoproteza tazobedrennogo sustava metodom oczenki krutyashhego momenta. (Implants for surgery. Total joint replacement. Determination of durability of friction unit's work of hip endoprostheses by a method of torque estimation.) MKC 11.040.40. Data vvedeniya 2015-01-01.

Mechanical tests and computer models for the evaluation of soft tissue parameters used in implants design

Aneta Liber-Kneć[1][0000-0001-8727-1361] and Sylwia Łagan[2][0000-0001-5734-9615]

[1,2]Cracow University of Technology, Warszawska 24, 31-155 Cracow, Poland
aliber@pk.edu.pl

Abstract. In aim to good fitting biomaterials to implants, some major requirements should be met. These are the biocompatibility, bio-inertion, bio-reactivity, non-pirogenity and non-cancerogenity. Some of these features have a biological character but most of them have a physicochemical basis. In this chapter, the mechanical properties of tissues and biomaterials will be described in terms of using them in computer modeling, which is often applied to predict their behavior in simulated conditions. This is the main task for biomedical engineers. The mechanical tests of biomaterials should be conducted with consideration of mechanical behavior of the biological tissues (hard and soft). These rules are important for proper design of the safe and strong implants. The conditions of tests conducted in laboratory should be similar to the body environment. There are some procedures as physical experiment, mathematical simulations and computer modeling, which can provide sufficient data for the implant design process.

Keywords: Animal Tissue, Mechanical Parameters, Constitutive Modeling.

1 Introduction

For the description of true physicochemical properties of biomaterials, complex experiments are needed. Many expensive and long-term pre-clinical tests should be realized before manufacturing safe biomedical implants. Mechanical experiments are required for each of a new kind of material and of a prototype of implant. Primary, all of these efforts should guarantee good utility properties of implants through the time of implant contact with the body. After basic experiments conducted on specimens, the crucial properties and behavior of tissues and biomaterials under different types of load (static, dynamic, rheological) are recognized. The next step is to repeat experiments in a simulated body environment (both: specimens of material and implant). These are very long and expensive procedures, thus engineers can use mathematical or computer methods to estimate physicochemical properties of biomedical implants, before the prototype will be implantable.

The analyses of literature and our investigations revealed some basic features of soft tissues and its mechanical behavior. These features are: a non-linear stress-strain relationship [2, 3, 7], viscoelastic properties (stress relaxation at constant strain) [4, 6], mechanical hysteresis in cyclic loading and unloading [9, 8, 5] and preconditioning in

repeated cycles [7, 12]. Experimental studies of soft tissues properties are associated with many following factors which can affect its properties: anisotropy, age, sex, anatomical site of specimens taken, method of specimens storage before tests, etc. There is no one and only standard conducted for soft tissue mechanical examinations, thus results obtained by different scientists vary. Other limitations are the legal regulations and ethical reasons which make obtaining human tissue for experimental research difficult. Therefore, the material of animal origin is widely used as a substitute for human tissue in the biomechanical studies. In many studies, including ours, pig's tissue is used because of good similarity to human tissue [3, 1].

In this chapter, properties of chosen biological soft tissues and polymers were described on the base on experimental tests and results of experiments were used in modeling procedures with the use of ORIGIN and ANSYS software.

2 Methods and results

2.1 Preconditioning and mechanical hysteresis

Biomaterials and biological tissues are complex materials, which properties should be correlated during the design of implants. This process requires the knowledge about the principal properties of tissues and biomaterials. Soft tissues shows mechanical hysteresis, which can be calculated from hysteresis loops obtained in loading-unloading tests (see Fig. 1). These loops continually decrease over time of loading and then stabilize, if the test is repeated indefinitely. This phenomenon was described by Fung as the preconditioning effect [12]. It means that a steady state behavior of the material is reached and will not change until the cyclic load is changed. Despite many studies concerning the preconditioning effect, changes that occur during it in soft tissues, e.g. tendons, skin and aorta are not well understood [9]. Mechanical hysteresis is also associated with energy dissipation.

The following mechanical parameters can be calculated from hysteresis curves:

- the value of dissipated energy (the area of the loop),
- the value of total work performed on the tissue during stretching (the area under curve for increasing strain),
- the value of energy returned (the area under curve for decreasing strain),
- the value of mechanical hysteresis (the proportion of strain energy that is dissipated by internal viscous damping in each extension cycle. It is calculated as the ratio of the area within the stress-strain loop (strain energy dissipated) to the area beneath the load portion of the curve (strain energy input) [5].

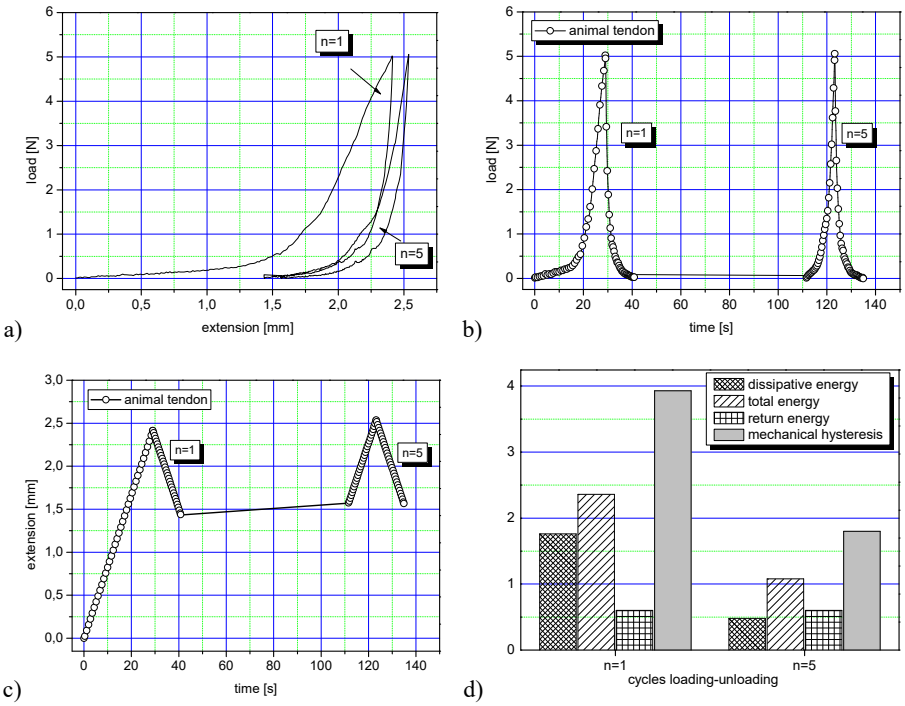


Fig. 1. The curves of loading-unloading cycles for animal tendon: a) hysteresis loops (Mullin effect), b) the level of load, c) extension versus time (in the first and fifth cycle of the loading-unloading),d) the change of value of dissipated energy, total work, energy returned and mechanical hysteresis between first and fifth cycle [for this example all in mJ unit].

Comparing the first and the fifth cycle, strong impact of preconditioning procedures on three parameters can be seen: the values of the energy of dissipation, the total energy and the mechanical hysteresis decreased, but the value of returned energy was almost on the same level. This is important information for the prediction of results of sport training or rehabilitation procedures, but also in design of artificial implants, for which physiological conditions of action are load-unload cycles, in the total time of life.

2.2 Preloading

Preparation of specimens (soft tissue or polymer) to the experimental test involves different methods and techniques as cutting, molding, injection, pressing, etc. These cause generation of pre-stress resulting from polymerization shrinkage, thermal shrinkage or loss of pre-stress during specimens isolating from animals. It can have an impact on results of mechanical tests and in consequence on the improper collection of input data to the computer simulation. Especially, when the group of materials is so large as for biomaterials (polymers, steels, alloys, ceramic, composites) and the places of its applications are also very different (soft and hard tissue). The set of preload in test procedure

enables introduction of comparative conditions and elimination of the artefacts resulting from the test machine type or clamping methods.

Correction of the input data is a very important step in future simulations. The value of 1, 2 or 5 [N] as preload is often used in experiments with soft tissues. This helps to limit dispersion of test results. In Fig. 2,the influence of preload on tensile curves was shown.

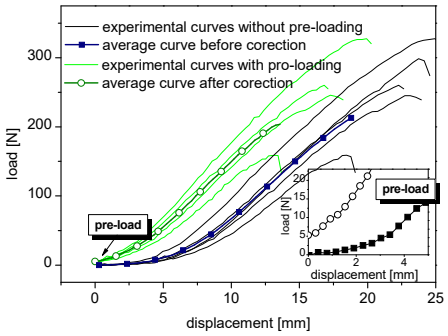


Fig. 2. Differences in load-displacement characteristic for uniaxial tensile test of pig's skin tissue without preloading and with preloading.

2.3 Test conditions

The specific test conditions are defined for each kind of engineering materials, but in case of matching implant materials with biological environment, the test conditions should be modified and correlated with physiological load conditions.
There are some important factors for the test:
- strain rate (load velocity) (see Fig. 3),
- direction of load,
- simply or complex load (uniaxial, biaxial, tensile load, tensile and shear load, etc.)
- type of load (force, moment, temperature (heat, cryogenic), etc.),
- time of acting (short, long-thermal),
- value of pre-loading,
- the method of fixing in grips machine (direct, indirect).

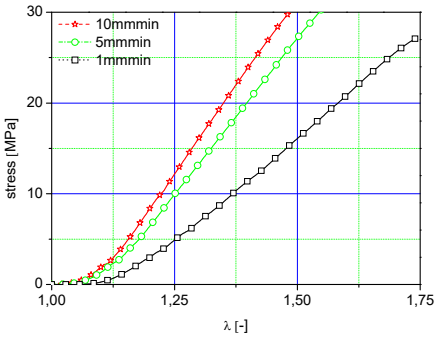


Fig. 3. Impact of strain rate on characteristic of uniaxial tensile curve for skin tissue [11].

The visible influence of strain rate on stress-strain relationship was shown in Fig. 3. These differences can determine e.g. the Young moduli value, which is one of material parameters defined in computer simulation of tissues.

Additional, the storage condition and the origin of tissues influence on the tests results(see Fig. 4). The biological samples should be properly prepared before the test. Fresh, wet or dry and short or long frozen (cryopreservated) samples can be considered. The type of storage medium is also important factor. These issues are important for collecting testing materials. When the object of test is tissue, the identifiably of origin is needed. Age, sex, kind and localization in the body should be described.

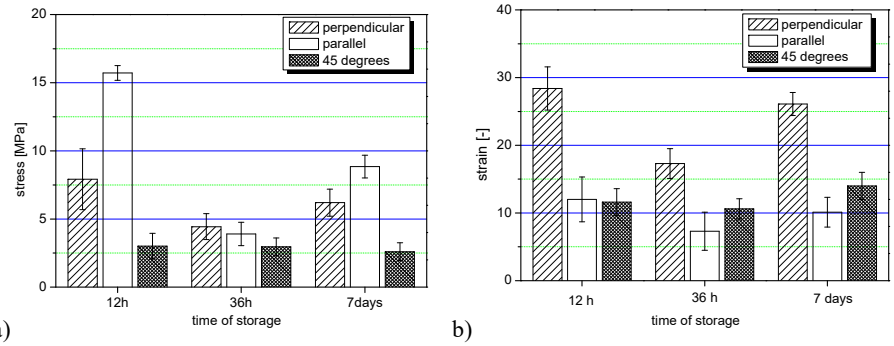


Fig. 4. Impact of storage conditions: a) on ultimate tensile stress (UTS) and b) strain at UTS in uniaxial tensile test.

2.4 Anisotropy

During computer simulation of biomedical implant behavior, features of biological tissue such as anisotropy, non-homogeneity, non-linearity, hyper-viscoelasticity and a remodeling skill of tissue must be considered. Thus, it is difficult to make the comprehensive description with mathematical formulas, and the analogous methods are used to simplify problems.

First simplifying assumption is to separate problems. Some of examples are given below:

- distinction of directions of cutting tissue samples (in engineering language are longitudinal, circumferential, perpendicular or parallel, in medical language are transversal, distal, mesial, occlusal sagittal, etc.) (see Fig. 5),
- distinction of the multiphase or multilayer material.

The use of experimental results is possible in FEM (Finite Element Method) modeling of biomechanical systems, where the materials parameters and character of material response under loading are needed. The cooperation of biologists and material engineers is needed to ensure safe use of implants, fitting of materials parameters and behavior in mutual relations. The mechanical characteristics of tissues shows a non-linear and hyperplastic character what must be considered in computer and mathematical simulations. The best way is to conduct multitask verification via many experiments. Such data can help to predict the behavior and implants-body relations under changing conditions in the future. Separation of problems and identification of mechanical parameters connected with directions of specimens taken is easy way to describe anisotropy of material.

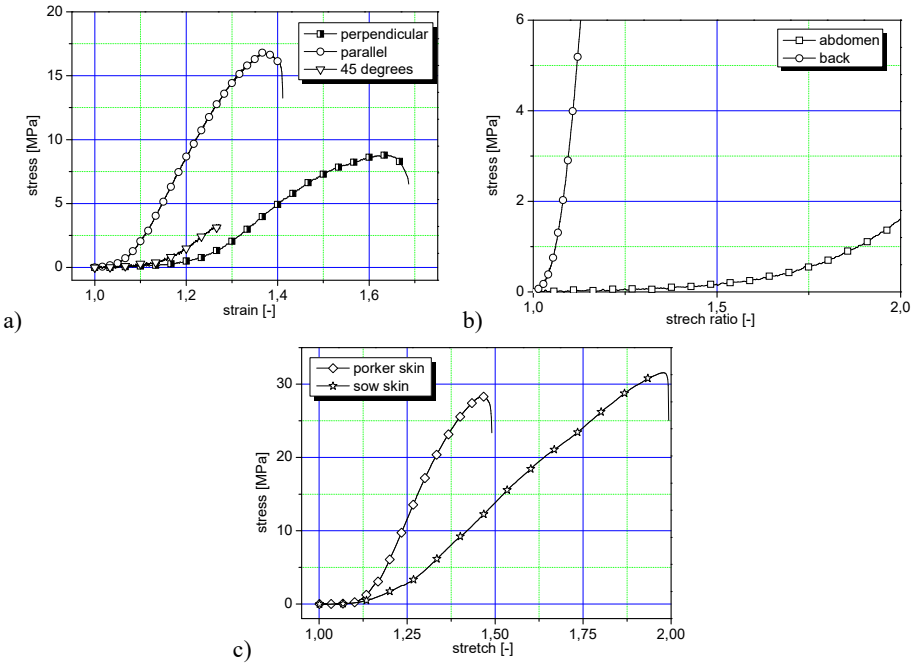


Fig. 5. The comparison of characteristic stress-strain curves of pig's skin tissue: a) for three different orientations of cutting (back area) [2], b) for two different regions of specimens taken [3], c) for two individuals (different) age and sex (perpendicular to the spine - orientation of taken).

2.5 Simulation of a body environment

The mechanical tests of biomaterials and tissues should be realized in conditions similar to the biological environment. The most important of them are body temperature, pH and humidity. These parameters have impact on properties of materials. During mechanical tests, it is necessary to simulate the biological environment by using simulated body fluids (SBF as a solution with an ion concentration close to human blood plasma), mild conditions of pH (7.25) and physiological temperature (36.6 °C). This is crucial for long-term tests, like the rheological test or fatigue test, not only for testing tissues but also for materials used for implants. The engineers should know, if the material has stabile dimensions and properties in dry and wet conditions. The thermal expansion and fluid absorption can induce local swelling and additional stresses.
The most common material used as biomaterials are polymers. Polymers have many shapes, structures (pure, as matrix or filaments in composites), can be dense or porous and gel or solid materials. These materials have many applications also as biodegradable implants (sutures, cardiac stents, scaffolds). Conducting tests in the simulated biological conditions provides data for computer and mathematical simulations used in predicting implant behavior in long-term or unstable conditions. The example of using a climatic chamber in relaxation test of polymer sutures is presented in Fig. 6 and described below.

a)  b) 

Fig. 6. The tensile test of chirurgical sutures: a) in special grips, b) in climatic chamber[4].

The relaxation test is carried out in room temperature conditions and in simulated biological environment conditions (0.9% NaCl solution, temperature of 37±1 °C). It was performed by attaching chamber with a temperature controlled saline solution bath. The strain level of 50% was selected for PP (polypropylene) and PGA (polyglycolide). For PDS (polydioxanone) suture, the selected level of strain turned out too high (the knot was broken), so the strain level was reduced to 23%. Each stress relaxation experiment lasted 24 hours.

The load, time and elongation were recorded, from which changes of stress versus time during stress relaxation tests were calculated. The stress relaxation curves are characterized by three specific phases: fast, transition and slow relaxation (see Fig. 7) [4]. The stress relaxation ratio, normalized stress and coefficient of fitting (material model parameters) can be calculated (see the formula 1-6) and used in computer simulations.

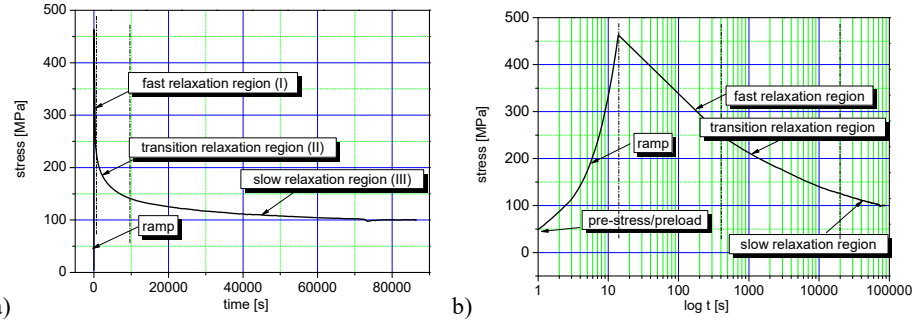


Fig. 7. Three characteristic phases of true stress relaxation curves: fast relaxation region (I), transition region (II), slow relaxation region (III)[4].

Also comparison of the value of load at break is important information for engineers and surgeons during selection suitable sutures for skin or tendons in different areas of the body from tissue support point of view. In Fig. 8 a comparison of the breaking force of three different sutures was presented.

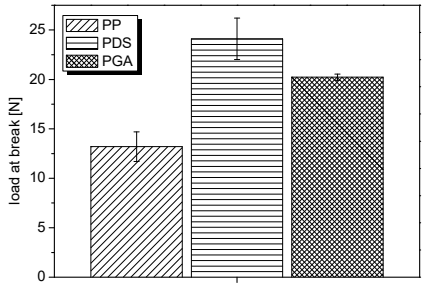


Fig. 8. Average breaking force for tested sutures[4].

2.6 Constitutive modeling of biomaterials and tissues

The modeling of biological and artificial materials is an important procedure which requires consideration of many aspects. Some of them were described below. The review of literature show that the QLV (quasi linear-viscoelastic) and hyper elastic descriptions are commonly use in computer simulations.

2.6.1 Quasi linear-viscoelastic (QLV)

The QLV constitutive model and the normalized stress function were investigated in this chapter. The normalized stress function is described as:

$$G(t)=\frac{\sigma(t)}{\sigma_{max}}, \tag{1}$$

$$G(0) = 1, \tag{2}$$

where σ(t) is the stress at time t, σ_{max} is the initial stress, amplitude at $t=t_p$ corresponding to the maximum stress. The temporary stress is defined by the following equation:

$$\sigma(t) = G(t) \cdot \sigma^e(\varepsilon), \tag{3}$$

where $\sigma^e(\varepsilon)$ is the stress temporary strain.

The time-dependent mechanical behavior of soft tissues is given as:

$$\sigma(t) = \int_0^t G(t-\tau)\frac{\partial \sigma^e(t)}{\partial \varepsilon}\frac{\partial \varepsilon}{\partial \tau}, \tag{4}$$

where $\frac{\partial \sigma^e(t)}{\partial \varepsilon}$ is the temporary elastic response, and $\frac{\partial \varepsilon}{\partial \tau}$ is the time-dependent strain of the sample. The reduced relaxation function is given as Prony's series:

$$G(t) = ae^{-bt} + ce^{-dt} + ge^{-ht}, \tag{5}$$

where a, b, c, d, g, h are constants, which could be determined from experimental data.

An exponential has been often used to describe the nonlinear elastic behavior of skin tissue:

$$\sigma^e(\varepsilon) = A(e^{B\varepsilon} - 1), \tag{6}$$

where A is a linear parameter which has the same dimension as stress, and B is non-dimensional function factor describing the nonlinearity of elastic response.

The optimization procedure was performed by using ORIGIN to generate data curve-fit for samples of skin tissue, and the results of simulation were presented in Fig. 9 for pig's skin samples.

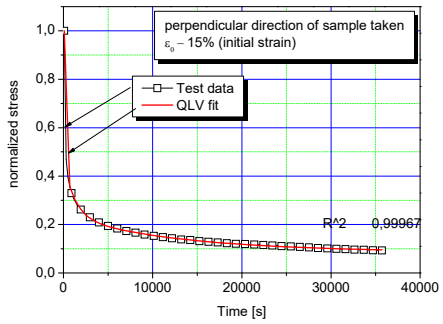


Fig. 9. Comparison of QLV fitting curve with experimental data of perpendicular skin sample with 15% strain [6]

2.6.2 Hyper-elastic material properties

Hyper-elastic material properties are a next type of characteristics which can be used to better design of tissue implants, with accurate bio-functionality. A hyper-elastic material model based on the definition of strain energy function, which is expressed in different ways, depend on the class or kind of materials considered.
Assuming isotropy of material, the strain-energy function can be written as depended on the strain invariants of the deformation tensor of Cauchy-Green $W(I_1, I_2, I_3)$. Considering the conditions of uniaxial tension of incompressible materials (σ_2=σ_3=0) and additional, if the hyper-elastic material is incompressible, the strain-energy function of few models can be given according to the formulas included in Table 1.

Table 1. Utilized strain energy functions formulas for material under uniaxial tension

Material model	Incompressible strain energy function
Mooney-Rivlin	$W = C_1(I_1 - 3) + C_2(I_2 - 3)$
Yeoh	$W = C_1(I_1 - 3) + C_2(I_1 - 3)^2 + C_3(I_1 - 3)^3$
Ogden	$W = \frac{2\mu}{\alpha^2}(\lambda_1^\alpha + \lambda_2^\alpha + \lambda_1^{-\alpha}\lambda_2^{-\alpha} - 3)$

To obtain an approximation of experimental data with mathematical a record, the fitting procedures with the use of software ORIGIN software were conducted. To the determine the values of C_i, algorithm of Levenberga-Marquardta (damped least-squares) was used. For the iteration process the initial values of parameters were defined as equal to 1. The results of modeling were presented in Fig. 10.

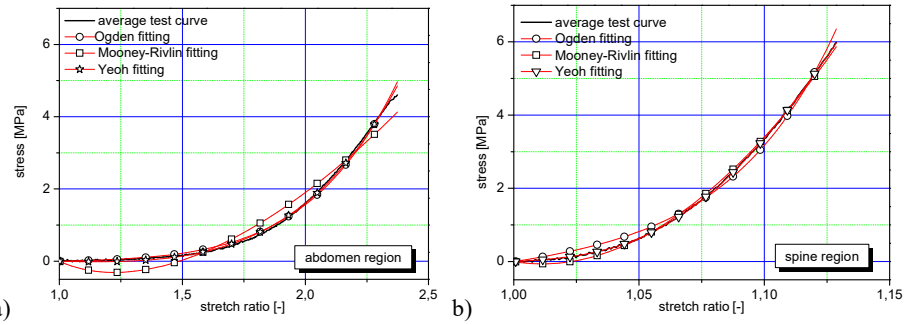
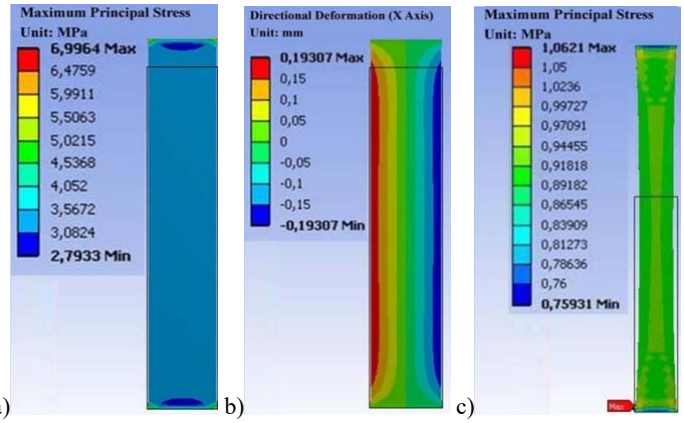


Fig. 10. The stress-stretch curves for skin samples including fitting of hyper-elastic material models: a) samples taken from the abdomen region; b) samples taken from the spine region [3].

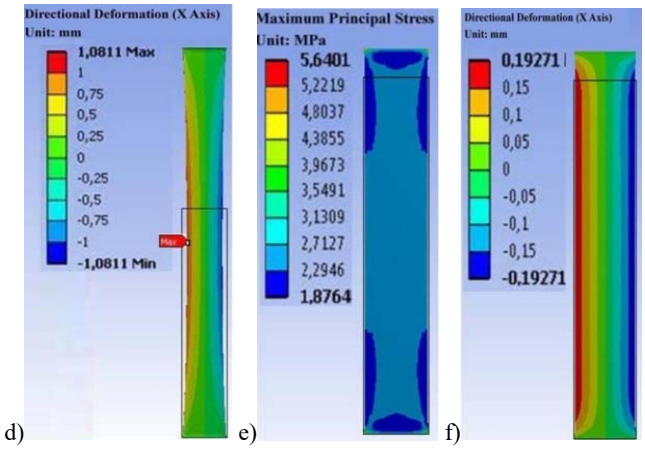
2.7 Finite element methods in hyper-elastic model of skin tissue

The animal soft tissues (skin, tendons, etc.) are often used as substitute of human tissues in experiments, as was shown. It is a good initial idea to provide the animal models as biological phantoms in surgical practice, as well as in pre-clinical tests of implants. When the correlation of behavior (e.g. bio-functionality or biocompatibility) of artificial implants and biological tissues is needed, the finite elements methods can be used. The hyper-elastic material models like the Mooney-Rivlin, the Ogden, the Yeoh are implemented in FEM software like ANSYS, so it can be used in complicated strength analysis of biomechanical implants, in the aim to estimate relationships between implant and tissue. The comparison of using the FEM analysis of skin tissue samples with using hyper-elastic Ogden and Yeoh models and with differentiation of body area (back and abdomen of animal) was realized in [10]. The results of analysis confirm the impact of location of skin tissue on behavior under load.

The different hyper-elastic behavior of skin tissue was shown in [10] (see Fig. 11) with the use of the Ogden and Yeoh material models. The different values of stress and deformation were observed depending on body location and type of model.



a) b) c)



d) e) f)

Fig. 11. The examples of distribution of stress and deformation specimen of skin tissue: a,b) skin from back – Ogden 1st order model; c,d) skin from abdomen region – Ogden 1st order model; e,f) skin from back – Yeoh 3th model [10].

3 Conclusions

Experimental investigations conducted in tensile test, stress relaxation tests or under loading-unloading conditions give the possibility of a better understanding soft tissues biomechanics and provide data for modeling and predicting tissues behavior under different loading conditions. This knowledge is useful both in medicine and engineering. Tests conducted on animal models provide information, e.g. for surgery, on how to close wounds with the appropriate tension, or in the analysis of pathological changes in tissues. For example, the processes of stress relaxation of skin tissue can be correlated with the properties of surgical threads. Our research in the field of stress relaxation of various types of surgical threads showed similar values of the relaxation coefficient of polypropylene thread to the values obtained for the skin [4]. Also in the design of implants, surgical instruments, medical robots or forensic analysis to assess mechanical damage or traffic accidents, this knowledge is essential.

References

1. Jor J.W.Y., Nash M. P., Nielsen P.M.F., Hunter P.J.: Estimating material parameters of a structurally based constitutive relation for skin mechanics. Biomechanics and Modeling Mechanobiology 10, 767-778 (2011).
2. Łagan S., Liber-Kneć A.:A characteristic of anisotropic mechanical properties of a pig's skin=Charakterystyka anizotropowych właściwości mechanicznych skóry świni.Engineering of Biomaterials 17(128-129), 61-63 (2014).
3. Łagan S., Liber-Kneć A.: Experimental testing and constitutive modeling of the mechanical properties of the swine skin tissue. Acta of Bioengineering and Biomechanics 19(2), 93-102 (2017).

4. Liber-Kneć A, Łagan S.: The stress relaxation process in sutures tied with a surgeon's knot in a simulated biological environment. Polymers in Medicine 46(2), 111-116 (2016).
5. Liber-Kneć A., Łagan S.: Mechanical hysteresis tests for porcine tendons. Engineering of Biomaterials 20(139), 26-30 (2017).
6. Liber-Kneć A., Łagan S.: Testing stress relaxation process of a porcine skin. Engineering of Biomaterials 19(134), 18-24 (2016).
7. Ni Annaidh A.N., Bruyere K., Destrade M., Gilchrist M.D., Maurini C., Ottenio M., Saccomandi G.: Automated estimation of collagen fiber dispersion in the dermis and its contribution to the anisotropic behavior of skin. Annals of Biomedical Engineering 40(8), 1666-1678 (2012).
8. Wiśniewska A., Liber-Kneć A.: Influence of a skin tissue anisotropy on mechanical hysteresis. Technical Transactions. Mechanics5-M(15),126-136 (2016).
9. Liber-Kneć A., Łagan S.: Comparison of mechanical hysteresis for chosen soft tissues taken from a domestic.Advances in Intelligent Systems and Computing, vol. 925, pp. 261-268. In: Tkacz E., Gzik M., Paszenda Z., Piętka E.(eds)Innovations in Biomedical Engineering 2018.Springer, Cham(2019).
10. Łagan S., Chojnacka-Brożek A., Liber-Kneć A.: FEM analysis of hyperelastic behavior of pig's skin with anatomical site consideration.Advances in Intelligent Systems and Computing, vol. 831, pp. 202-209. In:Arkusz K., BędzińskiR., Klekiel T., Piszczatowski S. (eds) Biomechanics in Medicine and Biology: Proceedings of the International Conference of the Polish Society of Biomechanics 2018. Springer Cham (2019).
11. Łagan S., Liber-Kneć A., Influence of strain rates on the hyperelastic material models parameters of pig skin tissue.Advances in Intelligent Systems and Computing, vol. 623, pp. 279-287. In: Gzik M., Tkacz E., Paszenda Z., Piętka E. (eds) Innovations in Biomedical Engineering 2017. Springer, Cham(2018).
12. Fung Y.C.: Biomechanics. Mechanical properties of living tissues. Springer, New York, 1993.

The Use of Information Technology in the Designing and Manufacture of Implants

Oleksandr Tarasov [1[0000-0002-0493-1529]], Alexander Altukhov[1[0000-0002-6310-3272]], Eduard Gribkov[1[0000-0002-1565-6294]], Anzhelika Parkhomenko[2[0000-0002-6008-1610]] and Andrii Kovalenko[1[0000-0003-3379-2000]]

[1]Donbass State Engineering Academy, 72, Akademichna str., Kramatorsk, 84313, Ukraine
[2] National University «Zaporizhzhia Polytechnic», 64, Zhukovskogo str., Zaporizhzhia, 69063, Ukraine
alexandrtar50@gmail.com

Abstract. The possibilities of automating the solution of problems of designing implants for various purposes and developing the processes of their manufacture are considered. It is shown that the use of integrated computer-aided design systems ensures the implementation of the end-to-end process of designing and manufacturing implants, as well as modeling the behavior of implants during operation. The use of information technologies, industrial CAD / CAE / CAM systems and approaches in the design of serial and individual implants of several types is considered. The process of manufacturing a U-shaped implant for spine surgery from titanium Ti6Al4V (VT6) is proposed. The process includes modeling in CAD / CAE systems for obtaining an implant workpiece by deformation and developing a process for processing the workpiece using CNC equipment in a CAM system. The application of this process allows the use of severe plastic deformation (SPD) when forming a U-shaped implant workpiece. An ultrafine-grained structure is formed in the workpiece and the physical and mechanical properties of the material are improved. Modeling was carried out using CAD / CAM systems and the finite element method (FEM) in the Abaqus CAE system. Calculations have shown that the degree of deformation during the deformation of the workpiece reaches 10 or more, which provides a high-quality structure in the workpiece material. The design of an implant made of sheet material for maxillofacial surgery has been completed. Preliminary shaping of the implant workpiece based on the patient's jaw model can significantly reduce the operation time during its installation. Modeling using CAD / CAE systems allows you to determine the quality of the implant attachment to the jaw and allowable loads. The modeling of the loading of the patient's teeth during the installation of the bracket system was also performed. The development of the patient's teeth model was performed in a CAD system. The model allows you to determine the places of rational installation of brackets, taking into account the real shape and position of the teeth. The presence of the model made it possible to calculate in the CAE system, the load on the teeth from the arch wire of the bracket system.

Keywords: Implant, CAD, CAE, CAM, Finite Element Method (FEM), bracket, arch wire

1 Introduction

The development of technology constantly requires new solutions: constructive, technological, organizational, etc. The process of creating a new one is creative, innovative, and creativity is also a process that is used from different points of view. The effectiveness of new technical solutions supports the interest in modern systems methods, algorithms for solving the tasks of automating the design process, including the solution of creative problems of creating new designs and technological processes. For the automated solution of various tasks in the field of design and development of technological processes, integrated computer-aided design systems are used, which ensure the implementation of the end-to-end process of designing, manufacturing and quality control of products.

Implants of various types were considered as an object of design and development of technological processes since they require an individual approach to design in each case, have a complex geometric shape, therefore automation tools are necessary for their design and manufacture. Materials that are used for implants should have high physical and mechanical properties, which increases the urgency of developing modern methods for improving the quality of the structure of the material of implants and, accordingly, methods for designing and modelling the processing of such materials. Such methods, in particular, include the technology of pressure treatment with the intensification of shear plastic deformation (SPD process) [1]. The use of these methods allows obtaining bulk blanks for implants with ultrafine-grained structure.

The capabilities of modern information technologies in the preparation and processing of three-dimensional models are used in various tasks of designing products, including medicine, in the development of designs of implants, prostheses, medical tools, etc.

Modern software allows to design the shape and composition of prostheses (CAD-system), check the size of the prosthesis according to the obtained images of the damaged parts of the patient's body. Next, perform a strength calculation under different load conditions (CAE-system) and, if necessary, change the design, both in the direction of decreasing strength and increasing. After obtaining the required shape of the prosthesis, prepare its model for manufacturing by 3D printing or machining on CNC machines.

Design and manufacture of medical implants are among the complex technological tasks since such products are subject to increased requirements according to some criteria. At the same time, obtaining implants with the help of modern methods of processing materials, such as SPD, creates conditions for expanding the range of implants produced and the materials used, which will allow their use in new, previously not possible areas of application.

When designing implants, individual and serial approaches should be considered. For implants used in maxillofacial surgery, an individual approach is required, which is connected with the peculiarities of the structure of the mandible bones and, as a rule, the specifics of injuries received for a given bone. In this case, it is required to obtain an image of the injured area of the jaw and to design the implant directly for the damaged area. Injuries to the hip joint and back have a typical character and lead to the installation of standard implants in strictly defined places. Hip and spinal implants are

used in typical cases and do not require an individual approach for each patient. In such cases, it is enough to have a set of implants with different sizes to cover all cases of injury for different patients.
At this time there is a lot of specialized software for designing highly specialized implants directly for each field of medicine. In dental implants , when designing dentures, the Delcam DentCad package and the DentMill package are widely used in the manufacture of prostheses on CNC machines using the prepared models.
The development of technology for the design and manufacture of implants is observed in many areas:

- ─the creation or improvement of the characteristics of the materials used;
- automation of methods for obtaining and processing information about defects of bone-stems on which implants will be installed;
- automation of the design of models of implants and technological equipment for their production;
- analysis of the conditions of functioning of the implants to ensure a given strength and rigidity of the structure;
- automation of the design process of manufacturing templates for the subsequent formation of the surface of the implant, if used parts from a flat sheet blank;
- automated production of the required elements on CNC machines and control of manufacturing accuracy;
- improvement of the technology of forming blanks by the method of plastic deformation to ensure the versatility of the process and the accuracy of obtaining a given geometry of implants.

Addressing these issues requires the participation of specialists in various subject areas, primarily medical professionals, specialists in the field of information technology design and various types of metal processing.
Tomography makes it possible to obtain a volumetric model of human bones for analyzing their condition and concluding on the need for prosthetics. The doctor chooses the type, place of installation and the shape of the implant, taking into account the existing experience of prosthesis, the applied methods, materials and tasks of the prosthesis, as well as the anatomical features of the patient. This information is the source for the designer, working in the CAD-system and performing the development of the project of the implant and tooling, tools for its manufacture. The main task of the preparatory stage is the development of a formalized design assignment for the designer and the technologist.
Technological processes for the manufacture of implants depend on their purpose, geometric shape, size and material used.
One of the most promising processes for the manufacture of implants for maxillofacial surgery is forming a sheet blank to obtain the required spatial form. At present, a rather wide range of sheet blanks of medical titanium Grade1 is being produced, from which sheet implants are made [2-7].
Sheet blanks for implants are relatively cheap, however, the nomenclature is focused on manual processes of giving them the required spatial shape. Therefore, implants with

a complex spatial shape do not always meet the requirements for structural rigidity, which significantly reduces their quality. The manufacture of such implants from bulk blanks is much more expensive, especially when using titanium, subjected to pretreatment using SPD. Given the size of the sheet blanks for implants, it is relevant to obtain them from modern materials through pressure treatment using universal process equipment and tooling. The objective of this stage is the choice of the method of obtaining blanks and implants, acceptable for technological complexity and availability of equipment.

It should be noted that the shape and size of the implant are determined not only by the patient's anatomy nevertheless also by the operating conditions. For example, after installation on the jaw, there is a partial restoration of its working functions to ensure the nutrition of the patient. Therefore, it is necessary to analyze the strength of the implant under conditions of loading during chewing and to ensure the high-quality fixation of the implant on the bones of the jaw. Installation site and bone strength characteristics are an important limiting factor in the design. The most effective for solving this problem is the application of the finite element method and the CAE-systems in which it is implemented. When setting this task, it is necessary to involve medical workers for the correct setting of initial data and their analysis. As a result of the analysis, an agreed decision on the design of the implant is made.

The technological process of manufacturing an implant depends not only on its geometry but also on the availability of raw materials and preforms. For example, when using titanium subjected to SPD to increase strength (while observing a given level of plasticity), sheet blanks are not produced by industry, therefore obtaining billets with the required thickness and dimensions in each particular case is also an urgent task.

Obtaining a flat billet with processing thickness can be performed on a CNC ma-chine, which requires the use of CAM-systems for the development of a control pro-gram, however, is not economically profitable. When a spatial form implant is fabricated by plastic deformation from a sheet, a sweep is required to determine the shape of the workpiece contour, therefore this task requires separate study using the CAE sys-tem. In any case, along the contour of the workpiece, technological allowance is necessary, followed by disposal of waste.

The design of technological equipment is made taking into account the universality of the process of manufacturing implants. Giving the blank implant a volumetric shape is possible by deforming polyurethane on a matrix - template. This method is the most universal since it requires only one working tool that determines the shape of the implant. The shaping of the workpiece is made in universal containers, the dimensions of which form a discrete number of values for processing workpieces of various sizes.

The template for shaping the workpiece is designed in the form of a matrix with a cavity or a punch, depending on the shape of the workpiece. Designing it is made taking into account the geometry of the part and the springing of the workpiece material in the process of deformation. Thus, the shape of the template differs from the surface of the implant, which must be considered when designing a template. After the design and manufacture of an implant, an analysis of the process results is needed to further improve the prosthetic technology and identify functions to develop the capabilities of the CAD subsystems.

Consider designs of implants of various shapes and purposes for spinal and maxillofacial surgery.

2 Development of a spinal U-implant design

In the world medical practice, obtaining Coflex U-implant spacers for spinal surgery [8-10] from titanium alloys Ti6Al4V (VT6) is well known. However, the use of technically pure titanium, which does not adversely affect the patient's body, is difficult due to insufficient strength without special processing methods. The use of SPD for pretreatment of bulk titanium blanks can increase the strength of the material while maintaining sufficient ductility [1].
In the manufacture of the implant, you must observe the following rules:

- the implant must be comfortable and functional for both the patient and the surgeon;
- the implant should provide wide parametric possibilities in fitting during prosthetics;
- the implant must ensure the long-term preservation of consumer properties, with the universality of the starting materials for its manufacture.

To build a model of an intervertebral implant, the initial information is obtained using x-ray methods for examining the patient's spine. The detection and analysis of irregularities are made by the doctor, who decides on prosthetics and builds a three-dimensional implant model in a specialized computer system. The result of the design is the visual implant model required to obtain it, as well as the results of calculating the stresses and deformations of the product during its use after installation to the patient. These data make it possible to guarantee the manufacture of a high-quality implant and the positive results of the prosthetic process.
The method of interstitial stabilization of the prosthetic segment is used in the treatment of intervertebral hernia in the lumbosacral spine. This is the most frequent surgical intervention, which applies up to 90% of all surgeons operating on the spine.
For the design of the U-implant design, we consider it in expanded form (see **Fig. 1**).
To achieve this shape of the workpiece, you can use stamping operations for several transitions. The stages of the construction of the workpiece and the implant are similar, consider them on the example of the workpiece:

- The first action: the creation of the main part of the workpiece, an indication of its coordinate system and the formation of its profile (see **Fig. 2.**).
- The second action: the formation of four petals (wings). To perform this action, it is necessary to create a contour of the petal on one side of the part and reflect it relative to the axis (see **Fig. 3**).

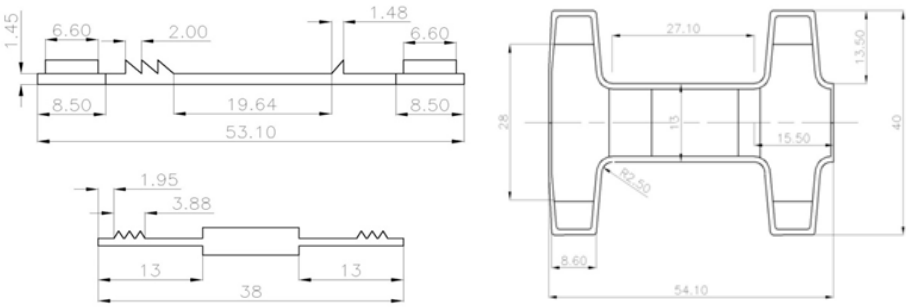


Fig. 1. Drawing of the intervertebral implant scan

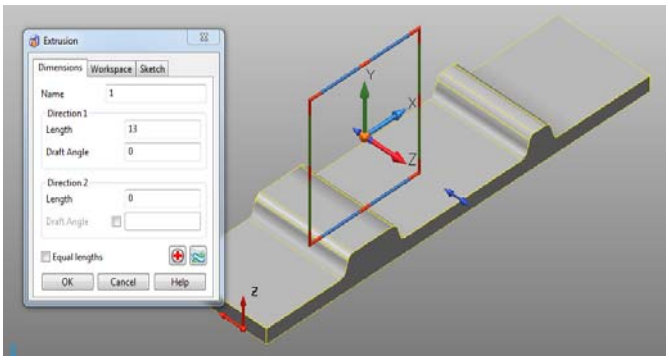


Fig. 2. Creating the main part of the workpiece

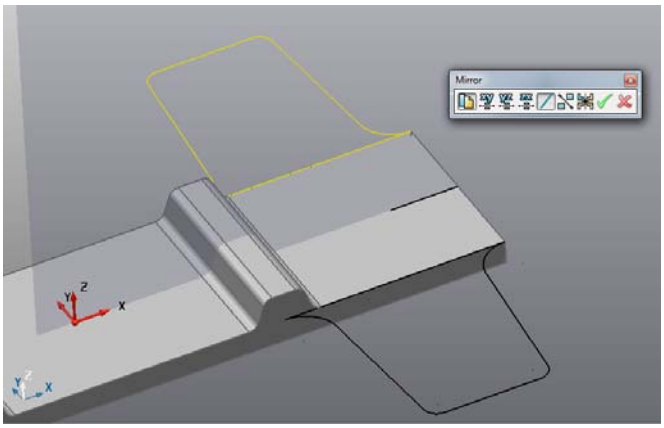


Fig. 3. - Formation of the contour of the petal and its reflection relative to the axis

- The third action: squeezing out the contours of the resulting petals and forming superstructures (protrusions) on them, from which teeth will be cut in the future (see **Fig. 4**).

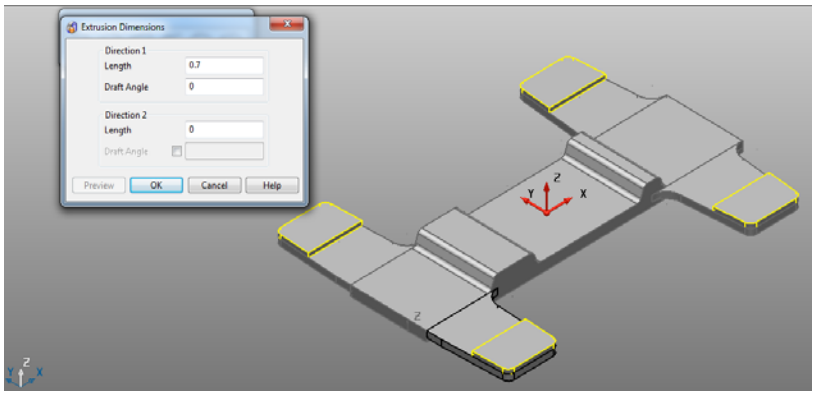


Fig. 4. - Formation of petals and superstructures

- Note: the part is built according to a similar algorithm, however, there are differences in the formation of the profile of the part and the formation of superstructures on the petals, since instead of them there are teeth (see **Fig. 5**).

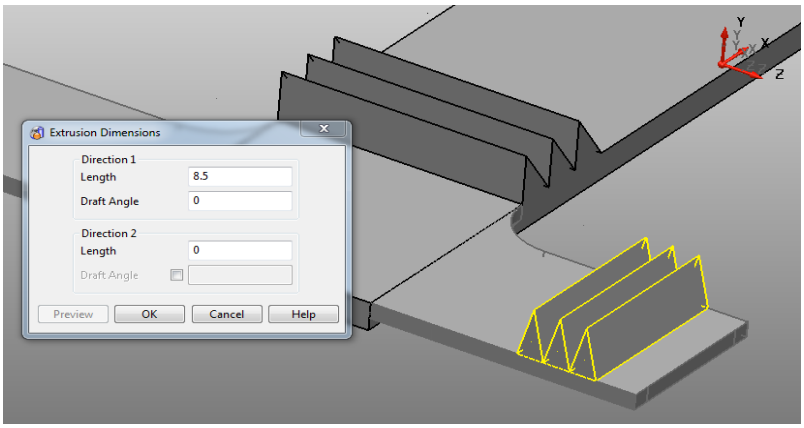
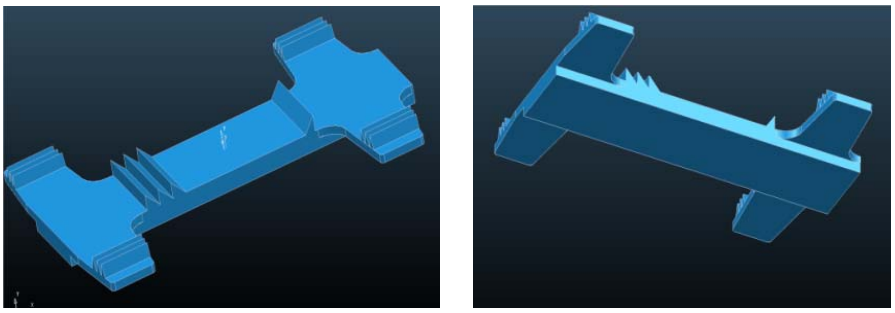


Fig. 5. - Formation of teeth on the petals of the intervertebral implant

As a result of the manipulations, a three-dimensional model of the intervertebral implant was built (see **Fig. 6**). 3D model of the workpiece to obtain the intervertebral implant (see **Fig. 7**).

According to the required shape of the workpiece, the design of a die tool: matrix and punches. The resulting tool models are used in the CAE system to simulate the behaviour of the workpiece during the deformation process [11]. The result of the simulation, namely the deformed preform is located in the matrix, is shown in **Fig. 8**.



a b

Fig. 6. - Three-dimensional model of an intervertebral implant: a - top view, b - bottom view

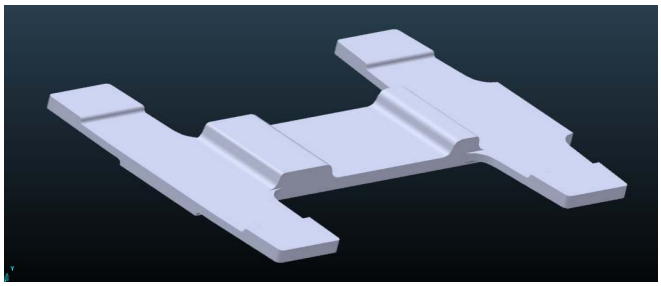


Fig. 7. - Three-dimensional preparation for creating a model of the intervertebral implant

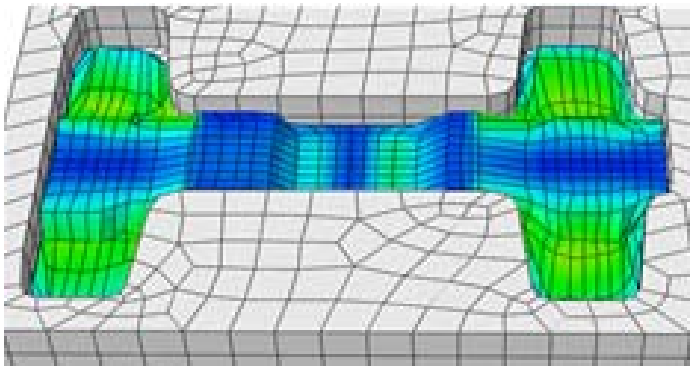


Fig. 8. – Matrix for modeling the deformation process of the implant workpiece

A feature of the proposed solution is the sequence of forming the implant blank (**Fig. 9**). During the first operation (b), the required length of the workpiece is formed, while the metal flows along the workpiece. When performing the second operation (c) of the workpiece deformation, the metal of the end zones flows in the transverse direction. This makes it possible to form workpiece zones from which four wings will be created. During the third deformation operation (d), four wings of the workpiece are formed. In this case, the central zone of the workpiece is not deformed. Then the spatial shape of

the implant is formed by bending in two directions. After processing by pressure, the final machining of the product is carried out.

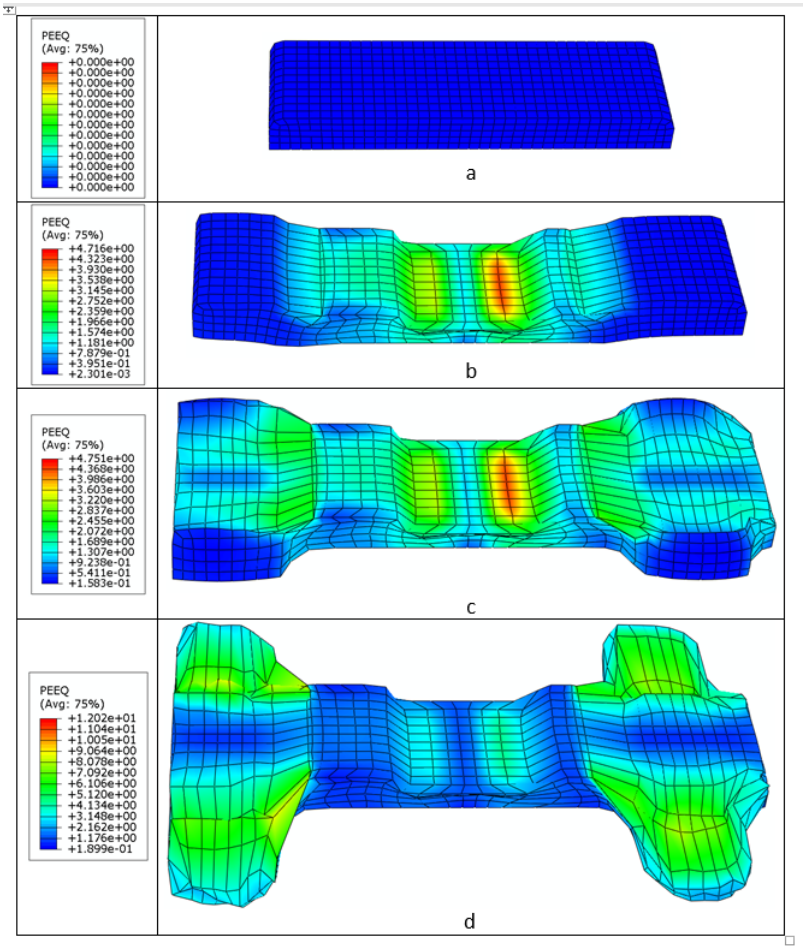


Fig. 9. – Stages of modeling the deformation of the implant blank in a CAE system

The general scheme for installing the implant on the spine is presented on **Fig. 10**. . This deformation technology provides a favorable direction of the fibers in the material of the finished product. Intense shear deformations during deformation contribute to the formation of a fine-grained structure in the workpiece material. The directions of bending when obtaining the spatial shape of the implant are made across the fibers, which increases the fatigue strength of the product.

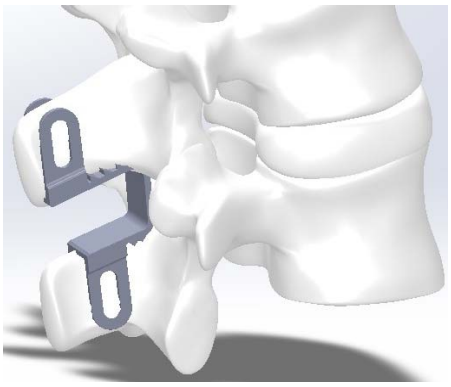


Fig. 10. – Spinal implant placement scheme

3 Development of manufacturing technology for U-implant in the CAM-system

To build a CAD system for the formation of a technological process for the treatment of an intervertebral implant, it is necessary to identify the main actions performed by the technologist. To do this, you need to build a typical technological process in the CAM system, and then in CAD to implement the interface necessary for building this process.

The formation of elements of the technological process of processing a part contains several interdependent items.

The first point is to load the part and the workpiece, apply a single coordinate system to them and monitor the compatibility of this coordinate system with the machine. It is believed that the X and Y axes are in the plane, and Z is located vertically.

The second point is the creation of the tool. Sometimes this item requires special attention and the tool needs to be created on a case-by-case basis, and sometimes it's completely simple to select it from the database of previously created tools. In the case of creating an intervertebral implant, it is necessary to create a special tool, a conical spherical cutter with an edge angle of 15 degrees, an edge radius of 0.1 and a small diameter (see **Fig. 11**).

The third point involves the creation of trajectories in such a way that when fixing parts on the frame, you can vary the trajectories so that they do not overlap with fasteners. In the process of performing this item, three trajectories were constructed to process the rear and nine trajectories for the front. Trajectories may not be in the same coordinate system relative to each other. For example, to process teeth, it was necessary to create a coordinate system that was at an angle of 15 degrees relative to the global one (see **Fig. 11**).

NC file, which is the product of the CAM system. When creating NC files, it is necessary to understand how the processing will take place, since it is necessary to specify the sequence of trajectories recording. As a result of this work, only two NC files were

built, for the front and rear part machining. As described above, the rear part is processed by three paths, and the front part by nine. **Fig. 12** shows a visualization of all part paths at once.

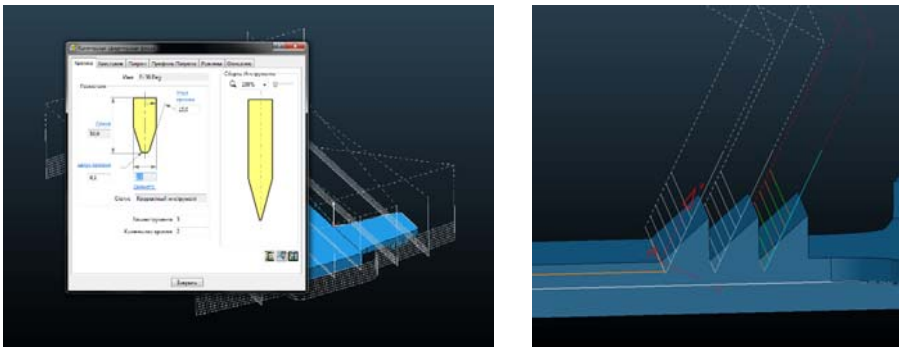


Fig. 11. - The process of creating a tool and the treatment process of the teeth on an implant

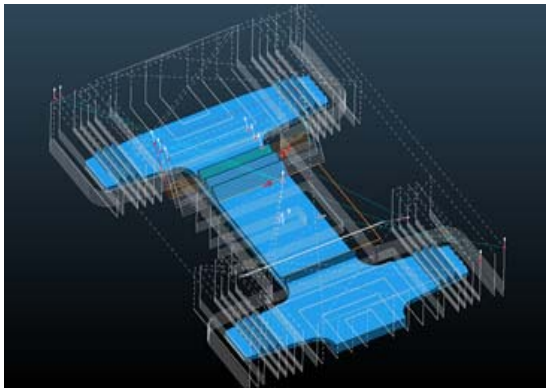


Fig. 12. - Visualization of all part paths at once

Based on the analysis of possible options for integration with Delcam PowerMILL, a CAD was developed that can directly access the Delcam PowerMILL CAM system API and then use the macros to perform the necessary work of process design.
Integration with Delcam PowerMILL CAM system is implemented using PowerSolutionDOTNetOLE DLL-library. Macros have been used to call the PowerMILL functions as well as the PowerMILL dialogue boxes.
In the development of CAD to ensure integration with the CAM-system PowerMILL the following objectives were set:

- the ability to run the window PowerMILL without the interface of the system itself;
- the ability to use tools PowerMILL, referring to them using CAD;
- the ability to import the project in PowerMILL using CAD;
- the ability to add tools to the CAD system while the program is running.

For integration with PowerMILL, the C # programming language was used.
Having provided the ability to integrate Delcam in its software products, Campaign developed a library for .Net platform called Pow-erSolutionDotNETOle. This library can access the base elements of the PowerMILL program and call them.
When developing the project of the integration program, it was necessary to establish a connection of the program with the library and obtain an object of the class of this library for further work with it. Thus, as a result of establishing a connection with the library and opening the PowerMILL window, all that remains is to execute the command using the "execute" method of the "PowerSolutionDotNETOle" class. As described earlier, the API is accessed by similar commands as in macros. That is, you can write a macro, open it with the macro editor, take the recorded string and insert it into the "execute" method of the "PowerSolutionDotNETOle" class. In the course of performing these operations, the CAD is addressed to the PowerMILL core and sends a dynamically generated macro to it. For the PowerMILL kernel, executing this line is as easy as executing a normal macro, which can be selected directly in the PowerMILL interface.
The next task for CAD was the definition and creation of text files of basic commands for PowerMILL. To do this, it was necessary to execute all the necessary commands in turn, while recording the macros and assign it the name of the corresponding command. After creating the macros, the CAD system should have referred to the directory in which all the macros are located. The CAD received a list of paths to the files of the specified format in this directory and formed an array with the paths to the files and their names. In addition to invoking a list of macros, CAD can dynamically update this list right during its work.
An example of a 3D sample operation:

- Import project.
- Creating a tool.
- The choice of processing strategy.
- Creating boundaries for a given strategy (see **Fig. 13**).
- Calculation of the processing trajectory (see **Fig. 13**).

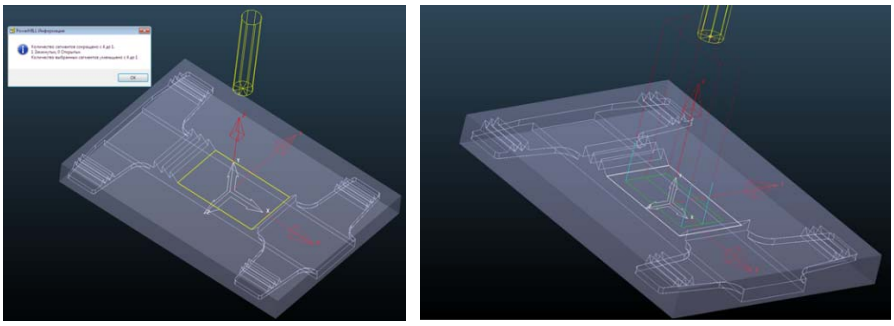


Fig. 13. - Creating boundaries for a given strategy and calculation of the processing trajectory

4 Designing an implant from sheet material

Baseline information is obtained using computed tomography of the patient's jaw. The detection and analysis of irregularities are made by the doctor, who decides on prosthetics and builds a three-dimensional implant model in a specialized computer system. As a result of the design, the following are obtained: visual models of the implant and accessories needed to obtain it (in the CAD system); as well as the results of the calculation of stresses and deformations of the product during its use after installation to the patient (in the CAE system). These data make it possible to guarantee the manufacture of a high-quality implant and the positive results of the prosthetic process.
The specialist starts the system by inputting the images obtained with the help of the tomograph. Pictures are needed to obtain a preliminary model of the jaw, and medical indications to determine the type of defects and the approach for prosthetics.
Next, the model of the damaged jaw is transmitted to the CAD system for correction and elimination of model defects. The transferred model has a format different from that of PowerSHAPE, therefore, the possibilities of working with it are limited. To make possible full-fledged work with the model, the construction of the implant and the equipment must be converted to it using standardized data exchange formats (STEP, IGES ISO, HSF, etc.).
After processing the model of the jaw, the specialist indicates on it the basic surfaces and points necessary for the construction and installation of the implant. Next, based on the binding of the implant to the jaw, it is built according to the appropriate algorithm. Consider the process of designing an implant with a specific example. The CAD model of the implant for osteosynthesis in oral surgery is shown in **Fig. 14**.

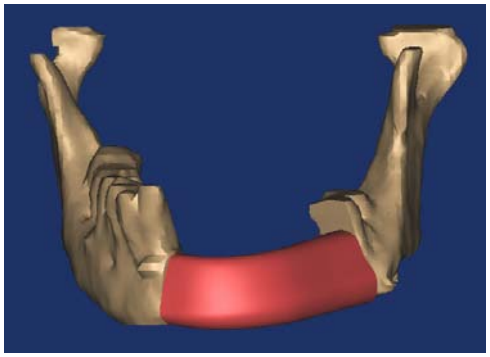


Fig. 14. - 3D-model of the "assembly" of the lower jaw fragments with an implant [14]

The technique for building the implant surface should be universal and allow for the possibility of adjusting the shape (see **Fig. 15**). To do this, we develop a generalized algorithm using cut planes, which allows us to obtain an array of implant sections with planes perpendicular to its generator. The variation in the number of sections allows you to change the accuracy of the construction of the surface of the product. The presence of sections provides the possibility of adjusting the profile of the implant with the

subsequent construction of its surfaces taking into account the thickness of the workpiece and the template (punch) to obtain it.

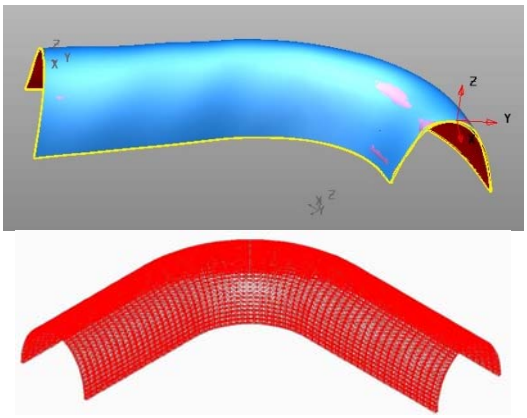


Fig. 15. - Construction of the surface for the template tool (punch) and implant with applied nodes (mesh)

To determine the loads acting on the implant during operation (when chewing), it was modelled in the CAE system [12]. For discretization of volume and calculations, four-node tetrahedral finite elements were used. A uniform grid of twenty elements along each of the generators and two elements along the plate thickness was used (see **Fig. 15**).

5 Designing Brackets

Examination of the patient allows the doctor to determine the need for orthodontic treatment, however, then it is necessary to carry out diagnostic studies, which include: obtaining control plaster models, facial photographs of the face and profile, intraoral photographs, X-ray diagnostics - orthopantomogram, radiography, scanning or tomography. When treating with the help of brackets, the patient is taken photos and smiles. Assessment of the patient's smile is also important since one of the main tasks of an orthodontist is to obtain an optimal treatment result. Diagnostic tools allow you to assess the true location of the teeth and roots, the position of wisdom teeth (third molars), the location of the jaws to the base of the skull, the angles of inclination of the teeth, makes it possible to evaluate the patient's profile, etc. [12].

Jaws, human teeth - are individual and have a complex spatial form. Therefore, to improve the quality of treatment, modelling tools are required to reduce the risks of adopting a rational treatment plan depending on the goals and the clinical situation. In the process of treatment with the help of the bracket system, the patient must undergo correction and check-ups approximately once a month, while, as a rule, the orthodontist

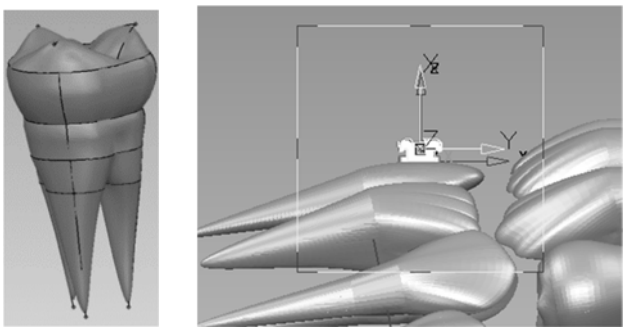
changes ligatures (rubber bands on brackets), arcs and other elements, monitors the dynamics of the displacement of teeth in the process treatment.

Solving the problem of modelling the installation of brackets and arcs can be done using specialized software that is integrated with the CAD system. The main purpose of the simulation is to determine the rational parameters of the mutual positioning (installation) of brackets and the patient's teeth and the loads on the teeth when installing the arch wire of the bracket system, which will make it possible to more accurately predict the behaviour of the entire system.

The choice of methods for modelling the bracket-system elements is determined by their geometry and the formulation of the modelling problem. To do this, we consider the structure and principle of operation of the bracket system.

Power arcs can be of different cross-sections and diameters; arch wire metal of various alloys with special properties are used. The power arc, as a rule, is created from a nickel-titanium alloy, which has the ability of shape memory and due to this, in the oral cavity, the arc tends to return to its original shape and moves the teeth, regardless of the initial deformation.

On the tooth, the vestibular, lingual, medial and chewing surfaces are distinguished, which determine the shape of the tooth and are used for protection [13]. Based on the overall dimensions of the required tooth, we build an approximate layout that will allow us to determine the boundaries of the contour (see **Fig. 16**, a).



a b

Fig. 16. - Construction of the lateral surface of the tooth (a) and the sequence of automated construction of brackets at specified points on the surface of the tooth models (b)

For visual inspection of the arch wire shape, a coordinate system is formed at the end of the end bracket (see **Fig. 16**, b). In this case, the gap relative to the arch wire lock is 0.1 mm. Then, between brackets, a G2-spline is created, which is fixed at two points of the body of each bracket to ensure that there is no intersection of the arc model with the bodies of brackets (see **Fig. 17**).

The SolidWorks Simulation environment was used to solve the problem of determining the stresses in the arch wire during bracket installation. Static analysis (Static) was chosen as the problem type. AISI 304 steel from the SolidWorks standard library was used as the workpiece material.

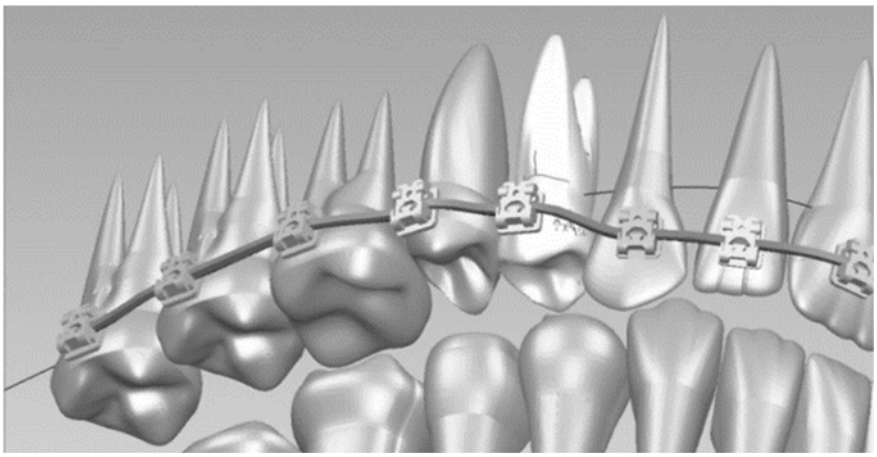


Fig. 17. - The constructed model of the bracket system with the arch wire installed

Rigid supports (fixed geometry) at points b, d, l, n (**Fig. 18**) were used as boundary conditions. At the rest of the points, using a force load of the "Specified displacement" type, displacements were set, which were fixed when the bracket was installed: a – (+0,5) mm; c – (-0,3) mm; e – (-0,2) mm; f – (+0,1) mm; g – (+0,7) mm; h – (+0,2) mm; k – (-0,3) mm; m – (+0,7) mm, where the positive sign corresponds to the movement to the outside, and the negative to the inside.
After solving the problem in Solidworks, the displacements of all points of the workpiece were obtained, which coincided with the specified ones, which indicates the adequacy of the calculation (**Fig. 19**).

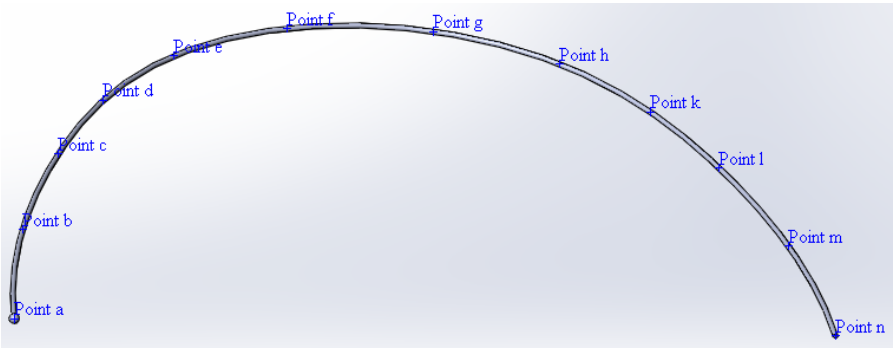


Fig. 18. - Location of control points

A detailed consideration of the loading of teeth in more complex cases, in particular, in the presence of a vertical deflection of the bracket systems arcs, was carried out in [15, 16]. The basic principles of using FEM in orthodontics are presented in [17].

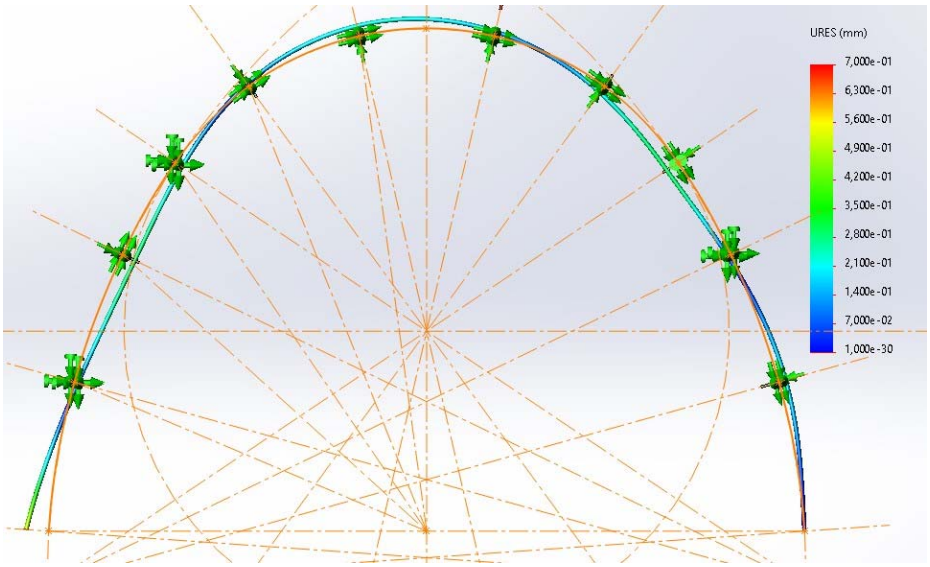


Fig. 19. - Field of workpiece points displacements

Conclusion

The design and modeling of various types of implants was carried out using modern CAD / CAM / CAE systems. A number of examples of designing and calculating the loading of various implants have demonstrated the functionality of modern computer-aided design technologies.

In particular, with the use of CAD / CAM / CAE systems, an end-to-end technological process of deformation of the workpiece was developed through a series of transitions and the manufacture of a U-shaped implant on CNC equipment. The use of FEM, implemented in the CAE-system, provides a distribution of the stress-strain state in the volume of workpieces and the body of patients. Knowledge of these parameters allows you to choose options for implant designs, places of their installation, characteristics of rigidity and strength, taking into account the peculiarities of the geometry, mechanical characteristics of the patient's body, limitations of the applied loads.

Analysis of the design and manufacturing process of implants for maxillofacial surgery made it possible to formalize the process of designing spatial implants from sheet blanks on the basis of tomography data. Modeling allows you to determine the stages and an algorithm for designing workpieces, as well as a universal tool for their receipt.

The design in the CAD system of the geometric model of the patient's teeth was carried out according to the scanning data. The presence of a CAD model allows you to determine the placement of brackets, taking into account the real shape and position of the teeth. Then the simulation is performed in the CAE-system of loads on teeth. The presence of the model ensures the correct choice of material and shape of the arch wire in order to limit the load on the periodontium to the permissible limits.

References

1. Nanostructured severe plastic deformation processed titanium for orthodontic mini-implants / Glaucio Serra [and al]. // Materials Science and Engineering C. – 2013. - Vol. 33 - pp.4197–4202.
2. Titanium in Medicine: Material Science, Surface Science, Engineering, Biological Responses and Medical Applications /Donald M. Brunette, Pentti Tengvall, Marcus Textor, Peter Thomsen // Springer-Verlag Berlin Heidelberg, 2001, 1018 p. ISBN: 978-3-642-63119-1, 978-3-642-56486-4
3. Titanium and titanium alloys: fundamentals and applications / edited by C. Leyens and M. Peters // Wiley-VCH, 2003. - 532 p.
4. Titanium, 2nd Edition (Engineering Materials and Processes) / Gerd Lutjering, James C. Williams // Springer, 2007. - 449 p.
5. Titanium alloys: modelling of microstructure, properties and applications / Wei Sha, Savko Malinov // Woodhead Publishing Ltd, 2009. - 598 p.
6. Spinal stabilization for patients with metastatic lesions of the spine using a titanium spacer / Hans Hertlein [and al]. // Eur Spine J. – 1992. - Vol.1 - pp.131-136.
7. Titanium alloy mini-implants for orthodontic anchorage: Immediate loading and metal ion release/ Liliane S. Morais [and al] // Acta Biomaterialia. – 2007. - Vol. 3 (3) - pp. 331-339.
8. Biomechanical effect of different lumbar interspinous implants on flexibility and intradiscal pressure/ Hans-Joachim Wilke [and al] // Eur Spine J. – 2008. - Vol. 17 - pp.1049–1056.
9. Shik Shim, Chan & Woo Park, Seoung & Lee, Sang-Ho & Jesse Lim, T & Chun, Kwonsoo & H Kim, Daniel. (2008). Biomechanical Evaluation of an Interspinous Stabilizing Device, Locker. Spine. 33. E820-7. 10.1097/BRS.0b013e3181894fb1.
10. Cheng-Chan Lo, Kai-Jow Tsai, Zheng-Cheng Zhong, Shih-Hao Chen & Chinghua Hung (2011) Biomechanical differences of Coflex-F and pedicle screw fixation combined with TLIF or ALIF – a finite element study, Computer Methods in Biomechanics and Biomedical Engineering, 14:11, 947-956, DOI: 10.1080/10255842.2010.501762.
11. New Interspinous Implant Evaluation Using an In Vitro Biomechanical Study Combined With a Finite-Element Analysis / Virginie Lafage, Nicolas Gangnet, Jacques Sénégas, François Lavaste, Wafa Skalli // Spine, 2007. 32(16):1706-1713 DOI: 10.1097/BRS.0b013e3180b9f429.
12. Application of the Finite Element Method in Implant Dentistry / N. Krishnamurthy // Springer-Verlag Berlin Heidelberg, 2008. - 147 p.
13. Computer-Guided Applications for Dental Implants, Bone Grafting, and Reconstructive Surgery / Marco Rinaldi, Scott D Ganz, Angelo Mottola // Elsevier, 2015. - 568 p.
14. Reconstruction of the post-traumatic defect of the lower jaw using modern computer technologies (Preparatory stage) / A.N. Chuiko, D.K. Kalinovsky // Modern dentistry (Suchasna stomatolohiya) – 2013. – Vol. 1. – pp. 76-81.
15. Jin-Gang Jiang , Yi-Hao Chen, Lei Wang, Yong-De Zhang, Yi Liu, Wei Qian**.** Modeling and Experimentation of the Unidirectional Orthodontic Force of Second Sequential Loop Orthodontic Archwire / Applied Bionics and Biomechanics. Hindawi. Volume 2020, Article ID 5786593, 11 pages https://doi.org/10.1155/2020/5786593
16. Jianlei Wu, Yunfeng Liu, Jianxing Zhang, Wei Peng, Xianfeng Jiang. Biomechanical investigation of orthodontic treatment planning based on orthodontic force measurement and finite element method before implementation: A case study. Technology and Health Care 26 (2018) S347–S359 S347 DOI 10.3233/THC-174689 IOS Press.

17. Prasad Konda.1, Tarannum SA. Basic principles of finite element method and its applications in orthodontics. Journal of Pharmaceutical and Biomedical Sciences (JPBMS), 2012, Vol. 16, Issue 16. ISSN NO- 2230 – 7885

Modeling and simulation of prosthetic gait using a 3D model based on perturbation functions

Sergey Vyatkin[1] [0000-0002-1591-3588], Olexandr Romanyuk[2] [0000-0002-2245-3364], Yurii Bezsmertnyi[3][0000-0002-1388-7910], Pavlo Mykhaylov[2] [0000-0001-5861-5970], Roman Chekhmestruk[2] [0000-0002-5362-8796]

[1]Institute of Automation and Electrometry, 630090, Novosibirsk, Ak. Koptyuga, 1, Russia
sivser@mail.ru
[2]Politechnika Lubelska, Lublin, Poland
waldemar.wojcik@pollub.pl
[2]Vinnytsia National Technical University, 21021, Vinnytsia, Khmelnytske shose 95, Ukraine
rom8591@gmail.com, psv@vntu.edu.ua
[3]Scientific Research Institute of Invalid Rehabilitation on the base of Vinnytsia Pirogov National Medical University, 21029, Vinnytsia, Khmelnytske shose 104, Ukraine
bezsmertnyiyurii@gmail.com

Abstract. The work describes the method of construction of a lower limb prosthesis, both below and above the knee. It describes the computer-aided design basis and the digital patient model with which the prosthesis is designed and tested in a fully virtual environment. The virtual patient model is the basis of the entire system, using a biomechanical model that takes into account patient characteristics such as anthropometric indicators. The program uses automated and knowledge-based approaches in order to replace the current development process, mainly manual, with a virtual one. A set of tools is used to design, configure and test the prosthesis. Prosthetic modeling allows configuring and generating a three-dimensional model of the prosthesis. The virtual test environment allows the prosthesis to virtually configure the artificial leg and simulate the patient's posture and movements, checking functionality and configuration. A gait model has been developed. For this purpose, the analogy of the double pendulum with the flesh segment of the lower limb was used, which allows to evaluate the gait model as a system using simple final figures or connections. The system representing association of segments is used, the energy exchange between links of the system is analyzed using the Lagrange method. The association analysis is developed by calculating inverse kinematics, which determines the trajectory and position of the system in order to find the forces that carry out this movement and energy transfer. Taken as a double pendulum, the lower limb has mass and length values based on the weight and height of each part. The obtained values were derived from anthropometric data of the subject.

Keywords: First virtual patient, Lower limb prosthesis, Virtual prototyping, Human modeling, Gait modeling.

1 Introduction

Many information and communications technology tools, such as computer-aided design (CAD) and automated engineering systems, have been developed recently to support the product development process, with the aim of reducing the need for physical prototypes and reducing costs and time. However, the spread of such systems is limited in some areas, especially when the product requires a high level of customization and represents an interface with the human body or its parts. An example would be artificial prostheses, which must be designed according to the shape of a specific anatomical area. This work is emphasized on modular prostheses of the lower limbs, both below the knee (transtibial, TT) and above the knee (transfemoral, TF), realized by assembling modern components in order to obtain maximum comfort and convenience of use [1]. Most of the components are standard, e.g. foot and knee can be selected from the manufacturer's catalogue, while the connection should be implemented based on the patient's anatomy. The connection is the main critical component and is designed and manufactured almost entirely manually, relying heavily on the experience and skills of prosthetic technicians. In addition, the patient plays a key role in the production process, since both standard and specially selected components are selected and projected according to their state of health and anatomical morphology.

Some CAD / CAM prosthetic systems such as Bioshape and Canfit are available on the market. Reverse engineering methods (usually laser scanning) can be used to obtain the external shape of the residual limb from which the connection is made and to modify the basic models stored in libraries. However, they are not integrated with modeling tools such as Finite Element Analysis (FEA) or multi-body systems for checking prosthetic construction. Various works can be found in the literature suggesting that FEA can be used to simulate the behavior of prosthetic components [2-4] and to analyze the interaction between joint and residual limb. Some studies have also demonstrated the feasibility of a fully computer-based connection design process based on the integration of CAD and FEA tools [5], but they are not able to manage all stages of the prosthetic design process in a unique environment and do not provide any assistance to the prosthesis.

This work represents the method of lower limbs prosthetics based on the virtual model of the patient with the use of perturbation functions and theoretical multiple subtraction operation [6]. The main objective was to develop a digital model of an amputated person, which will be used to design and test a prosthesis in a fully virtual environment. In order to achieve the goal of replacing the current process (mainly manual) with a virtual one, several issues were considered: capturing and formalizing the knowledge of orthopedic technicians about this process, obtaining information about the patient and morphology using diagnostic tools, developing comprehensive solutions for designing and testing standard and specially selected components, and using digital methods to simulate the behavior of the prosthesis while walking. The adoption of a digital model to represent the patient is in line with the trend towards multiscale human modeling as a tool for a wide range of applications, from ergonomics to occupational safety and health [7].

The method provides the prosthetic professional with a set of interactive tools for designing, adjusting and testing the prosthesis. It consists of two parts: prosthetic modeling and virtual testing (Fig. 1).

The first part configures and generates a three-dimensional model of the prosthesis, while the second part virtually adjusts the artificial leg and simulates the patient's posture and movements, checking the functionality and configuration of the prosthesis.

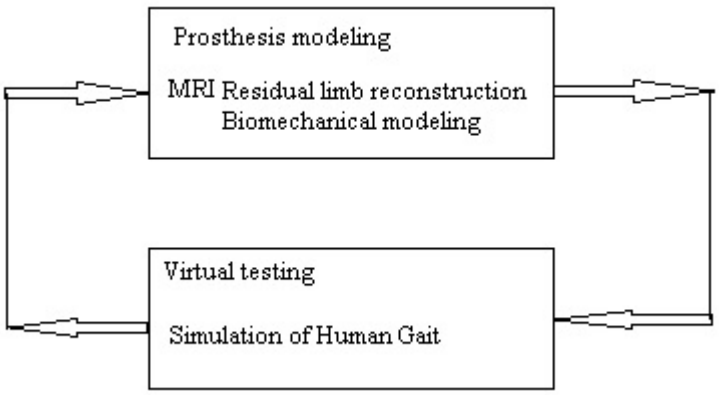


Fig. 1. Virtual patient

MRI is used to scan the patient's leg. Volume data is a three-dimensional array of cubic elements (voxels) representing units of 3D space. At the same time, typical data sets describing physical objects or phenomena are usually represented by files up to several gigabytes in size, since these files contain data about each point of the scanned three-dimensional space. If the interpolated density at the sample point exceeds a certain threshold density, then a density gradient G is calculated from these points *D0,..., D7*, which can be used as a normal to the surface.

An efficient method for converting voxel data arrays into a functional description based on perturbation functions is applied.

Organization

The chapter is organized as follows. Section 2 gives a brief survey of the work related to function representation of objects, rendering and interactive geometric modelling. provides the overview of our approach to computation of the ray-surface intersection. Section 3 describes prosthesis modeling, design of the connection and and configuration of the prosthesis. Virtual testing (the program allows you to set and evaluate the functionality of the prosthesis by simulating the positions and movements of the virtual patient) is presented in Section 4. Analysis of the method described in Section 5. Summary work is discussed in Section 6.

2 Virtual patient

The virtual patient consists of a biomechanical model and a data set. The data are used to adjust the model of the amputated patient. They represent the elements by which the development process is managed, a selection of standard components prior to joining and configuration of the prosthesis. The tasks and decisions made by the prosthetists depend on the characteristics of the patient and their health status. For example, a special kind of foot, identified as a high-energy foot, is suitable for young and healthy patients because it can better maintain high blood pressure. Patient data and information have been grouped into three main categories, namely patient assessment, residual limb assessment and anthropometric indicators.

The first group refers to general patient data, mainly used to select standard components. The second deals with parameters for assessing residual limb condition and creating a three-dimensional joint model. The latter refers to the anthropometric indicators of the patient and the residual limb. They are used for the correct size of the digital patient model, standard components and connection.

The biomechanical model of an amputated person is determined at different levels of detail, depending on the task at hand. For example, compound modeling and simulation require a detailed residual limb model (skin, bone and muscle).

Three main tools are used to create a biomechanical model: a universal human modeling system, medical images of the residual limb (obtained, for example, with MRI or CT) and a special software module for 3D reconstruction of the residual limb.

Perturbation functions are used to create and model the virtual patient [6]. This allows creating a detailed biomechanical model of the human body using rigid links connected through joints to simulate the skeleton and flexible elements to represent muscles, tendons and ligaments. Thus, it is possible to create an individual model modifying anthropometric data. Two reference models were considered, one for amputated TF and one for amputated TT, which should be customized for each individual patient. The following data are needed to characterize the virtual patient: anthropometric patient data and a digital model of the lower limb. This means that the biomechanical model of an amputee is implemented in two stages: the first stage is to determine the model size and the second stage is to determine the residual limb binding [8,9,10].

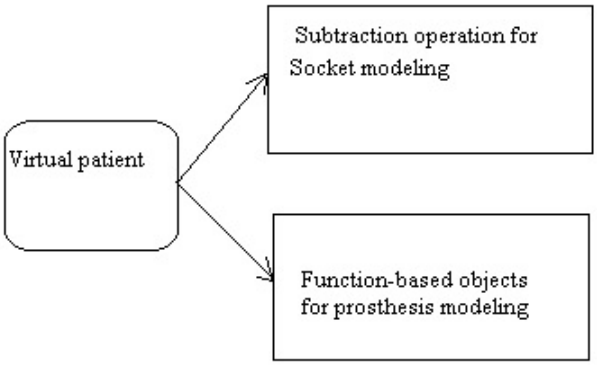


Fig. 2. Prosthesis modeling

2.1 Perturbation functions

Using free forms based on perturbation functions, we model the shape of the prosthesis, using the union operation, we construct the entire prosthesis. Using the subtraction operation, we model the connection socket (Fig. 2).
Characteristic feature of a offered method is that the basic primitive things are chosen as surfaces of the second order - quadrics. The quadric is defined by ten coefficients. The same amount of data is necessary for description of a spatial arrangement of one triangular facet. More compact representation of model allows to reduce memory expenses sharply at database storage and, hence, to unload the bus between computing components of system and memory blocks. On the other hand, it is possible to construct smooth objects without defects inherent in polygonal models. On this basis the class of free forms is under construction. It means possibility to describe complex geometrical objects, setting function of deviation from the second order base surface.

The algorithm of recursive multi-level ray-casting (ray-casting) performs an effective search for elements of volume, voxels, involved in the formation of the image. Let us deal with some object representing a three-dimensional scene, which has the property to respond to the request for intersection with a rectangular parallelepiped, or bar. And a negative answer guarantees that this object has no common points with the bar, and a positive answer implies that possibly common points exist. Only the object's functional abilities are important for the algorithm's work, and we won't concretize its internal structure yet. Let's introduce the coordinate system as shown in Figure 3. The search for voxels is done in the space inside the cube, from -1 to 1 for each coordinate, so that the center of the cube corresponds to the origin of the coordinates. At each step of the recursion, the initial volume is broken down into four bars in the x0y screen plane. For each bar, the object is queried to determine the intersection. If the result is positive, the bar is subjected to the next level of recursion. And the bars that gave a negative result of the test for an intersection with the object are not subject to further

immersion in recursion, which corresponds to an exception from the consideration of square areas of the screen, on which the bar (and therefore the surface of the object) is not displayed [11,13].

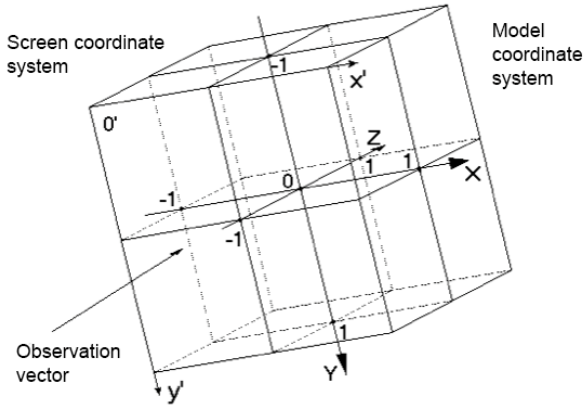


Fig. 3. Unit cube

This stage of the algorithm's work is well illustrated in Fig. 4, where there are empty squares and there are areas of the image not involved in the formation of the image. At a certain step of subdivision we reach bars that are displayed exactly in one virtual pixel of the image. Now we should start subdivision of "thin" bars, rays, along the Z-axis, first performing the request to intersect with the nearest half of the ray. Thus, for each ray at the final level of recursion we will define a voxel containing the surface of the object and closest to the observer. In other words, while searching for volume elements, voxels containing the areas of the object surface forming the image, the given algorithm bypasses the cubic space along the quad tree whose leaves are the roots of binary sub-trees [14,15,16].

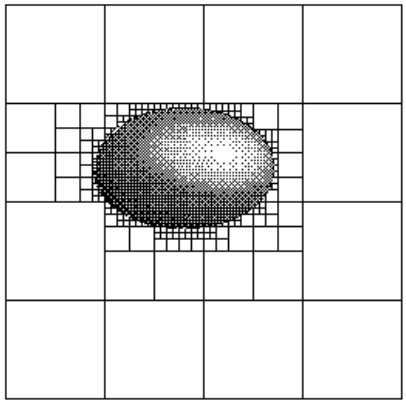


Fig. 4. Quadtree subdivision

The object-square is the basis for building all other objects. The surface forming it is a zero level surface o the square three-dimensional function, i.e. the function of its surface is set implicitly and is defined by ten coefficients $(A,B,C,...K)$:

$$\begin{aligned} &Q(x,y,z) = Ax^2 + By^2 + Cz^2 + Dxy + \\ &+Exy + Fyz + Gx + Hy + Iz + K = 0 \end{aligned}, \tag{1}$$

where x, y and z are spatial variables. You can write this equation in a matrix form:

$$(x,y,z,1)\begin{pmatrix} A & D/2 & E/2 & G/2 \\ D/2 & B & F/2 & H/2 \\ E/2 & F/2 & C & I/2 \\ G/2 & H/2 & I/2 & K \end{pmatrix}\begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix} = 0 \tag{2}$$

We agree to further call the space in which the algorithm performs the division of cubic volume, working or model and mark it with the letter M. Now let's enter the coordinate system directly related to the real space metric (camera or eye coordinate system). Let's allocate in this space the volume limited by the quadrilateral truncated pyramid, or in other words, the visibility pyramid, which is characterized by depth and angle of field of view (Fig. 5). Let us call this space pyramidal and denote it by the letter P [17,18,19].

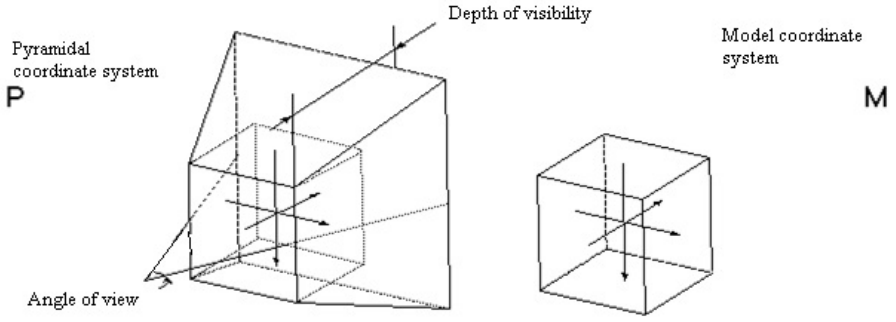


Fig. 5. Visibility pyramid

If the geometrical transformation is described by matrix (C), then the new transformed matrix of coefficients $(Q) = (C)^T (Q)(C)$ operates in space M accordingly to matrix (Q) in space P. In particular, at the stage of space division into a quadric tree, where the compression is performed twice and the transfer to the ± 1 on two coordinates, the recursive transformation of quadric coefficients looks like this:

$$A' = A/4;$$
$$B' = B/4;$$
$$C' = C;$$
$$D' = D/4;$$
$$E' = E/2;$$
$$F' = F/2; \quad (3)$$
$$G' = G/2 + i*A/2 + j*D/4;$$
$$H' = H//2 + i*D/4 + j*B/2;$$
$$I' = I + i*E/2 + j*F/2;$$
$$K' = K + i*G/4 + j*H/4;$$
$$K'' = K' + i*G'/2 + j*H'/2$$

where the odds without a stroke are taken from the previous step of recursion; variables $i, j = \pm 1$ are determined by the algorithm depending on the direction of immersion in recursion (i - in x axis, j - in y axis). The obtained coefficients are used in the intersection test.

If in the equation of the quadric $Q(x, y, z) = 0$ the values of the variables x, y, z change within the interval [-1, 1], then

$$\max[|Q(x, y, z) - K|] \le$$

$$\max F = |A| + |B| + |C| + |D| + |E| + |G| + |H| + |I| ; \quad (4)$$

Now note that if $|K| \pounds \max[|Q(x, y, z) - K|] \le \max F$, then perhaps there is a point $M_0 = (x_0, y_0, z_0)$ $(-1 < x_0, y_0, z_0 < 1)$ such that $Q(x_0, y_0, z_0) = 0$. In case $\max F < |K|$, there are no such points, and the sign of the coefficient K distinguishes the position of the bar inside or outside from the surface of the square $Q = 0$. Based on the results of this test, the algorithm continues to separate only the bars that lie inside the quadric as a whole or, possibly, in part, rejecting deliberately external bars.

The described algorithm displays only the parts of objects that have entered the space inside the cube. The application of projective transformation generalizes the given algorithm into pyramidal volumes, which allows synthesizing images with perspective.

The branch of geometrical knowledge that studies the projective properties of figures is called projective geometry, where homogeneous coordinates are entered to determine the metric in space. In three-dimensional space the point with Cartesian coordinates (x, y, z) corresponds to an infinite set of homogeneous coordinates (x', y', z', a) such that $x = x'/a$, $y = y'/a$, $z = z'/a$, i.e. identical coordinates are determined to the total non-zero multiplier. Of interest is the transformation matrix acting on homogeneous coordinates as follows:

$$\begin{pmatrix} C11 & C12 & C13 & C14 \\ C21 & C22 & C23 & C24 \\ C31 & C32 & C33 & C34 \\ C41 & C42 & C43 & C44 \end{pmatrix} \begin{pmatrix} x_m \\ y_m \\ z_m \\ 1 \end{pmatrix} = \begin{pmatrix} x_p \\ y_p \\ z_p \\ a_p \end{pmatrix}, \text{ or } \quad (C)(M) = (P), \quad (5)$$

where (C) is the transformation matrix, (M) is the homogeneous coordinates of the space point M, and (P) is the display coordinates in P.

Within the limits of projective geometry the theorem [8,§108] that projective display of space M on space P is unequivocally proved is defined by the task of five pairs of points corresponding on display, provided that from five points set in space M, any four do not lie in one plane. We select five pairs of such reference points (M^i) and (P^i) (the upper index corresponds to the number of a pair) and make a system of equations:

$$(C)(M^i) = r^i(P^i), \quad (6)$$

where $i = [1..5], r^1, r^2, r^3, r^4$ - unknown multipliers, but $r^5 = 1$. Solving these equations we find the coefficients of the projective transformation matrix (C) , which is further used to transform the geometric primitives.

The free forms are one more wide class of geometric model objects. The surface of the quadrica is a constant level surface of the second order function. $Q(x, y, z)$. Let's define a certain perturbation function, $R(x, y, z)$, which has the property of smoothness.

Now let's consider the function $Q'(x, y, z) = Q(x, y, z) + R(x, y, z)$. Due to its smoothness R and Q, Q' it will retain this property and therefore the surface of the zero level will also be smooth. Thus, we have built a complex surface using the second order function as the base. But in order to get the necessary shape of the surface, we must learn to set the corresponding perturbation function in an acceptable way. Smooth perturbation is constructed from a square function as follows. Let's define the inner area

Q as part of the space where $0 < Q$. Then in the outer part of Q we put R equal to zero, and in the inner part we erect Q into a cube:

$$R_i(x,y,z) = \begin{cases} Q_i^3(x,y,z), if\ Q_i(x,y,z) \geq 0, \\ 0, if\ Q_i(x,y,z) < 0 \end{cases} \quad (7)$$

We got the function R, which is naturally calculated in the process of recursive subdivision, so that the quad can act as an outrage to other objects. And the object, which we call free form, remains a quadric in its essence but takes into account the additive introduced by the perturbation to its own basic function (Fig. 6). In turn, the free form can be a perturbation function for some other object. Because max $\max[Q+R] \leq \max[Q]+\max[R]$, it means that to estimate the maximum Q on some interval it is necessary to calculate the maximum of the perturbation function on the same interval.

Fig. 6. Free form

The calculation of all components of pixel color is performed similarly to the following formulas:

$$C = (Q_{ambi}C_{ambi} + Q_{diff}C_{diff} + Q_{spec}C_{spec}) / (Q_{ambi} + Q_{diff} + Q_{spec}) \ , \quad (8)$$

where the index "ambi" refers to the characteristics of scattered radiation, and "diff" and "spec" refer to the diffuse and mirror parts of reflected light respectively; C - color components; Q - weighting factors.

Calculations of color components are made on the basis of vector model of illumination [9, Chapt.5]. Four vectors are involved in calculations: normal to surface (n), vector to illumination source (l), direction of reflected light (r) and vector to observer (v):

$$C_{diff} = (n,1)C_{lite}C_{surf} \ ; \quad (9)$$

where C_{lite} - source color, C_{surf} - surface color.

$$C_{spec} = (r, v)^p C_{lite} ; \quad (10)$$

where p - surface roughness coefficient.

To create a complex scene, it is necessary to describe in it some certain quantity of the primitives necessary for a concrete problem. Displayed object, with which the rasterization algorithm interacts by means of queries, is the whole three-dimensional scene. Therefore the geometrical model should allow to construct objects and their compositions of unlimited complexity (Fig. 7.).

It is reached first of all by application of Boolean operations of association and intersection. All scene represents a kind of a tree which each knot is the object-designer who is carrying out logic operations over the descendants, and tree tops are the primitive things used by system. At the moment when the rastering algorithm addresses any query to the constructor object, this object addresses to the descendants, transforms the received result, and gives the corresponding answer to query. In this case, the descendant can be a primitive, or another object-constructor. At application of geometrical operations, turns, movings, scaling, to the object-designer, it makes all these operations with the descendants [20,21,22,23].

Fig. 7. Combining the three free forms

To create a "connection" in free form, a multiple theoretic subtraction operation is applied (Fig. 8).

$$F(x, y, z) = F_1(x, y, z) \setminus F_2(x, y, z) \quad (11)$$

Deducted residual limb volume from free form

$$G_3 = F_i(G_1, G_2) \tag{12}$$

where

$$G_1 : f_1(x, y, z) \geq 0 \quad G_2 : f_2(x, y, z) \geq 0 \tag{13}$$

Fig. 8. Free form and "connection".

Using function-based object and subtraction operation, the prosthetic can model the socket emulating the traditional procedures carried out during socket manufacturing (Fig. 9).

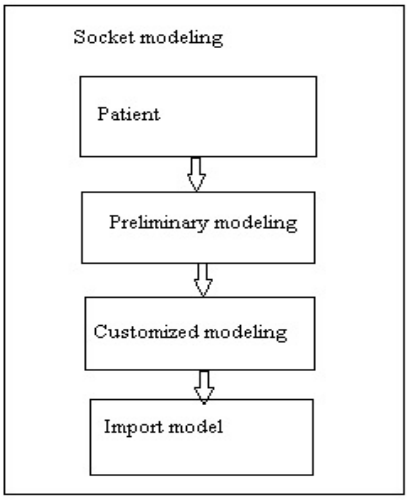


Fig. 9. Virtual socket modeling

2.2 Model size

The correct size of the virtual patient is a key factor in fitting the prosthesis. This means that the anthropometric data of the patient is needed. A distinction is made between general patient data to determine the correct size for the patient and residual limb measurements to position and link the prosthesis to the patient.
Some of these are included or can be derived from the above patient characteristics. Some examples are the hip joint (HJ) and knee joint (KJ), height, weight, and length of the amputated foot. Others, such as shoulder height, can be obtained using a motion capture system.
Once the necessary data has been entered or automatically obtained, the program automatically applies the first level of setting to the virtual patient generating the skeleton, mass, joints and soft tissue based on anthropometric measurements.

2.3 Residual limb attachment

A virtual patient configured with anthropometric data does not include the residual limb relevant for prosthetic design and gait analysis simulation.
A detailed model of the residual limb, including external (skin) and internal (muscles and bones) parts, is based on medical images obtained with an MRI. MRI is preferable to CT because it is less invasive for the patient. However, the end users of the program are podiatrists who do not have special competences and skills in the field of automated tools (for example, tools for medical image processing). Therefore, a software module has been implemented that automatically reconstructs a three-dimensional model of the residual limb without requiring human intervention, starting with the MRI volume [10]. The implemented algorithm first converts the MRI scan into a three-dimensional graph, where the nodes correspond to voxels, and the ribs have weights representing the similarity of neighbors. Aggregate clustering is then performed using features such as intensity and dispersion in regions. Once the graph is segmented in different regions, the algorithm highlights internal bone voxels by analyzing size and shape. The outer surface corresponding to the cult is obtained by a threshold operation. Ultimately, clusters belonging to bones and cult are automatically converted to functional surfaces. The end result is a three-dimensional geometric model, which allows the exchange of CAD information between the program modules.
The residual limb model is imported and linked to the amputated model in two stages:
- The bone segment is first connected to the virtual femur (VF), or to the knee (KJ), using the height of the hip joint and knee joint respectively;
- The residual soft tissue of the limb is then positioned accordingly.
Once an individual virtual patient has been created, the designer can start designing a prosthesis based on the biomechanical model and patient characteristics.

3 Prosthesis modeling

The modeling allows the creation of a three-dimensional model of the prosthesis, which is crucial for the practical study of the installation of the prosthesis and walking of the patient. The modeling and selection of standard and specially selected components is based on the digital layout of the respective anatomical area, the characteristics of the patient (e.g. anthropometric data) and the level of use of the prosthesis. It integrates three basic modules:

- Simulations of the connection (2.1.6).
- A finite element simulation program to analyze the interaction between the stump and the "compound".

3.1 Design of the "connection"

The first two modules are combined, the interface is implemented using an object-oriented language C++.

Using the simulation program, a "connection" (2.1.6) can be created in the denture by simulating the traditional procedures performed during fabrication of the "connection". It is controlled either automatically or semi-automatically and includes a set of design rules derived from the analysis of the traditional process, such as where and how to change the shape of the "connection" or automatically determine the thickness of the "connection" based on the patient characteristics.

Four main stages of the controlled simulation procedure are implemented: patient history, pre-modeling, individual simulation, and final simulation of the "connection" on which other specific modifications will be applied to achieve functional form. The main operations at this stage are performed almost entirely automatically according to the patient's characteristics and the traditional process.

In the next step, the model "connection" is formed directly on the digital residual limb model to be perfectly adjusted to the anatomy of each individual patient. The program provides interactive tools [11] that allow emulating tasks traditionally performed by a technician. It starts to change specific zones by simulating the operations of adding and removing layers. For this purpose, these areas are divided into two categories:

- Loading areas where there are no bone protrusions or tendons and it is necessary to squeeze the "joint" closer to the limb and create pressure to withstand the body weight.
- Unloading zones, where there are no bone protrusions or tendons and the "joint" should not press on the limb and at the same time should not be too wide, because it can cause other physical problems.

The program also offers the correct percentage of material to be added or removed according to the "tone parameter". The designer can decide to automatically apply the values calculated by the program or change the shape with a virtual tool [11]. It is an interactive deformation tool that allows to perform operations of adding and removing material, as the technician manipulates the plaster cast to be emulated. For example, in

the unloading zones the technician adds material from the plaster cast, because the "connection" should not press on the stump and can be quite loose, while in the loading zones the material should be removed to have a tight "connection".
As for the amount of material to be added or removed, eight levels of manipulation with a thickness between 1 and 8 mm are highlighted, which correlate with the stump tone.
Finally, the designer forms the top edge automatically or semi-automatically and the program offers the final thickness of the "connection". As a rule, the prosthesis determines the thickness empirically based on the weight of the patient. The following formula is used in the program:

$$S_t = \frac{P_w}{20} \quad , \tag{14}$$

where S_t - joint thickness (mm), P_w - patient's weight (kg).

A more sophisticated algorithm based on engineering knowledge and mechanical properties of materials can replace this formula.
After modeling a "compound", a simulation is automatically performed to analyze the pressure distribution, which is considered the most important evaluation parameter. Consideration is given to the pain threshold (minimum pressure that causes pain) and pain tolerance (maximum tolerable pressure) for the most important zones [12].
Some modeling assumptions and rules have been adopted:
- Residual limb geometry and "connection". Bones and soft tissues are combined to create a unique piece without geometrical breaks.
- Characteristics of the material. The mechanical properties of joint, bone and residual limb are considered linear, homogeneous and isotropic.
- Boundary conditions and loads. The upper surface of the residual limb is limited. In the first two stages of the analysis no external load is applied, while in the last stage the weight of the amputee is applied to the center of mass of the "connection" in a vertical direction. The movement of the "coupling" and the static constant load are applied gradually in the analysis stages. The interaction between the stumps and the "connection" is simulated using an automated surface-to-surface contact algorithm [13].
- The stages of analysis. The simulation is carried out in three stages: application of the residual limb with prestressing on the residual limb, the adjustment phase to achieve better movement of the "connection" around the residual limb and to achieve maximum comfort, and application of a constant static load (amputated weight).
The program extracts the input analysis from the patient parameters (e.g. patient weight and residual limb length), frees up the files needed to generate the FE model, and selects the scenario (language-specific C++) for analysis. The FE model is created automatically without human intervention. The prosthetist cannot change the characteristics of the FE model, but if necessary, the program allows you to set some parameters, such as material properties. The program then provides analysis and generates an output file containing pressure values that are imported and visualized. The program evaluates the

pressure distribution and highlights the zones to be changed using interactive tools. The program then performs the simulation again until satisfactory results are achieved.

3.2 Selection of standard components and configuration of the prosthesis

A special program is used to select standard components and create the final assembly of the prosthesis. The prosthesis program selects the most appropriate standard components for the individual patient and the program offers possible configurations of the entire prosthesis according to the characteristics of the patient.
To simulate a prosthesis, the lower limb prosthesis is divided into modules as follows:
- Connection module: includes inlay, connection and adapters.
- Double adapter: Includes double male or female pyramidal adapters that connect the "joint" and knee in the TF denture and can replace the support in both the TT and TF dentures.
- Knee module (TF amputated only): includes knee prostheses and knee adapters.
- Tube module: includes a connecting pylon and tube adapters.
- Foot module: includes a foot prosthesis, leg adapters and a heel, also called a "virtual heel". Two key issues were considered: component modelling/calibration and component selection.

Calibration and selection rules were extrapolated from commercial catalogues provided by the main prosthetic brands. Since our goal is to correctly assemble and validate the virtual prosthesis and perform virtual gait analysis, a library was developed containing three-dimensional parametric models for each module. Only those characteristics that affect static and dynamic behavior such as joint weight and position are considered, not the actual shape and appearance of the components. The reference dimensions were taken from real and commercial catalogues. Particular attention was paid to the foot and knee as they are standard components responsible for the behaviour of the entire prosthesis and their selection is a key factor in achieving a satisfactory configuration. There is a huge variety of components on the market, and each manufacturer has its own model. For this reason, instead of taking into account specific models that are subject to frequent variations, a set of stable categories has been identified. As an example for a patient with a TF reputation, the main knee joint typologies have been grouped as follows:
- Fixed: knee with locked center of rotation during walking, which can be unlocked manually if necessary.
- Monocentric: elbows with one centre of rotation, separated for self-braking, friction, pneumatic and hydraulic operation.
- Polycentric: knees with a large number of rotation centres are also divided into self-braking, friction, pneumatic and hydraulic modes.

To select the right components, two setting procedures (one for TT and one for TF) and a database to automatically select the right components for each type of amputated limb and their respective size.

The program automatically collects all possible combinations of different selected parts and provides the specialist with all relevant specifications. The user can select the most suitable one or modify some components according to the patient's needs. When assembling the components, the system ensures that the alignment of the prosthesis is similar to the skeletal structure of the other leg. Traditionally, this operation is called plumb alignment [24,25,26].

4 Virtual testing

The program allows you to set and evaluate the functionality of the prosthesis by simulating the positions and movements of the virtual patient using the method [14].
First of all, the virtual patient must wear the assembled prosthesis. This is imported from the virtual simulation program and correct positioning is obtained with regard to the prosthesis height and the rotation of the foot relative to the vertical line. In particular, the prosthetic leg must be aligned with the other and the "joint" must fully support the residual limb.
The amputated virtual patient can then be used to perform static alignment and gait analysis during various activities. The basic idea is to provide the prosthetist with a movement law library specialized for patients wearing a prosthesis. This requires learning several movements and poses of the patient during typical daily activities and then deriving the laws of movement for the artificial jointsc [27].
This approach was tested using a motion capture system without markers to determine the laws of motion pertaining to the patient's joint. A data collection program consisting of the following components is used: four cameras, resolution 640 × 480 pixels at 60 Hz, markerless motion capture, computer with processor Intel Core2 CPU E8400 3.0 GHz, and a graphic accelerator GeForce 8800 GTX.
The decision taken is inexpensive and does not require the patient to wear markers because it is based on image and silhouette analysis [15]. It automatically recognizes the different segments of the body and then calculates the position and orientation in three-dimensional space. In this environment, the patient must perform typical daily activities such as walking, sitting, walking up/down, and so on. Obviously, the accuracy of the tracking data is crucial, and the quality of the webcam solution is tested.
In order to replicate the movement with the virtual patient, two issues need to be considered: data conversion and display of the received data on the virtual patient defined in the previous step. Two software modules have been developed that perform these tasks automatically.
In order to achieve accurate simulation of muscle and joint movement, reverse dynamic simulation is first performed to record angles and muscle contraction history for the target body segments (links). "Movement agents are located on the model to control movement and 'teach' joints and soft tissues how to move. In our case, in order to reproduce the functionality of the residual limb, we introduced "augmented movement

agents" associated with the prosthetic segments: three associated with the prosthetic leg, one with a tube below the knee representing the lower part of the leg, one with the knee (for TF) and one with a "connection". Once the movements have been recorded in the reverse dynamic simulation, the compiled movement histories are ready to control the forward dynamic simulation.

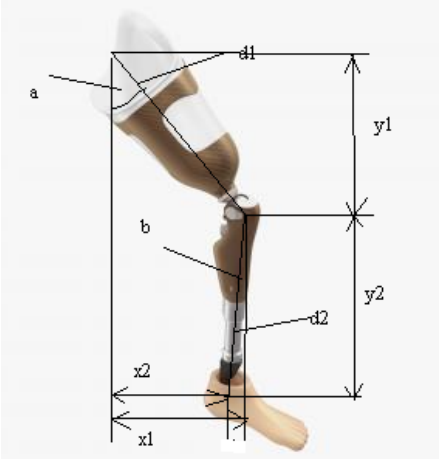


Fig. 10. Lower limb model is set to double pendulum. The lower limb model is used to obtain the gait model. Two compound pendulum segments corresponding to the femur and tibia respectively are considered.

The values d1 and d2 are defined as the lengths of the femur and tibia (Fig. 10). The femur corresponds to the femur segment and the tibia to the leg segment. The position of the femur corresponds to the x and y coordinates in the equations.

$$x_1 = d_1 \sin\alpha \tag{15}$$

$$y_1 = d_1 \cos\alpha \tag{16}$$

The task is to find the energy relations of the system, and for this purpose it is necessary to determine the components of position, speed and acceleration. Concluding (15) and (16) with respect to time, the velocity components are obtained as equations (17) and (18).

$$\dot{x}_1 = d_1 \dot{\alpha} \cos\alpha \tag{17}$$

$$\dot{y}_1 = d_1 \dot{\alpha} \sin\alpha \tag{18}$$

The total speed for the femur is equal to the sum of equations (17) and (18), resulting in equation (19).

$$vms^2 = \dot{x}_1^{\,2} + \dot{y}_1^{\,2} = d_1 \dot{\alpha} \cos^2 \alpha + d_1 \dot{\alpha} \sin^2 \alpha = d_1^2 \dot{\alpha}^2 \qquad (19)$$

The kinetic energy of the contralateral thigh is determined by the equation (20).

$$E_{k1} = \frac{1}{2} m_1 vms^2 \qquad (20)$$

Let's substitute (19) in (20) with the expression for hip kinetic energy, equation (21).

$$E_{k1} = \frac{1}{2} m_1 d_1^2 \dot{\alpha}^2 \qquad (21)$$

Equation (22) represents the potential energy for the thigh.

$$E_{p1} = -m_1 g y_1 \qquad (22)$$

Substitution (16) in (22) gives equation (23), which corresponds to the potential thigh energy.

$$E_{p1} = -m_1 d_1 g \cos \alpha \qquad (23)$$

The above described mathematical procedure used for the hip is repeated for the leg. The position of the leg is determined by equations (24) and (25).

$$x_2 = x_1 + d_2 \sin \beta = d_1 \sin \alpha + d_2 \sin \beta \qquad (24)$$

$$y_2 = y_1 - d_2 \cos \beta = -d_1 \cos \alpha - d_2 \cos \beta \qquad (25)$$

The speed components are determined by equations (26) and (27)

$$\dot{x}_2 = x_1 + d_2 \dot{\beta} \cos \beta = d_1 \dot{\alpha} \cos \alpha + d_2 \dot{\beta} \cos \beta \qquad (26)$$

$$\dot{y}_2 = y_1 + d_2 \dot{\beta} \sin \beta = d_1 \dot{\alpha} \sin \alpha + d_2 \dot{\beta} \sin \beta \qquad (27)$$

The full speed for the leg is determined by equation (28).

$$vps^2 = \dot{x}_2^{\,2} + \dot{y}_2^{\,2} = (d_1 \dot{\alpha} \cos \alpha + d_2 \dot{\beta} \cos \beta)^2 + \left(+ d_1 \dot{\alpha} \sin \alpha + d_2 \dot{\beta} \sin \beta \right)^2 \qquad (28)$$

Equation (29) corresponds to kinetic energy for the leg, and equation (30) refers to kinetic energy for the leg.

$$E_{k2} = \frac{1}{2} m_2 vps^2 \tag{29}$$

$$E_{k2} = \frac{1}{2} m_2 (d_1^2 \dot{\alpha}^2 d_2^2 \dot{\beta}^2 + 2 d_1 d_2 \dot{\alpha} \dot{\beta} \cos(\alpha - \beta)) \tag{30}$$

Equation (31) describes the potential energy of the leg.

$$E_{p2} = -m_2 (d_1 g \cos\alpha + d_2 g \cos\beta) \tag{31}$$

The total kinetic energy of all segments of the body, thigh and tibia is presented in equation (32).

$$E_{kt} = E_{k1} + E_{k2} \tag{32}$$

$$E_{kt} = \frac{1}{2} m_1 d_1^{\,2} \dot{\alpha}^2 + \frac{1}{2} m_2 (d_1^2 \dot{\alpha}^2 d_2^2 \dot{\beta}^2 + 2 d_1 d_2 \dot{\alpha} \dot{\beta} \cos(\alpha - \beta))$$

The total potential energy for the entire lower segment of the facility is shown in equation (33).

$$E_{pt} = E_{p1} + E_{p2} \tag{33}$$

$$E_{pt} = -m_1 d_1 g \cos\alpha - m_2 (d_1 g \cos\alpha + d_2 g \cos\beta)$$

After obtaining the final energy values, Lagrangian is calculated in equation (34).

$$L = E_{kt} - E_{pt} \tag{34}$$

Equation (35) is the Lagrangian equation of motion, and equation (36) is for β .

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{\alpha}}\right) - \frac{\partial L}{\partial \alpha} = 0 \tag{35}$$

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{\beta}}\right) - \frac{\partial L}{\partial \beta} = 0 \tag{36}$$

5 Results

A comprehensive program check is quite complex. In addition, it requires the active participation of prosthetics and patients.

At present, the proposed program has been tested in three stages at various levels to gradually test the new design process and the development of automated tools. The first

step concerns the prosthetic configuration, the second step concerns the prosthetic simulation process and the last step concerns the entire design process, from simulation to testing.
Testing was done on a computer with a processor Intel Core2 CPU E8400 3.0 GHz, and a graphic accelerator GeForce 8800 GTX.

5.1 Prosthetic Configuration

The goal was to check the validity of the knowledge-based configuration procedures and the selection and calibration rules adopted for standard components. Several amputated TF and TT models were considered [16-18]. Several patients were selected to cover a variation of all parameters. Four parameters required for the selection procedure were collected for each patient: patient weight, lifestyle factors, residual limb length and patient strength. As soon as the data were obtained, the program automatically selected the most appropriate foot and knee prostheses (for TF). The program is designed to facilitate selection by providing a list of the most effective components. This solution is preferable to a single result, as some other factors (e.g. aesthetic or cultural aspects) may influence the final choice. The configurations proposed by the programme have been compared with those identified through the traditional process.
Test cases for which no suitable matches were found were explained by the fact that some patient characteristics could not be quantified. For example, for a middle aged patient who is in good health but not motivated, a less effective prosthesis is preferable because it will be easier to use and shorten training time. Another example concerns the aesthetic aspect of components: Some people may prefer a small component, while the system may choose a larger but more effective component.

5.2 Prosthetic modeling

In this case, the purpose of the experiments was to test the prosthetic modeling approach and the results were mostly qualitative. Two cases were examined: amputated TT, 41 years old, height 173 cm, and TF, 47 years old, height 176 cm.
First, the "connection" was simulated and then the entire prosthesis was adjusted.
The geometry of the residual limbs of both amputated limbs was obtained and reconstructed; then patient data were entered, a three-dimensional model of the "connection" was created according to the procedure and rules proposed by the program, according to the characteristics of the patient. This is particularly important because most operations depend on the patient's characteristics, such as changing critical zones (loading and unloading zones), which significantly affect the shape of the "connection". The main objective was to verify the effectiveness of simulation/strain tools and the feasibility of a knowledge-based approach to designing a product, a "connection", that is precisely tuned and characterized by a strict interaction with the human body.

After model creation of "connection" (Fig. 6) the program automatically starts modeling and visualizes results. The program evaluates the pressure values of the critical areas and offers the necessary modification. The simulation results for the amputated TT show a uniform and consistent pressure distribution, except for the medial region of the tibia. To reduce the pressure in the medial region of the tibia (the "loading area"), the program automatically changes the geometry of the "connection" in this area and the prosthesis can also change it using virtual tools. The program then performs the simulation again.

The next step was to adjust the entire prosthesis. Based on patient data, the system provides the specialist with the appropriate components and adapters, dimensioning them and assembling all the parts. For example, for a patient with TF, the program automatically offered two types of feet (mono-axial high-energy and multi-axial sandwich feet) and two elbows (mono- and polycentric pneumatic knee). Once the respective components had been confirmed, the program extracted three-dimensional parametric models of the standard components from the database and assembled them together with the "connection" to obtain a virtual prototype of the full denture.

Some modifications are required to facilitate the use of simulation tools, especially for users who do not own a computer. For this purpose, the use of new interaction tools, such as a tactile device, to make simulations of "connections" more natural is evaluated.

5.3 Projecting: from modeling to testing

This test has allowed to check feasibility of all process, from modeling to the virtual approach to testing.

The entire design process was tested for an amputated TF. A virtual patient model with anthropometric indices and residual limb morphology was created. The prosthesis was then configured, simulated and tested. A virtual patient wearing an artificial limb was generated and two typical situations were simulated: patients walk on a flat floor and climb a step.

The preliminary results are considered promising, but further programme improvements are needed.

6 Conclusion

This paper presents virtual patients and a software platform designed to design and configure both TF and TT dentures using only digital models and virtual instruments. The digital patient is the core of the entire program around which the prosthetist designs and tests the virtual prosthesis, following the rules during the development of the proposed product. Experiments were conducted with a variety of test cases and with the involvement of highly specialized orthopaedic clinicians in order to get feedback directly from end users, check the effectiveness and usability of the instruments and check the suitability of the digital patient.

Technical experts evaluated the simulation tools for their ease of use. The configurations offered by the program are consistent with those obtained through traditional procedures, with the advantage that different configurations can be generated and compared more easily. The residual limb model has been found to be adequate either to simulate the connection or to simulate the connection and residual limb interaction. With regard to the latter topic, experimental work will be carried out on the best characteristics of material properties by means of indentation tests and obtaining real values of pressure in critical areas of the human part under consideration to test the given FE model. We also consider the possibility of introducing tactile interaction tools to support the simulation phase of compounds and make the procedure of compounds formation more natural and similar to the traditional one. Preliminary tests were performed using a low-cost haptic device. The results are interesting and a more sophisticated device is planned, taking into account the inexpensive solutions.
The approach to virtual testing and using a digital patient of a real person wearing a prosthetic device is promising, but further development and improvement are needed to make it easily applicable. The proposed program can reduce the number of prototypes and the psychological impact on the patient's life. The automated approach allows for all stages of the traditional "compound" development process.

References

1. Smith D. G., Michael J. W., Bowker J. H. (eds). 2004. Atlas of Amputations and Limb Deficiencies: Surgical, Prosthetic, and Rehabilitation Principles. Rosemont, IL: American Academy of Orthopedic Surgeons.
2. Frillici F. S., Rissone P., Rizzi C., Rotini F. The Role of Simulation Tools to Innovate the Prosthesis Socket Design Process. In Intelligent production machines and systems" (eds D. T. Pham, E. E. Eldukhri, A. J. Soroka), pp. 612–619. 2008. Dunbeath, UK: Whittles Publishing.
3. Dou P., Jia X., Suo S., Wang R., Zhang M. Pressure Ddistribution at the Stump/Socket Interface in Transtibial Amputees During Walking on Stairs, Clinical Biomechanics, Vol: 21, Issue 10. pp. 1067-1073 (2006).
4. Faustini M. C., Neptune R. R., Crawford R. H. The Quasi-Static Response of Compliant Prosthetic Sockets for Transtibial Amputees using Finite Element Methods. Medical Engineering & Physics, Volume 28, Issue 2, pp.114-121(2006).
5. Goh J. C., Lee R. S., Toh S. L., Ooi C. K. Development of an Integrated CAD-FEA Process for Below-Knee Prosthetic Sockets. Clin. Biomech.20, pp. 623–629 (2005).
6. Vyatkin S., Romanyuk A, Nechiporuk M., Romanyuk O., Troianovska T. Visualization of volumetric data and functionally defined surfaces using graphics processing units. CONFERENCE 2019.Big Data Processing: Methods, Models and Information Technologies. Series: Applied Research in Computer Science. pp. 216-228. Edited by Oleg I. Pursky. Shioda GmbH, Steyr, Austria, (2019).

7. McFarlane N. J. B., Lin X., Zhao Y, etc. Visualization and Simulated Surgery of the Left Ventricle in the Virtual Pathological Heart of the Virtual Physiological Human. Interface Focus, Volume 1, Issue 3, pp. 374–383 (2011).
8. Woods F. Higher Geometry. Publisher: Applewood Books, 2012.
9. Rogers D. Algorithmic foundations of computer graphics. Mir, Moscow, 1989.
10. Vyatkin S. I, Romanyuk A. N, Krupoderova L. M. Method of three-dimensional compression of volumetric data. CONFERENCE 2020. Actual problems of science and practice, Abstracts of XIV International Scientific and Practical Conference, Stockholm, Sweden, pp. 290-293 (2016).
11. Vyatkin Sergei. I., Romanyuk Olexander N, Al-Maitah Mohammed, Romanyuk Oksana V, Nykiforova Larysa E, Sawicki Daniel, Demsova Natalia. Function-based interactive editing of decoration and material properties. Proceedings SPIE 10808, Photonics Applications in Astronomy, Communications, Industry, and High-Energy Physics Experiments 2018. Vol. 10808.
12. Lee W. C., Zhang M., Mak A. F. Regional Differences in Pain Threshold and Tolerance of the Transtibial Residual Limb: including the effects of age and interface material. Arch. Phys. Med. Rehabil. 86, pp. 641-649.
13. Vyatkin S. I., Romaniuk O. N., Kyrylashchuk S. A., Nechiporuk M. L. Animation of three-dimensional objects using iterative methods. International periodic scientific journal. Modern engineering and innovative technologies. Is. 7, Part 1. - Karlsruhe, Germany. pp.. 59 – 62 (2019).
14. Vyatkin Sergey, Romanyuk Alexander, Romanyuk Oksana, Nechyporuk Mykola, Troyanovskaya Tatiana and Tsikhanovska Olena. Photorealistic object reconstruction using perturbation functions and features of passive stereo projection. CONFERENCE ACIT' (2020).
15. Bezsmertnyi Y.O., Shevchuk V.I., Grushko O.V., Tymchyk S.V., Bezsmertna H.V., Dzierżak R, et al. Information model for the evaluation of the efficiency of osteoplasty performing in case of amputations on below knee. Proc. SPIE 10808, Photonics Applications in Astronomy, Communications, Industry and High-Energy Physics Experiments 2018, 108083H (2018).
16. Shevchuk V.I., Bezsmertnyi Y.O., Kyrychenko V.I. M'iazova plastyka pid chas amputatsii i reamputatsii homilky [Muscular plastic during amputation and abdominal shaft. Orthopedics, traumatology and prosthetics]. Ortopedija, travmatologija i protezirovanie. 4, (2010), pp. 13-18 9 (2010). (in Ukrainian).
17. Shevchuk V.I., Bezsmertnyi Y.O., Maiko V.M. Kistkova plastyka pry amputatsiiakh ta reamputatsiiakh nyzhnikh kintsivok [Bone plastics with amputations and reamputs of the lower extremities]. Ortopedija, travmatologija i protezirovanie. 1, (2011), pp 47-55 (2011). (in Ukrainian).
18. Yurii O. Bezsmertnyi, Viktor I. Shevchuk, Sergii V. Pavlov, "Prognosis of efficacy of medical and social rehabilitation in disabled individuals with respiratory diseases", Proc. SPIE 11176, Photonics Applications in Astronomy, Communications, Industry, and High-Energy Physics Experiments 2019, 1117633 (6 November 2019); https://doi.org/10.1117/12.2537340.
19. Yurii O. Bezsmertnyi, Sergii V. Pavlov, and etc. "Information model for forecasting of violation reparative osteogenesis of long bonds", Proc. SPIE 11176, Photonics Applications in Astronomy, Communications, Industry, and High-Energy Physics Experiments 2019, 111762A (6 November 2019); https://doi.org/10.1117/12.2536250.
20. Sergey I. Vyatkin, Olexander N. Romanyuk, Sergii V. Pavlov, and etc. "Transformation of polygonal description of objects into functional specification based on three-dimensional

patches of free forms", Proc. SPIE 11176, Photonics Applications in Astronomy, Communications, Industry, and High-Energy Physics Experiments 2019, 1117622 (6 November 2019); https://doi.org/10.1117/12.2537043.

21. Sergey I. Vyatkin, Olexander N. Romanyuk, Sergii V. Pavlov, and etc. Offsetting and blending with perturbation functions // Proc. SPIE 11045, Optical Fibers and Their Applications 2018, 110450W, 2019; doi: 10.1117/12.2522353; https://doi.org/10.1117/12.2522353.
22. Sergey I. Vyatkin, Olexander N. Romanyuk, Sergii V. Pavlov, and etc. A GPU-based multivolume rendering for medicine // Proc. SPIE 11045, Optical Fibers and Their Applications 2018, 1104513, 2019); doi: 10.1117/12.2522408.
23. Sergey I. Vyatkin, Olexander N. Romanyuk, Sergii V. Pavlov, and etc. Offsetting and blending with perturbation functions // Proc. SPIE 10808, Photonics Applications in Astronomy, Communications, Industry, and High-Energy Physics Experiments 2018, 108082Y, doi: 10.1117/12.2501694..
24. Sergey I. Vyatkin, Sergii A. Romanyuk, Sergii V. Pavlov, and etc. Using lights in a volume-oriented rendering // Proc. SPIE 10445, Photonics Applications in Astronomy, Communications, Industry, and High Energy Physics Experiments 2017, 104450U; doi: 10.1117/12.2280982..
25. Oleg G. Avrunin; Maksym Y. Tymkovychhttp://profiles.spiedigitallibrary.org/summary.aspx?DOI=10.1117%2f12.2229040&Name=Sergii+V.+Pavlov; Sergii V. Timchik; Piotr Kisała, et al. Classification of CT-brain slices based on local histograms, *Proc. SPIE* 9816, Optical Fibers and Their Applications 2015, 98161J (December 18, 2015); doi:10.1117/12.2229040.
26. Olexander N. Romanyuk; Sergii V. Pavlov; Olexander V. Melnyk; Sergii O. Romanyuk; Andrzej Smolarz, et al. Method of anti-aliasing with the use of the new pixel model, *Proc. SPIE* 9816, Optical Fibers and Their Applications 2015, 981617 (December 18, 2015); doi:10.1117/12.2229013.
27. S. O. Romanyuk, S. V. Pavlov, O. V. Melnyk. New method to control color intensity for antialiasing. Control and Communications (SIBCON), 2015 International Siberian Conference. - 21-23 May 2015. - DOI**:** 10.1109/SIBCON.2015.7147194.