ADVANCES IN BIOMEDICAL ENGINEERING

JOLANTA PAUK
AGNIESZKA DARDZIŃSKA-GŁĘBOCKA
PIOTR MROZEK
ZBIGNIEW OKSIUTA

Politechnika
Białostocka

ADVANCES IN BIOMEDICAL ENGINEERING

Editors
Jolanta Pauk, Agnieszka Dardzińska-Głębocka,
Piotr Mrozek, Zbigniew Oksiuta

Politechnika
Białostocka
OFICYNA WYDAWNICZA POLITECHNIKI BIAŁOSTOCKIEJ
BIAŁYSTOK 2022

Reviewers:
Assoc. Prof. Kristina Daunoravičienė, PhD, Vilnius Gediminas Technical University
Assoc. Prof. Paweł Kostka, PhD, DSc, Eng., Prof. of SUT

Science editor in the discipline of biomedical engineering:
Prof. Jan Ryszard Dąbrowski, PhD, DSc, Eng.

Copy editor:
Emilia Kiszycka

DTP & cover of a book:
Marcin Dominów

ISBN 978-83-67185-40-0
ISBN 978-83-67185-49-3 (eBook)
DOI: 10.24427/978-83-67185-49-3

Printing: PPH Remigraf sp. z o.o.

Publishing House of Bialystok University of Technology
Wiejska 45C, 15-351 Białystok
e-mail: oficyna.wydawnicza@pb.edu.pl
www.pb.edu.pl

Contents

Introduction

Over the thirty years, as the discipline of biomedical engineering has evolved at the Bialystok University of Technology, it has become clear that it is a youthful and developing interdisciplinary field. Biomedical engineers provide innovative technological solutions for the medical industry by designing prosthetic limbs and artificial organs, the material that is used to manufacture them, and the software that's used for medical equipment.

Although it is not possible to cover all of the biomedical engineering domains in this book, we have made an effort to focus on most of the major fields of activity in which scientists at Bialystok University Technology are engaged. The chapters throughout the book feature current research and developments in, for example, biomechanics, mechanobiology, design of medical structures, technical rehabilitation measures, modern engineering materials and their production, biomedical signal processing, acquisition and exploration of information from medical databases, and medical 3D printing.

This book is dedicated to students of biomedical engineering and high school students interested in studying the interdisciplinary course of Biomedical Engineering. Our aim is to help young people understand the nature of biomedical engineering, which links engineering techniques with biological sciences and medicine to improve the quality of human health and well-being.

DOI: 10.24427/978-83-67185-49-3_1

Chapter 1

Implants of the temporomandibular joint

Piotr Borkowski, Łukasz Chwiedosik

Bialystok University of Technology, Institute of Biomedical Engineering,
e-mail: p.borkowski@pb.edu.pl, luukess@gmail.com

Abstract: Disorders of the temporomandibular joint are the result of degenerative diseases of the musculoskeletal system associated with morphological and functional deformities. Joint reconstruction is indicated in patients in whom conservative treatment has failed and a significant proportion of the joint has been lost. The use of total implants reconstructing the head and acetabulum has become a standard. Also, an individual design of plates adjacent to the bone is becoming an accepted norm. The developed procedures include the process of obtaining data using computed tomography, individual implant fitting and verification engineering analyses. The paper presents an example of an individual design process for a temporomandibular joint implant.

Keywords: temporomandibular joint, reconstruction, custom made implants

1.1. The temporomandibular joint

The temporomandibular joint (TMJ) consists of two bone elements, the mandibular fossa and the mandibular head. Both of these elements are separated from each other by the intra-articular cartilage that divides the joint into the so-called two floors surrounded by an articular capsule reinforced with ligaments [1].

Temporomandibular joints are functionally and anatomically linked with each other, thanks to mandibular bones articular heads. Therefore, they are different from the rest of the joints in the human body. The main stimuli that affect suction include pressure on the articular surfaces by the disc and mandibular bone heads, contractions and changes in the length of muscle fibers, and muscle tension [1, 2, 3]. Pathogenic factors, and more precisely their effect on the temporomandibular joint, may cause inflammatory processes not only in adulthood, but also in the child's development. Inflammation can affect individual parts of the joint, the entire joint, or the joint and the muscular apparatus. The harmful effect of pathogens may disrupt the proper development of temporomandibular joints [4].

Implants are more and more often used to restore the proper functions in the temporomandibular joint. Quite a large percentage of patients qualified for implantation have distorted anatomy of the temporomandibular joint due to previous surgical interventions or due to pathological conditions of the joint. Such situations additionally increase the difficulty of performing a stable reconstruction with the use of standard implant components. Contraindications for performing surgical procedures with the use of a temporomandibular joint implant are chronic infection in the site after surgery and hypersensitivity related to the material from which the implant was made. The most commonly used biomaterials for implants are titanium and cobalt-chrome alloys for the condyle, and high molecular weight polyethylene for the acetabular insert [5, 6].

Due to the development of craniofacial radiology, prostheses of the temporomandibular joint are more and more often adapted to the patient's anatomy, which allows for a more precise adjustment of the implant to the mandibular bone and better adaptation of the implants. Implants of this type are performed with the aid of computer support, based on radiographic images of the patient made with the use of computed tomography [7]. Additionally, individual prostheses of the temporomandibular joint enable better load transfer, reduction of micromovements, as well as better adjustment and positioning with screws. At the same time, they allow to shorten the time needed to perform the surgery [8, 9].

1.2. Analysis of the existing solutions of the temporomandibular joint implant

Surgical procedures of the temporomandibular joint with the use of joint implants are used only in cases where traditional procedures are insufficient to improve its functioning.

The main advantage of using temporomandibular joint implants is immediate restoration of joint functions without the need for jaw blockage after surgery [10]. Additionally, alloplastic reconstruction makes it possible to start physiotherapy immediately after implantation and shorten the duration of the operation [6].

Temporomandibular joint implants faithfully reflect the structure of the natural joint, they consist of a concave acetabulum attached to the skull and a convex head with a plate attached to the mandibular branch. Among the implants of the temporomandibular joint, one can distinguish two, three and four-piece constructions. Most ompanies offer implants in several sizes with the possibility of individual adjustment.

An example of a two-piece implant is the ZIMMER BIOMET implant (Fig. 1.1). It is made of a cobalt chrome plate with a head and a polyethylene acetabular cup. This type of solution allows to reduce production costs and limit the introduction of metal elements into the human body [11].

An example of a three-piece design is the OrthoTiN implant (Fig. 1.2). In this case, the cup is made of two elements: a plate made of titanium alloy and a polyethylene insert, while the head implant is a monolith [12].

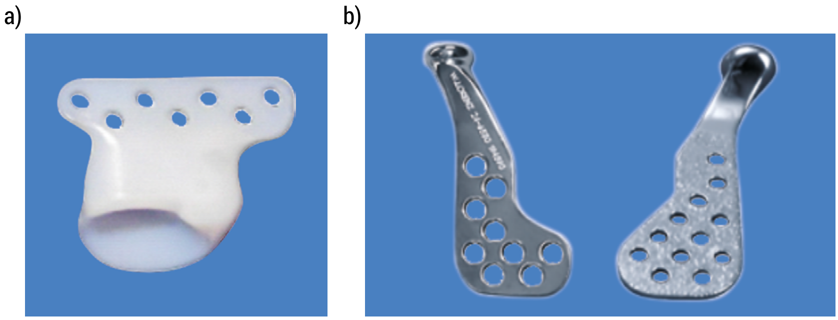


FIGURE 1.1. The implant cup (a) and plate (b) by ZIMMER BIOMET [11]

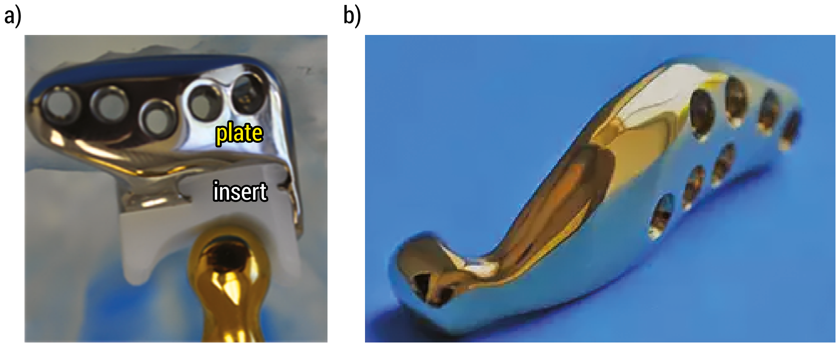


FIGURE 1.2. OrthoTiN implant cup (a) and plate (b)[12]

An example of a four-piece implant is the design of the PRÓTESE DE ATM CUSTOMIZADA company (Fig. 1.3). It is a type of blocked implant of the temporomandibular joint intended for total joint arthroplasty. The ATM implant plate is made of Ti6Al4V titanium alloy, while the plate head is made of CoCrMo alloy. The implant cup is, as before, two-part and consists of a plate made of Ti6Al4V alloy and a polyethylene insert [13].

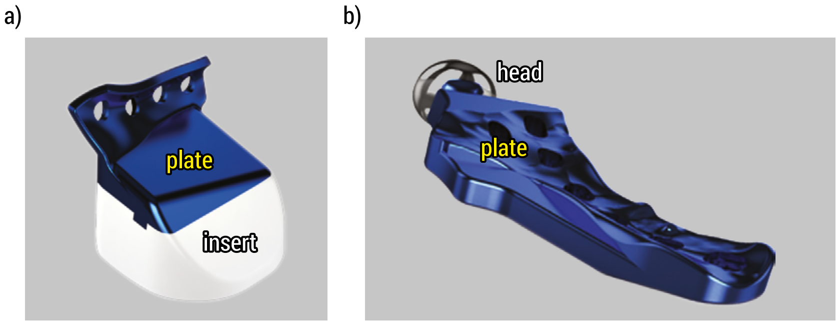


FIGURE 1.3. The implant cup (a) and plate (b) by Próteses Customizadas De ATM [13]

1.3. Designing an individual implant of the temporomandibular joint

Designing patient-specific implants is particularly applicable to irregularly shaped bones (skull, pelvis).

When designing an individual implant of the temporomandibular joint, the following design assumptions were adopted:

- individual adjustment of the temporomandibular joint implant,
- restoration of the basic functions of the temporomandibular joint
- the use of fixing screws intended for the cortical bone,
- the use of biocompatible materials used to manufacture implants for the temporomandibular joint, while maintaining the best possible structural strength parameters.

Designing an implant specially adapted to a given patient begins with radiological examinations using computed tomography (CT). It is thanks to this radiological examination that at a later stage it is possible to create a three-dimensional model of the jaw and skull (Fig. 1.4). Currently, the methodology of human tissue segmentation based on computed tomography is widely known, and more and more software can be found on the market. [14,15]. The presented models were made with the use of free and open source 3D Slicer image computing platform.

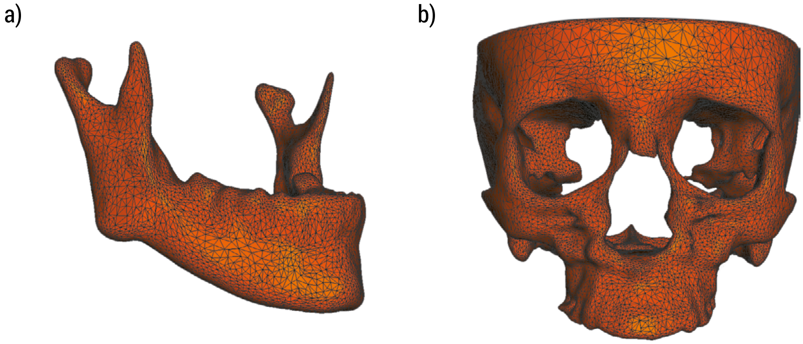


FIGURE 1.4. **Model of (a) the jaw and (b) the skull**

In the next step, based on the obtained geometry, an implant model is created that reflects the individual characteristics of the patient. The models were made in the CAD3D SolidWorks engineering software. The anatomical, individually fitted contact surface of the implant with the bone was obtained using Boolean operations. The received implant model reflects the individual characteristics of the patient's surface bone. Fig. 1.5a shows the process of creating the shape of the mandibular branch implant, Fig. 1.5b – the fitting of the implant to the bone and, eventually, Fig. 1.5c – the resulting implants with individually reconstructed surfaces.

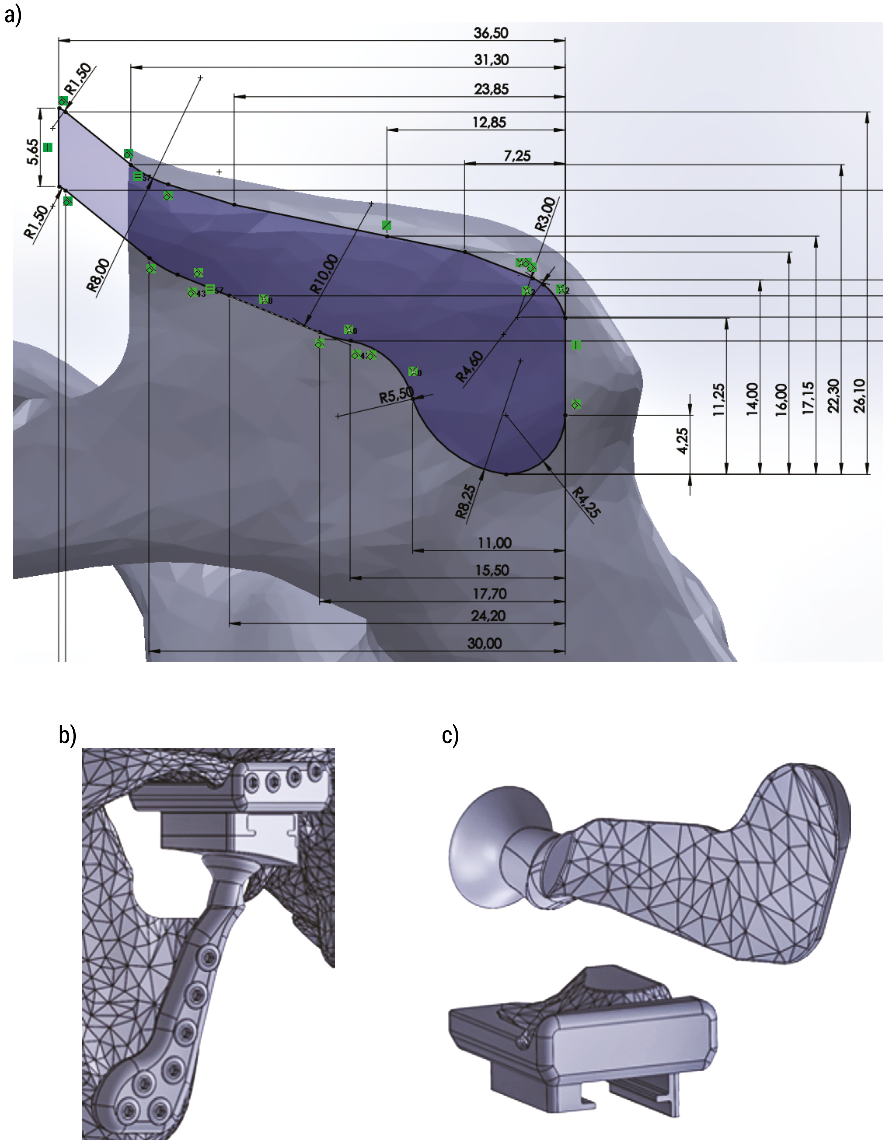


FIGURE 1.5. **Designing an individual implant of the temporomandibular joint: (a) determining the boundary of the implant, (b) fitting the implant, (c) implant with the mapped geometry**

After designing the implant, a very important element of the entire process is its validation, during which the geometry of the implant and the method of its mounting are analyzed. Validation is carried out through strength tests with the use of finite element analysis [16, 17, 18].

The paper presents an example of a strength analysis of a mandibular branch implant. A force loading the plate model of 200 N was assumed, all holes were used to stabilize the implant (Fig. 1.6). The plate material was made of Ti6Al4V alloy. The analysis was carried out for 3 variants of the force, which was set at an angle of 5°, 10° and 15° in order to simulate different angles of the mandible.

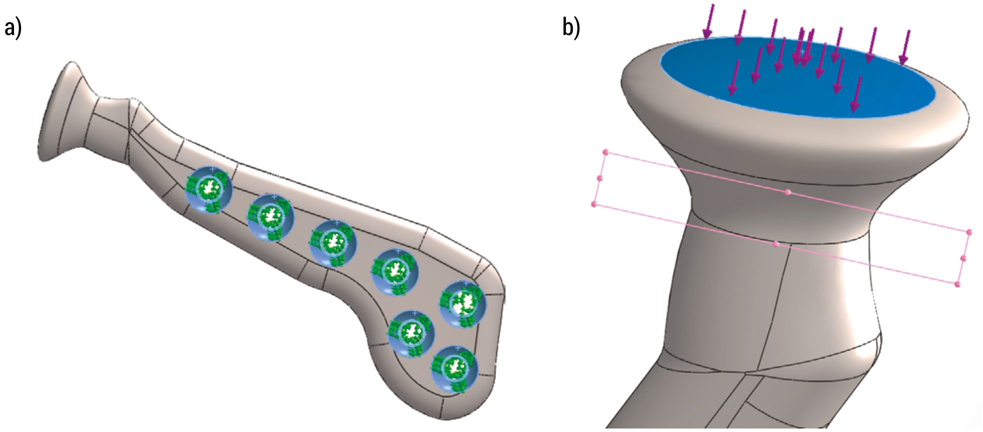


FIGURE 1.6. **Boundary conditions for the implant model of the condylar plate**

The figures below (Fig. 1.7a,b,c) show the result of stress simulation for the implant of the condylar plate of the temporomandibular joint. As a result of the analysis, the readings of the stress distribution according to the Huber Mises hypothesis were obtained. As expected, the concentration of stresses can be observed in the narrowing of the implant and in the area of the first fastening hole. The maximum value of stresses increases with the angle of the force (mandible opening) and is for the angle of 5° – 62.652 MPa, 10° – 89.425 MPa and 136.276 MPa for 15°. It corresponds to the increase in the value of the moment resulting from the change of the direction of the applied force. The yield point of 830 MPa for the material used was not exceeded in any of the cases.

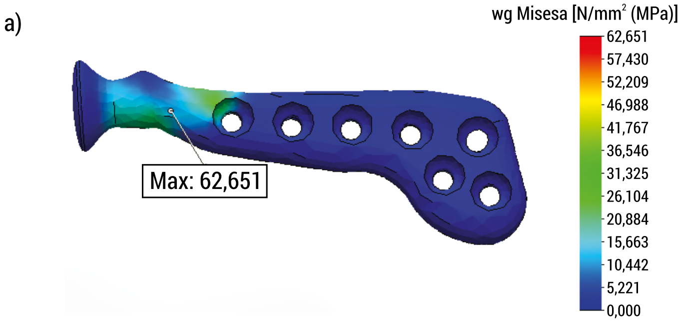


FIGURE 1.7. **Readings of the resulting stresses for the condylar plate in the test with the force loading the implant at an angle of (a) 5°**

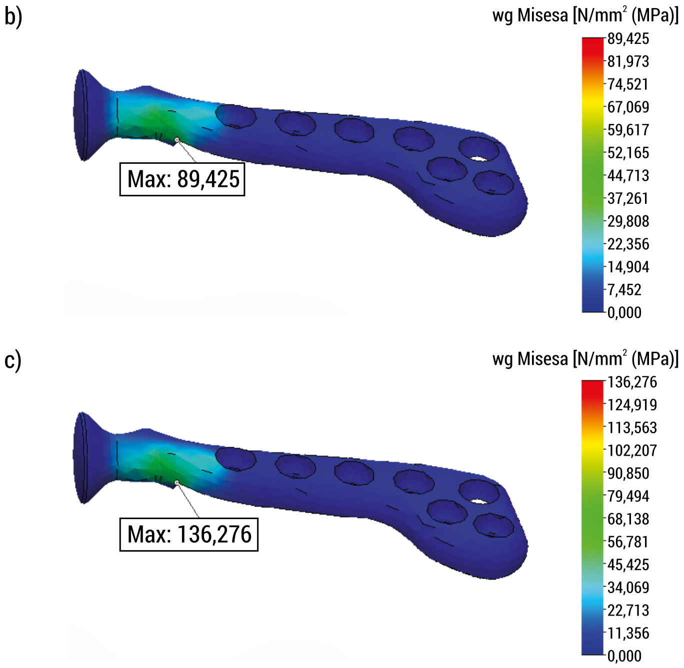


FIGURE 1.7. Readings of the resulting stresses for the condylar plate in the test with the force loading the implant at an angle of (b) 10°, (c) 15°

1.4. Conclusions

The temporomandibular joint plays an important role in the functioning of not only the masticatory system but the entire body. It accounts for the chewing and breathing processes. Therefore, any disturbances within it have serious consequences. In reconstructive surgery, implants are increasingly used to reconstruct the diseased joint. The development of computer technology supports the planning of operations and manufacturing technology, which makes it possible to manufacture "custom-made" implants precisely matched to the irregular geometry of the skull. In the case of such constructions, it is important to perform verification calculations using the finite element method.

References

1. Bartkowski SB. et al.: Chirurgia szczękowo-twarzowa, Podręcznik dla studentów i lekarzy, Wydanie III poprawione i rozszerzone, Collegium Medicum UJ, Kraków 1996.
2. Balcerzak J. et al.: Chirurgia szczękowo-twarzowa, Wydanie V zmienione i uzupełnione, Wydawnictwo Lekarskie PZWL, 2015.
3. Aleksandrowicz R., Ciszek B., Krasucki K.: Anatomia Człowieka, Repetytorium, Na podstawie anatomii człowieka, A. Bochenka, M. Reichera, Wydawnictwo PZWL, Warszawa, 2014.
4. Szpringer-Nodzak M., Wochna-Sobiańska M.: Stomatologia wieku rozwojowego, Wydawnictwo PZWL, Warszawa, 2003.
5. Mercuri LG.: "The role of custom-made prosthesis for temporomandibular joint replacement", 35, 1, 1–10, Revista Española de Cirugía Oral y Maxilofacial, 2013. doi:10.1016/j.maxilo.2012.02.003.
6. Ferreira FM. et al.: "Total temporomandibular joint alloplastic reconstruction", Revista Dor São Paulo, 15(3), 211–214, 2014.
7. Alkindi M. et al.: "In vitro biomechanical simulation testing of custom fabricated temporomandibular joint parts made of electron beam melted titanium, zirconia, and poly-methyl methacrylate", Applied Sciences, 9 (24), 5455, 2014. doi:10.3390/app9245455.
8. Moreira CVA. et al.: "Total bilateral TMJ reconstruction for pain and dysfunction: case report", International Journal of Surgery Case Reports, 42, 138–144, 2018. doi: 10.1016/j.ijscr.2017.11.063.
9. Kim SSY., Miles Somers M.: "Total temporomandibular joint replacement without a submandibular incision", Journal of Oral Medicine and Oral Surgery, 26, 4, 2020. doi:10.1051/mbcb/2020036.
10. Junior JTC. et al.: "Reconstruction of TMJ with prosthesis joint", A Textbook of Advanced Oral and Maxillofacial Surgery, Volume 3, 2016. doi:10.5772/63804.
11. www.zimmerbiomet.com/content/dam/zimmer-biomet/medical-professionals/cmf-thoracic/total-joint-replacement-system/total-joint-replacement-system-brochure.pdf [access: 25.05.2022].
12. www.orthotin.com/tmj-implants [access: 25.05.2022].
13. www.proteseatm.com.br/en/ [access: 25.05.2022].
14. Dąbrowska-Tkaczyk A. et al.: "Virtual and material models of human thoracic-lumbar spine with compressive fracture based on patients' CT data and the rapid prototyping technique", Archive of Mechanical Engineering 58.4 (2011): 425–440.
15. Borkowski, P., Skalski K.: "Expandable endoprosthesis for growing patients – Reliability and research." Biocybernetics and Biomedical Engineering 34.4 (2014): 199–205.
16. Mrozek P., Borkowski P.: "Application of electronic speckle pattern interferometry for evaluation of the static stiffness of a complex orthopedic implant", Advances in Science and Technology. Research Journal 15.2 (2021).
17. Mazzocco AG. et al.: "Custom-made temporomandibular joint prosthesis: computer aided modeling and finite elements analysis", Chemical Engineering Transactions, 74, 1489–1494, 2019.
18. Mesnard M. et al.: "Biomechanical analysis comparing natural and alloplastic temporomandibular joint replacement using a finite element model", Journal of Oral and Maxillofacial Surgery, 69, 4, 1008–1017, 2011.

DOI: 10.24427/978-83-67185-49-3_2

Chapter 2
Numerical and experimental analysis of orthopedic brace for treatment of idiopathic scoliosis

Sławomir Grycuk, Piotr Mrozek

Bialystok University of Technology, Institute of Biomedical Engineering,
e-mail: s.grycuk@doktoranci.pb.edu.pl, p.mrozek@pb.edu.pl

Abstract: The main objective of the work presented in this chapter is to identify the working scheme of the brace's structure, subsequently enabling identification of the possibilities of potential brace design optimization. This chapter presents the results of numerical simulations concerning a Boston-type orthopedic brace with the use of the finite element method (FEM) in the Ansys Workbench environment. Reverse engineering methodology was employed to develop the geometric model, based on the results of digitization performed by an optical 3D scanner. To experimentally verify the FEM model, a test stand for measurement of the brace's field of displacements, equipped with a laser electronic speckle pattern interferometer (ESPI), was used. Special attention was paid to representing the loads and boundary conditions of numerical simulations in experimental tests. Ultimately, the relative difference between numerical and experimental results in displacements in the central part of the brace's front wall did not exceed 0.9%. The force flow trajectories characteristic of the brace's shell were determined using the experimentally confirmed FEM model, revealing the overall way of its operation. The main areas of the brace carrying loads correcting the spine and sites performing little effort, from the perspective of their participation in the brace's fundamental therapeutic application, were identified. The conducted analysis is universal in nature and can be adapted to other types of orthopedic braces. The results obtained allow for the proposal of methods for mechanical optimization of the brace's design.

Keywords: brace, FEM, ESPI, out of plane displacement

2.1. Introduction

Scoliosis is a severe, three-dimensional deformity of the spine [1]. Non-surgical treatment of scoliosis is based on bracing [2]. A brace is a tailor-made thoracic orthosis. It is typically relatively heavy and stiff. For treatment to yield results, the brace must be worn by the patient for over 23h per day, practically from diagnosis until the end of puberty [3]. The corrective action of a brace that is easiest to implement involves applying forces at the apex of the spinal curve and at two support sites on the opposite side. This is referred to as the three-point pressure system [4]. Designs that are more advanced, the Cheneau brace, for example, involve applying a system of multiple three-point corrective systems onto a 3D space [5].

Currently, computerized tools are used more and more frequently to design braces. Here, one can distinguish two main approaches. The first approach, as applied in the design of CtrlBrace (CAD/CAM), for example, is based on measuring the patient's torso without posture correction and then modeling the geometry of the corrective brace in a CAD/CAM operation; this process is performed by a professional orthotics designer [2, 6–8]. The second approach involves using a finite-element torso model (FEM) for the purposes of designing the brace's geometry [9–12], as in the case of NewBrace. For this purpose, computed tomography (CT) or magnetic resonance (MR) medical imaging results are used, perhaps including a 3D scan of the torso, and they are then used to develop a model of the body that represents the real-life musculoskeletal structure. After correction by the computer-modeled 3D field of forces, the geometric image of the torso's FEM model is used as the model for brace shell geometry. On account of the difficult problem of contact and complex anatomical structures in the human torso, it is impossible to represent FEM-modeled fields of corrective forces on the torso faithfully in practice. In the works of [13, 14], the authors compare the efficacy of braces designed according to the above-described approaches. They stress that every individual approach has its upsides and downsides, and integrating them may result in enhanced bracing efficacy. The designed braces' efficacy of action can be verified in practice, e.g. by X-ray or CT imaging of the brace while it is worn by the patient.

It can be observed that present advancements in the field of rigid braces pertain to a better grasp of their corrective functions [4], which are described by, among others, the requisite three-dimensional field of forces applied to the torso and, to a lesser degree, by the mechanical properties of braces indispensable for the performance of corrective functions. The literature provides examples of FEM applications for describing individual braces as options on top of FEM modeling of the torso [15], but these brace models are simplified and do not sufficiently represent the actual brace's geometry. Reference [16] presents a more accurate model, but here FEM analysis results only concern the narrow problem of topological design optimization, lacking an in-depth analysis of the stress and strain distribution in the brace's structure.

Computer modeling of the orthosis's mechanical structure appears to be difficult owing to the character of this construction. This type of orthosis is a thin-walled shell with an open cross-section, under the influence of a 3D distribution of loads. The literature dedicates attention to the relationship between the brace material's mechanical properties and the loads exerted by the orthosis [1]. Stiffer braces exert corrective forces more effectively, whereas the more flexible ones are more comfortable in everyday use. The literature concerning determination of corrective force distributions is extensive [3, 6, 7], but it provides little information about the brace structure's mechanical properties which are responsible for these forces [17, 18]. Literature analysis reveals that critical factors in the efficacy of the brace's action are proper values and directions of the corrective forces it exerts. Thus, it would appear that the conditions that the mechanical construction of a brace must fulfill for the purposes of implementing the required field of forces are equally important.

It could be hypothesized that the appropriate approach to analyzing mechanical features of current brace designs is to formulate a credible numerical model of a real-life orthosis, whose efficacy has already been confirmed in orthopedic practice. In-depth modeling results of such structures are absent in the literature. The numerical simulations and experimental tests presented in this work are aimed towards developing a Boston brace FEM model that would be reliable and experimentally verified. In-depth analysis of the model will serve to identify the working scheme of the brace's structure, subsequently enabling identification of the possibilities of potential brace design optimization. To clarify the methodology of the work a flow chart of the conducted scientific research is presented in Fig. 2.1.

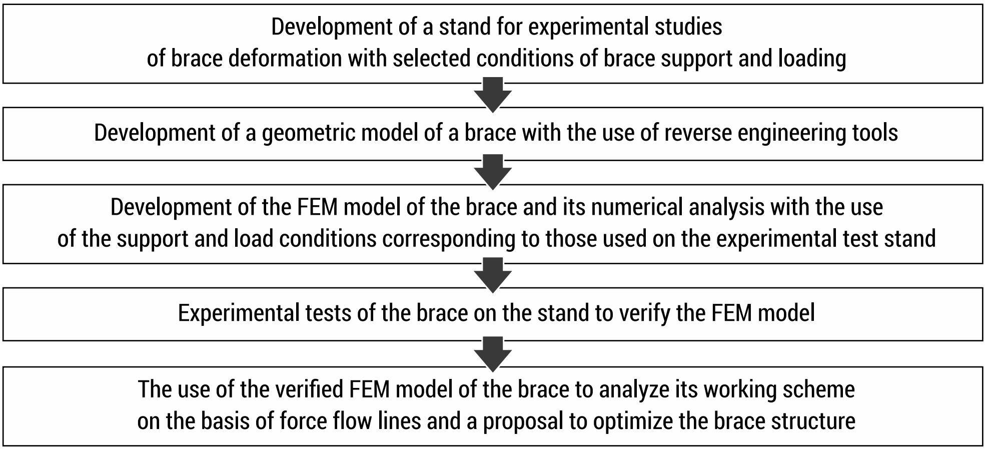


FIGURE 2.1. Flow chart of the conducted scientific research

2.2. Materials and methods

2.2.1. Research object – Boston brace

The Boston brace system has the purpose of enabling non-surgical scoliosis treatment through prevention of scoliosis progression in juvenile and adolescent patients [19]. Treatment with a brace is initiated in cases where scoliosis progression is highly probable. In adolescent idiopathic scoliosis, the patient starts wearing a brace when the curvature of the spine is over 20° according to the Cobb scale. Curves with an apex in the T-6 to L-3 segment are typically treated effectively using Boston braces. Such an orthosis cannot usually treat curves with apexes outside these limits effectively [20]. Fig. 2.2 shows an example of a Boston brace and a diagram of the above-described scoliosis.

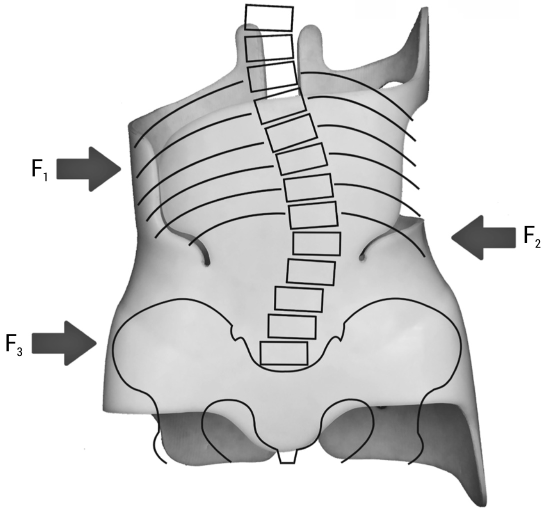


FIGURE 2.2. **Boston brace – illustration of left thoracic-lumbar arch with simplified correction scheme (front view). Forces exerted by the brace: F_1 and F_2 – forces acting on the thorax, F_3 – force applied to the pelvis**

Boston braces are used for lumbar and sacrolumbar scoliosis as well as for thoracic-lumbar scoliosis. Because of this, the upper part of a Boston brace cannot be rigid and may only be supported on the patient's armpit in order to prevent the orthosis from shifting. Braces of this type are produced by tailoring ready-made plastic shapes appropriately to the patient's personal dimensions. After that, more precise adjustments are made to the patients requirements using pads fastened on the interior of the orthosis.

2.2.2. Stand for experimental verification of FEM results

The stand presented in Fig. 2.3 was created for experimental testing of orthopedic braces. The stand consists of: optical bench 3, brace fastening device 7 and pressure sensor 8, system for applying additional load 9 in the form of a tension member drawn over a rotary pulley with a 10 g weight, DPSS laser 1 emitting a beam with 50 mW power and wavelength λ=532 nm, optical set 2, digital camera 4, beam splitter 5 and computer with software 10. Altogether, components: 1, 2, 4 and 5 make up the ESPI interferometer (described more precisely in paper [21]). The brace fastening instrument 7 makes it possible to apply preliminary load, stabilizing the brace, from the brace's interior, and this load is measured by the pressure sensor 8, whose mandrel is at point *A* (Fig. 2.5b) on the interior of the orthosis. Preliminary force *A*, with a value of 35.85 N (Fig. 2.5b) is applied by a support positioned inside the brace, by tensioning of the brace's straps. The value of force *A* falls within the range of real-life forces occurring in braces [22]. Brace displacements induced by the initial force exceed the interferometer's measuring range and are not measured. Next, an additional force *E* of 0.1 N (Fig. 2.5b), inducing brace displacements within the measuring range of the interferometer, is applied close to where the preliminary force was applied and implemented by means of system 9 comprising a tension member, a 10 g weight and a pulley. Displacements of the brace as a result of loading with the additional force are logged and processed by computer with software 10. The test stand simplifies the number and directions of forces acting on the brace in order to reflect real-life boundary conditions in the numerical model as accurately as possible. The application of the ESPI interferometer makes it possible to determine the values of "out of plane" displacements for every point of the brace's surface in the field of view of the interferometer's camera (formula 2.1):

$$\Delta z = n \cdot \frac{\lambda}{2}, \tag{2.1}$$

where: Δz – displacement value in nm, n – order of a correlation fringe, λ – emitted light's wavelength (532 nm in this example).

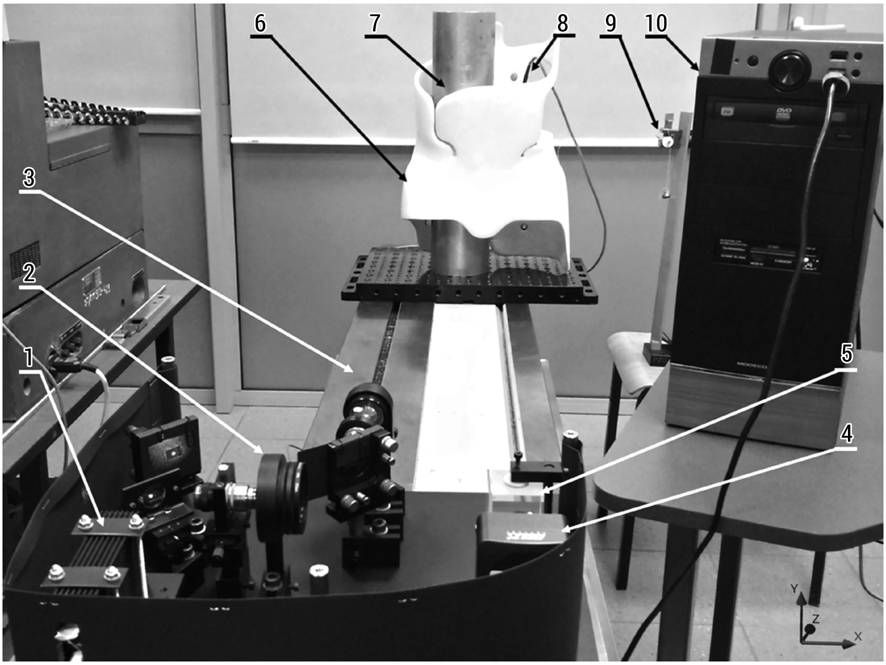


FIGURE 2.3. Brace test stand: 1 – laser, 2 – optical elements/optical set (beam splitter, lenses, polarizers), 3 – bench, 4 – digital camera (CMOS), 5 – beam splitter with matte element, 6 – tested object, 7 – holder for brace fastening and tightening, 8 – pressure sensor, 9 – additional load application system, 10 – computer with software

The test result is the increase in displacements after the application of the additional force. The obtained experimental results from the basis for verification of whether numerical simulations were correct.

The brace was tightened with a force of 35.85 N, which was measured by pressure sensor 8, whose head was located on the inside of the brace at point A (Fig. 2.5b). Gravity also acts on the brace that has been immobilized in this manner. In numerical simulations, an additional force of 0.1 N was applied at the site corresponding to point E after the brace was stabilized (Fig. 2.5b). The results of the tests were determined by the increase in displacements after the additional force was applied. The results of the experimental measurements serve as the foundation for ensuring that the numerical simulations are correct [23].

2.2.3. Brace Geometric Model

Reverse engineering techniques (3D scanning) were used to create a 3D geometric computer model of the brace, based on the existing Boston brace presented in Fig. 2.2. The orthosis underwent preparatory treatments for proper scanning (Fig. 2.4a). For the purposes of 3D measurement, the brace was stiffened using adjustable buckles at the level of tightening straps 1. Matting material 2 was placed on the brace's surface, and reference points 3 enabling a 3D correlation of subsequent 3D scans were glued on. The point cloud was created using an Atos Core 200 optical scanner (GOM GmbH, measurement accuracy of 0.03 mm, positioning repeatability ±0.05 mm, optical system with two CCD 5 Mpx cameras). A mesh of STL triangles defined as a binary representation consisting of 2,753,159 elements was obtained from the point cloud following initial, rough processing. For this purpose, 35 exposures were performed with 28 markers. Geomagic Wrap (3D Systems) point cloud processing computer software was used to remove fastening buckles from the model, repair artifacts, and fill in losses. The mesh, modified in this fashion, was used to create a representation of the brace's exterior surface. The next step involved creating a shell with a thickness of 4.0 mm, based on this surface, corresponding to the brace structure described by NURBS splines. The model representing the actual brace was created through such processing of the point cloud (Fig. 2.4b). An interior pad and a corrective ring made of soft material were additionally modeled at the height of the apex of the spinal curve. The brace computer model was imported for FEM numerical simulation into the ANSYS Workbench environment, in which numerical analyses were then performed.

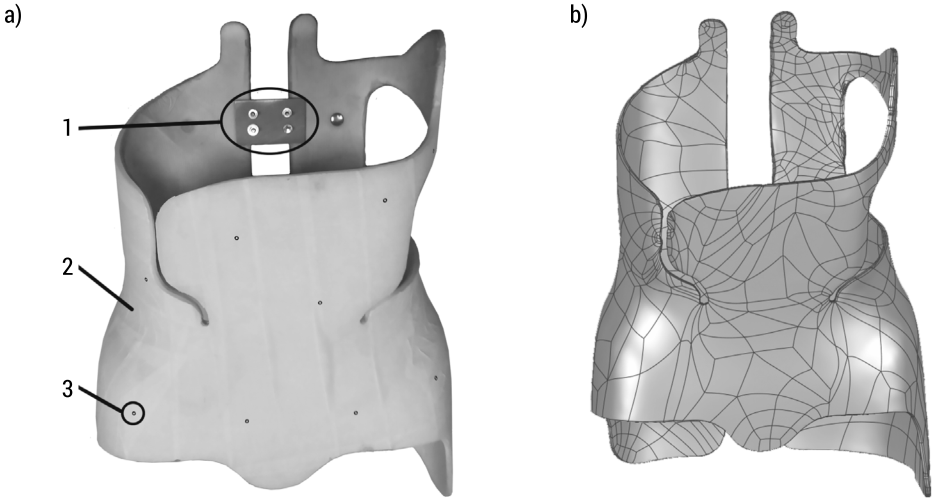


FIGURE 2.4 Reverse engineering – the process of turning a real object into a computer model of an orthopedic brace: a) the brace that is ready to be scanned: 1 – stiffening the brace by fastening, 2 – the orthosis' matted exterior layer, 3 – markers of reference, b) the orthosis as a CAD computer model

2.2.4. FEM Model

The exterior body's material was polypropylene (E = 1000 MPa, v = 0.3) [16, 17], and the middle layer was modeled using soft foam (E = 100 MPa, v = 0.45) [17]. The FEM model was generated using the 10-node tetrahedral element (SOLID187). For detailed determination of principal stress vector value and direction distributions, which were determined separately after the problem was solved for each element individually, a relatively fine mesh of tetrahedral finite elements was generated (Fig. 2.5a). Tetrahedron shapes seem to be rather regular, in spite of the brace's complex geometry. At critical sites where the geometry was more complex, the mesh was made denser, and this was done to improve the accuracy of stress calculation in the case of loading at these sites. Discretization was conducted up to the point where further node mesh concentration did not alter the obtained von Mises stresses by more than 3%. After mesh concentration, there were 156,926 nodes and 91,649 finite elements with a maximum edge length of 7.5 mm (Fig. 2.5a).

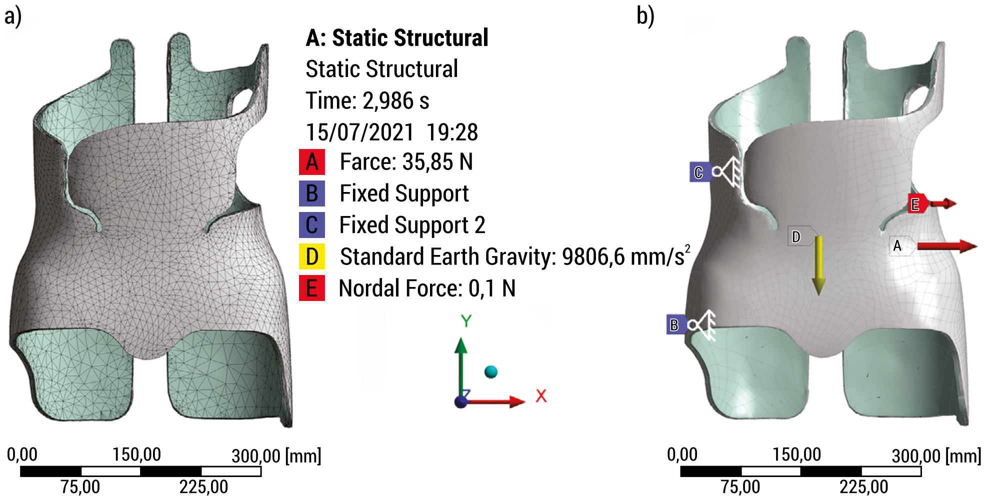


FIGURE 2.5. Numerical brace model: a) tetrahedral finite element mesh generated, b) boundary conditions on the imported geometric model (NURBS): *A* – preliminary force applied to orthosis, *B* and *C* – fixed supports, *E* – additional force applied after the brace has been pre-loaded, *D* – gravitational force

After the finite element mesh was generated and the brace model's parameters were defined, boundary conditions feasible for representation in experimental tests were introduced (Fig. 2.5b). Two immobile supports, defined as very small surfaces *B* and *C*, were immobilized in space. Preliminary loading of the brace was implemented by applying force *A*, with a value of 35.85 N. The value of force *A* falls in the range of real-life forces acting on braces [22]. Preliminary loading was required to stabilize the position of the brace on the experimental test stand, described in section 2.2.2. While the brace is immobilized in this manner, the force of gravity *D* also acts on it, although it does not seem to play a large role. The displacement field was computed for the loads and supports defined in this way. Following the first stage of calculations, as described above, an additional force *E*, with a value of 0.1 N – as per the experiment's conditions – was applied to the brace in the same direction as force *A*. Simulations ultimately yielded a result obtained by computing the difference in displacements of the shell before and after the input of force *E*. Because the interferometer measures displacements on the axis running lengthwise along the camera lens, displacement calculation results with respect to this direction are presented. The distribution of displacements on the Z axis was obtained as the result, and then experimentally verified on the test stand.

2.3. Results

2.3.1. Model verification tests

The brace's displacement distribution, computed with the use of the FEM numerical analysis, is shown in Fig. 2.6a, and the distribution measured with the use of the ESPI is shown in Fig. 2.6b. Displacement values in the experimental tests were determined according to formula (2.1) by means of the correlation fringe distribution with orders of interference from 0 to 6 (Fig. 2.6b). The maximum displacement on the Z axis in the numerical simulations was 1.612 μm, and in experimental measurements 1.596 μm. A detailed comparison of displacements in areas is shown in Fig. 2.6c and 2.6d.

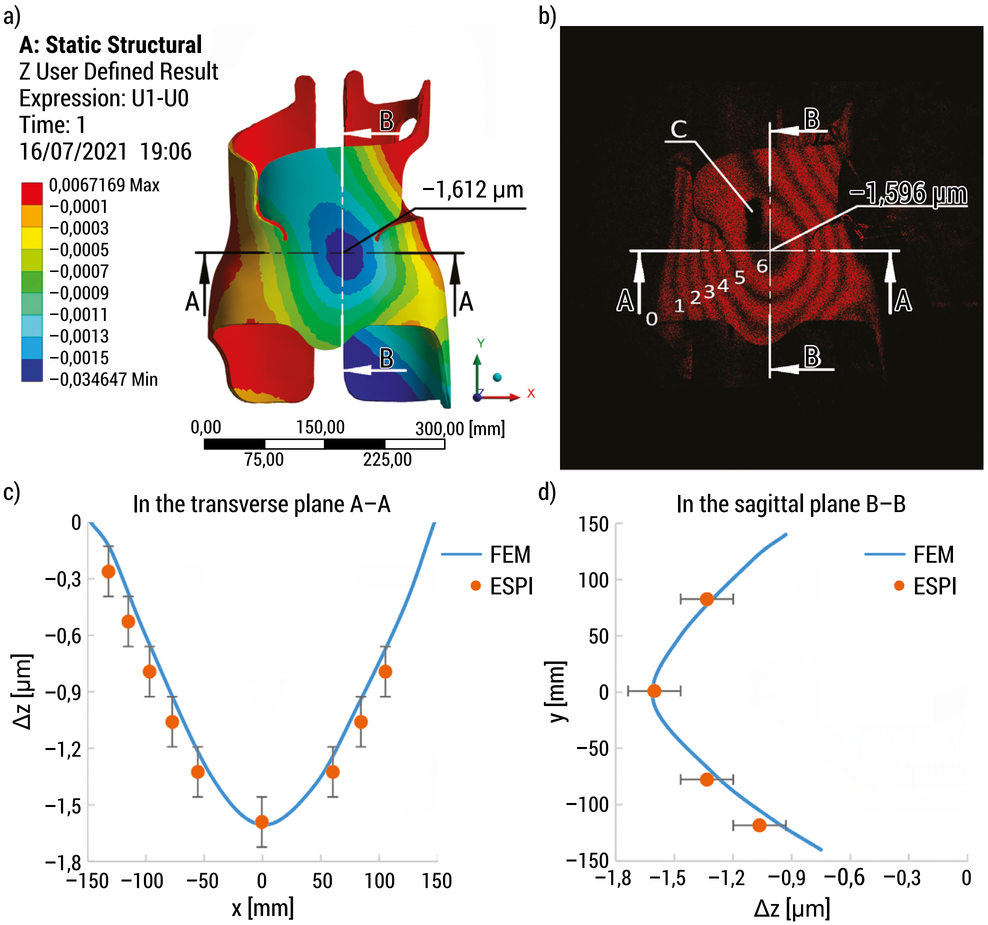


FIGURE 2.6. Displacement distribution results in the Z direction: a) displacement distribution yielded by FEM numerical analysis, b) displacement distribution yielded by ESPI measurements: n = 0, 1,..., 6 are the numbers of successive interference orders (*C* – artifact caused by an overly intense laser beam reflex), c) deformation in the transverse plane A-A, d) deformation in the sagittal plane B-B

2.3.2. Tests using verified numerical model

Tests of the brace were carried out under 28 N load, applied at point *A* (Fig. 2.5b) consistently with typical corrective loading, using the numerical model, verified according to the method described above [22]. The distribution of von Mises stresses shown in Fig. 2.7 was obtained. Besides sites relatively small in area, where the loading force and supports were applied, the value of stresses in the brace was less than 4.4 MPa (Fig. 2.7), falling in the linear range of polypropylene's stress-strain curve [24], which confirms prior assumptions concerning the brace model.

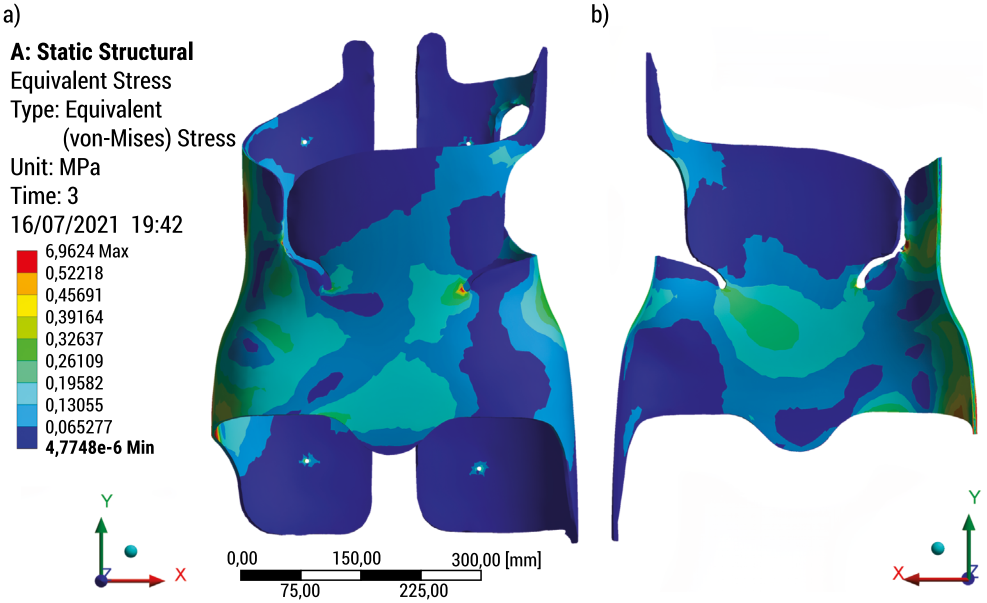


FIGURE 2.7. Brace view with von Mises stress values: a) external view of the front, b) front view from the inside

It is difficult to figure out how the brace works just by based on Fig. 2.7. To present the method of load carrying by the brace's mechanical structure more clearly, Fig. 2.8 shows the front view of the brace, presented in a manner enabling illustration of principal stress vectors on the interior and exterior side of the orthosis's front wall.

Fig. 2.9 shows principal stress vectors only for tensile stresses σ_1. Maximum principal stress σ_1 trajectories are also plotted in this figure. Lines are consistent with the directions of stress vectors, beginning and terminating at loading sites. Stress trajectories had the form of splines with mild changes of direction, tangent to σ_1 vectors.

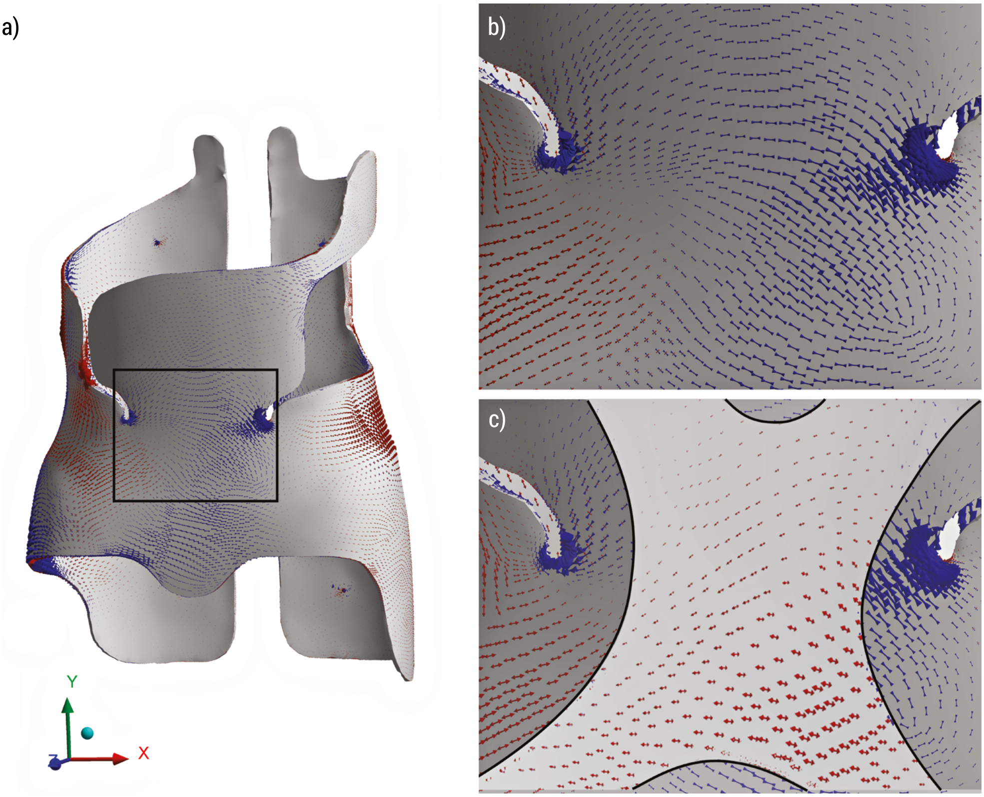


FIGURE 2.8. Principal tensile stresses σ_1 (red) and compressive stresses σ_3 (blue) distribution in the area of front wall of the brace's marked: a) area of the brace's front wall, b) stress distribution on the marked area's exterior surface, c) stress distribution in a cross-section of the front wall – after "removal" of a portion of the front wall's material, the layer closest to the interior surface is visible

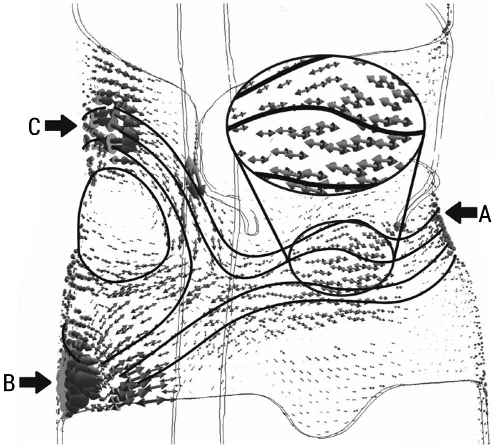


FIGURE 2.9. Principal tensile stress vector distribution with principal stress trajectories

2.4. Discussion

Model verification test results, shown in Fig. 2.6, reveal the distinct similarity of displacement contour line distribution in the FEM image to the ESPI correlation fringe pattern. Comparing the obtained results of the numerical simulations with the experimental results quantitatively, the relative difference of displacements in the brace's central front wall, between both methods, is small – approximately 0.9%. Although only out of plane displacements of the brace were measured, the conclusion that this method of model verification is sufficient was borne out by their good compliance with the results of FEM calculations on the entire front surface, covering almost half of the brace. One may draw conclusions about stress and strain values associated with the displacement field by using linear-elastic relationships. Therefore, the proposed method of describing brace geometry and mechanical parameter values applied to model materials, and above all, the mesh concentration and type of finite elements used for FEM analysis, can be deemed suitable for developing a model that would be highly consistent with the characteristics of the real-life object. Using the FEM model for further tests involving different loading variations and values and distributions similar to those in real life may serve as a foundation for drawing conclusions concerning the work done by a real-life brace.

Applying the so-defined approach, the standard FEM analysis was performed, involving determination of von Mises stress distribution for loads with values on the same order of magnitude as typical corrective forces applied in the elementary three-point pressure system. The results shown in Fig. 2.7 show that the polypropylene shell's thickness provides the orthosis with the requisite strength properties. Von Mises stress values are significantly lower than the tolerable stress values for polypropylene. The brace sites at the greatest risk of damage, in a mechanical regard, and areas performing negligible work, can be identified. Even if similar three-point pressure systems were used, the irregular distribution of stress values in the brace's body makes it difficult to generalize the results obtained across braces with different load values and geometries.

The distributions of principal stresses, as shown in Fig. 2.8 for various parts of the structure, may reveal more about the general nature of the brace's work. The exterior surface of the brace's front wall is compressed, and the interior surface of this wall is under tension, as can be seen. Absolute stress values in the exterior layer are about half of the stress values in the interior layer. By comparing the signs and values of acting stresses, one may conclude that the brace shell is subject to a complex state of bending and tension (Fig. 2.10). Other parts of the brace's shell work similarly. Stresses are distributed throughout most of the brace's shell in such a way that tensile stresses have a significantly greater value on the interior side than on the exterior, and this is due to the superposition of stresses stemming from two individually acting types of load.

In areas where these stresses have relatively high values, distinct regularities can be seen influencing the shape of force flow lines based on determined vector fields of principal stresses and plotted trajectories of maximum principal stresses σ_1 (Fig. 2.9). When summed up, the principal stress trajectories are arranged in a characteristic way, connecting corrective loading sites along paths of relatively short length. This illustration is consistent with the force flow lines concept, which is currently under development in optimal design theory [25]. This concept is based on the observation that, when considering the given geometric boundaries of the object, the “flow” of forces through a mechanical structure between loading sites is most intense near topographic lines linking these points in space. Principal stress trajectories are one way of putting the concept of force flow lines into practice; however, there are other, competing ideas in the literature [26]. A brace can be described as a three-dimensional mechanical system that primarily carries loads over relatively short trajectories between corrective loading sites, using this concept as a foundation. The area of the brace responsible for implementing the brace’s corrective function can be identified based on the 3D distribution of the trajectories depicted in Fig. 2.9. Other areas serve a supporting role, such as stabilizing the torso brace. In Fig. 2.11a, the area performing the corrective function is highlighted in dark color, while less-active areas of the brace’s structure and those performing a supporting function are highlighted in light color. The correction area’s position is comparable to that determined in the publication [16]. In comparison to the topological optimization approach based on elastic energy minimization, it appears that the method of analyzing the structure’s work presented above can be understood more easily in the context of stress distribution analysis.

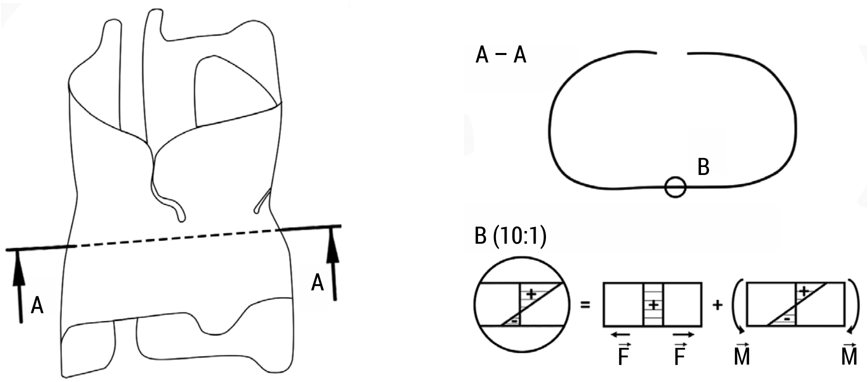


FIGURE 2.10. Stress distribution in the front wall of the brace in cross-section A-A (the interior side of the brace wall corresponds to the top side of the diagram): superposition of stresses caused by tension under force F and bending under moment M

Based on the above-mentioned results of the numerical model that has been experimentally confirmed, one can identify methods of their practical application. Based on the test results, one may conclude that reducing the brace’s weight is possible without changing its mechanical parameters significantly, particularly its stiffness at corrective loading sites. Manufacturing the brace using a thinner material, with additional stiffening in the dark-colored area shown in Fig. 2.11a, responsible for carrying corrective loads, may be the simplest solution. In this location, the brace can be strengthened by integrating composite overlays using a polymer matrix reinforced with carbon or glass fibers, directed along the primary stress trajectories indicated in Fig. 2.11b. When comparing figures 2.8, 2.9, and 2.10, it is evident that overlays should be placed to the inside side of the brace's front component, as the interior side is subjected to substantially more load than the exterior side in this location. Near corrective loading points, the converse is true: the outside side carries significantly more loads than the interior side, hence overlays should be put to the exterior as shown in Fig. 2.11c. In a similar way, overlay application locations in additional areas holding remedial loads can be found.

The discussion above shows that one can predict the trajectory of force flow lines, for the purpose of identifying areas for reinforcement, in the brace rather easily, in the case of a three-point system of forces. The Boston brace uses a corrective method based on the three-point pressure principle, involving fixation above, below, and on the apex of the curve [27]. The experimental test stand provides conditions close to actual, physiological loading. The experimental stand and numerical calculations should account for more complicated physiological loading systems in more complex circumstances; however, it appears that the presented technique may also be effective for optimizing brace design in these more complex cases. Another technique to put the results of the study into practice is to eliminate material, such as by drilling holes in sections of the brace that are not important for its correcting characteristics (light color area in Fig. 2.11a).

The current trend in orthopedic brace research is to find a material that can replace polypropylene and be 3D printed using FDM (Fused Deposition Modeling) [28]. The findings are in line with this trend, as they could serve as the foundation for a method of fabricating a 3D printed structure with openwork geometry that corresponds to the distribution of primary stress paths while also meeting the structural stiffness requirements [16].

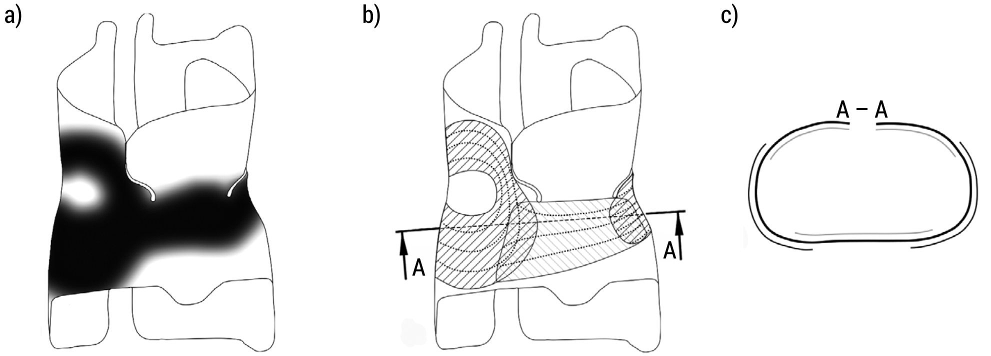


FIGURE 2.11. Boston brace: a) the area of the brace's structure that is placing the biggest effort and performing a corrective function (dark color), and the remaining area, playing a supporting role (light color), b) stiffening by combining thin overlays of composite with a polymer matrix reinforced by glass or carbon fibers that are oriented along the principal stress trajectories, c) the positions of the brace overlays are shown in a cross-section; dashed lines indicate the orientation of reinforcement fibers

The findings of this study could act as a springboard for more research on stiffening and lowering the mass of the brace. The insights gained from the modeling, notably the distribution of principal stress directions and values, allow for the recommendation of optimal stiffening component positioning based on the determined principal stress directions. Simultaneously, the mass of the brace can be reduced by removing material from portions of the brace that are not under load.

2.5. Conclusions

The force flow lines illustrated by plotting primary stress trajectories based on stress distributions calculated in the brace's structure proved to be useful to define the brace's working scheme. In this approach, locations of the brace that were particularly important in terms of corrective treatment were identified. Real-life measurement results confirmed the numerical results. It can be concluded that the force flow concept based on well-developed computer model calculations may serve as the foundation for refining the brace's design and its optimization.

References

1. Chan WY., et al.: "Mechanical and clinical evaluation of a shape memory alloy and conventional struts in a flexible scoliotic brace", Annals of Biomedical Engineering, 46 (8), 1194–1205, 2018.
2. Raux S., et al.: "Tridimensional trunk surface acquisition for brace manufacturing in idiopathic scoliosis", European Spine Journal, 23 (4), 419–423, 2014.
3. Chan A., et al.: "Design and validation of transducers to measure interface force distribution in a spinal orthosis", Medical Engineering & Physics, 34 (9), 1310–1316, 2012.
4. Rigo MD., Villagrasa, M., Gallo, D.: "A specific scoliosis classification correlating with brace treatment: description and reliability", Scoliosis, 12 (1), 1–11, 2010.
5. Rigo M., et al.: "TLSO biomechanics of correction (investigating the rationale for force vector selection)", Scoliosis, 5 (1), 1–8, 2006.
6. Clin J., et al.: "A biomechanical study of the Charleston brace for the treatment of scoliosis", Spine, 35 (19), E940–E947, 2010.
7. Clin J., et al.: "Comparison of the biomechanical 3D efficiency of different brace designs for the treatment of scoliosis using a finite element model", European Spine Journal, 19 (7), 1169–1178, 2010.
8. Cobetto N., et al.: "3D correction of AIS in braces designed using CAD/CAM and FEM: a randomized controlled trial", Scoliosis, 12 (1), 1–8, 2017.
9. Clin J., et al.: "Biomechanical modeling of brace treatment of scoliosis: effects of gravitational loads", Medical & Biological Engineering & Computing, 49 (7), 743–753, 2011.
10. Cobetto N., et al.: "Braces optimized with computer-assisted design and simulations are lighter, more comfortable, and more efficient than plaster-cast braces for the treatment of adolescent idiopathic scoliosis", Spine Deformity, 2 (4), 276–284, 2014.

11. Courvoisier A., et al.: “Prediction of brace effect in scoliotic patients: blinded evaluation of a novel brace simulator – an observational cross-sectional study”, European Spine Journal, 28 (6), 1277–1285, 2019.
12. Nijssen JPA., et al.: “Design and analysis of a shell mechanism based two-fold force controlled scoliosis brace”, Mechanisms and Robotics, 5 (41), 1–12, 2017.
13. Cobetto N., et al.: “Effectiveness of braces designed using computer-aided design and manufacturing (CAD/CAM) and finite element simulation compared to CAD/CAM only for the conservative treatment of adolescent idiopathic scoliosis: a prospective randomized controlled trial”, European Spine Journal, 25 (10), 3056–3064, 2016.
14. Weiss HR., Kleban, A.: “Development of CAD/CAM based brace models for the treatment of patients with scoliosis-classification based approach versus Finite Element Modelling”, Asian Spine Journal, 9 (5), 661–668, 2015.
15. Guan T., Zhang Y.: “Determination of three-dimensional corrective force in adolescent idiopathic scoliosis and biomechanical finite element analysis”, Frontiers in Bioengineering and Biotechnology, 8, 1–13, 2019.
16. Liao YC., et al.: “Shape modification of the Boston brace using a finite-element method with topology optimization:, Spine, 32 (26), 3014–3019, 2007.
17. Périé D., et al.: “Biomechanical modelling of orthotic treatment of the scoliotic spine including a detailed representation of the brace-torso interface”, Medical & Biological Engineering & Computing., 42 (3), 339–344, 2004.
18. Périé D., et al.: “Personalized biomechanical simulations of orthotic treatment in idiopathic scoliosis”, Clinical Biomechanics, 19 (2), 190–195, 2004.
19. Fortin D., et al.: “A 3D visualization tool for the design and customization of spinal braces”, Computerized Medical Imaging and Graphics, 31 (8), 614–624, 2007.
20. Chung CL., et al.: “A mechanical analog thoracolumbar spine model for the evaluation of scoliosis bracing technology”, Journal of Rehabilitation and Assistive Technologies Engineering, 5, 1–9, 2018.
21. Mrozek P.: “The use of electronic speckle pattern interferometry for evaluation of machine tool static stiffness”, Lasers in Engineering, 43 (1), 81–99, 2019.
22. Gignac D., et al.: “Optimization method for 3D bracing correction of scoliosis using a finite element model”, European Spine Journal, 9 (3), 185–190, 2000.
23. Sanz-Pena I. et al.: “Characterising the mould rectification process for designing scoliosis braces: Towards automated digital design of 3D-printed braces”, Applied Sciences, 11 (10), 1–13, 2021.
24. Khlif M., Masmoudi N., Bradai C.: “Polypropylene tensile test under dynamic loading”, Journal of KONES, 21 (1), 132–138, 2014.
25. Wawruch P., Czarnecki S.: “Construction of stress trajectories in optimal, non-homogeneous elastic bodies”, Advances in Mechanics: Theoretical, Computational and Interdisciplinary Issues, 137–140, 2016.
26. Kelly DW., Tosh MW.: “Interpreting load paths and stress trajectories in elasticity”, Engineering Computations, 17 (2), 117–135, 2000.
27. Boston Brace International, Inc: Reference manual for the Boston scoliosis brace, Boston, 2003.
28. Ng KJ., Duke K., Lou E.: “Investigation of future 3D printed brace design parameters: Evaluation of mechanical properties and prototype outcomes”, Journal of 3D Printing in Medicine, 3 (4), 171–184, 2019.

DOI: 10.24427/978-83-67185-49-3_3

Chapter 3
A two-dimensional finite element analysis of the influence of fracture gap and angle on bone healing process

Piotr Prochor, Patrycja Taraszewicz

Bialystok University of Technology, Institute of Biomedical Engineering,
e-mail: p.prochor@pb.edu.pl, patrycja.taraszewicz@gmail.com

Abstract: Over the years, many mechanoregulation concepts of fracture healing have been developed. They all assume that higher values of the stimulus in the fracture gap cause the formation of a fibrous tissue and only lower values can lead to the formation of bone tissue. A literature review of the numerical evaluation of bone fracture healing has revealed some shortcomings in the existing studies. Scientists often focus on only one of the three phases of fracture healing. In addition, studies are often carried out with the use of only transverse fractures. The aim of the study is to develop an algorithm for the numerical evaluation of bone fracture healing, which can be used for both transverse as well as oblique fractures. To conduct the research, the authors proposed an algorithm that considers all phases occurring in secondary healing: 1 – hematoma formation, 2 – tissue differentiation, 3 – bone tissue remodeling. Appropriate CAD models of simplified fracture were designed that additionally considered two variables: fracture angle (from 0° to 45°) and fracture gap (from 2.5 mm to 10.0 mm). Axial loading was used in order to generate mechanobiological stimulus. The obtained results presented a significant influence of the analyzed variables on the studied process. This proved that the designed algorithm can be used to speed up the treatment process and enable designing a more effective rehabilitation procedure for an individual patient.

Keywords: numerical analysis, bone, fracture, secondary healing, evaluation algorithm

3.1. Introduction

Long bone fractures are a common problem in healthcare, which causes at least a temporary decrease in the quality of human life and in critical cases may even lead to a permanent disability resulting from an abnormally healed bone [1]. This phenomenon creates economic problems and a burden for the society, resulting from its universality [1–3].

Fracture types are classified according to several criteria, such as soft tissue penetration [4] or morphology and topography [5,6]. In the first case, the classification includes closed fractures, in which the continuity of the skin is not broken, and open fractures, in which at least one of the bone fragments penetrates soft tissues [4]. Closed fractures are much more common, while open fractures only occur in about 11.5 per 100,000 people every year [4]. It is estimated that fractures of the long bones of the lower limbs are much more dangerous than those of the upper limbs, which may be caused by significantly greater loads transmitted by the lower limbs [4].

According to the next criterion, that is morphology and topography of the fracture, there can be distinguished transverse, linear, oblique, spiral, greenstick, comminuted fractures [5,6]. Transverse and oblique fractures are the most common types, and they can additionally occur in nondisplaced and displaced forms, with the second one significantly hindering the healing process due to the increased fracture gap [7]. Moreover, oblique fractures can occur at different angles [8].

Healing of a fractured bone can occur through direct (primary) healing or indirect (secondary) healing [9]. Primary healing rarely occurs in the natural process of fracture healing, as it requires elimination of the fracture gap with internal or external stabilization [10]. It occurs when the gap is not greater than 0.01 mm and the deformations in this area are less than 2%. This enables the formation of cutting cones that form cavities through the fracture site with the participation of osteoblasts, which are then supplemented with Haversian canals by osteoblasts [9]. This process does not produce callus. Depending on the type of fracture, primary healing may take from a few months to several years [9].

Secondary healing includes both intramembranous and endochondral ossification [11]. This type of bone healing does not require the reduction of the fracture gap or the rigid stabilization of the bone fragments. On the contrary, secondary healing is stimulated by micromovement and load transfer [12]. However, too large bone fragments displacements or excessive load may delay or prevent bone healing [13]. Secondary healing is often observed in inoperable fracture treatment, although there are cases where it occurs after clinical procedures with insufficient fracture stabilization [14].

So far, a number of mechanobiological concepts describing the process of secondary fracture healing have been developed, such as concepts of Carter [15], Claes and Heigele [16] or Huiskes [17]. These models are often complemented by other concepts, including the concept of diffusion of mesenchymal cells into the fracture gap [18]. These concepts are commonly modified and used in numerical analyses evaluating the influence of various factors on the bone healing process [19–24].

A literature review has shown that scientists often limit numerical studies to single stages of bone fracture healing. The first stage, i.e. hematoma formation, is the least frequently considered. At this stage, biological and biochemical factors, rather than mechanical ones, have a much greater impact on improving the process. Moreover, the material properties of a hematoma and a granulation tissue have not yet been clearly defined. Most often, the second phase of bone union, i.e. the differentiation of tissues in the fracture area, is a subject of analyses.

On this basis, it was concluded that analyses of secondary healing of non-transverse fractures are not usually performed. Therefore, it is necessary to supplement the currently known data with analyses of all stages of bone fracture healing, taking into account the morphology and topography of the fracture. For this reason, the aim of this study was to develop an algorithm that takes into account all stages of secondary healing for the most common types of fractures, i.e. transverse and oblique. Such an algorithm would allow to estimate time and mechanobiological conditions needed to individually design rehabilitation procedure for every patient, which is crucial for his appropriate recovery.

3.2. Materials and methods

The evaluation of the studied phenomenon was carried out with the use of an algorithm that takes into account all phases of secondary healing of a bone fracture: callus formation, tissue differentiation and bone remodeling. The algorithm was conducted with the use of the finite element method and Ansys Workbench 2022 software. A simplified diagram of the proposed algorithm is shown in Fig. 3.1.

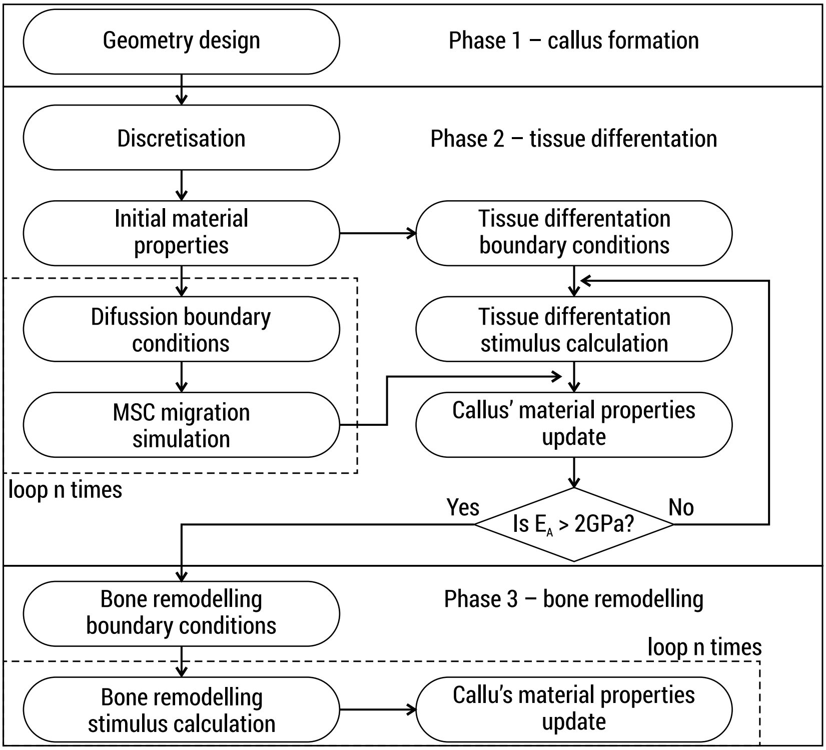


FIGURE 3.1. Proposed algorithm to evaluate the secondary healing of a bone fracture

3.2.1. Phase 1 simulation – callus formation

The first part of the research involved the simulation of the hematoma, which was obtained by designing appropriate geometries in the SolidWorks 2022 software. The models took into account two variables: the width of the fracture gap and the fracture angle. A total of 16 models were analyzed and the included values of the studied variables, as well as the method of marking individual models are presented in Tab. 3.1.

TABLE 3.1. Designations of the analyzed models

		Fracture gap (d)			
		2.5 mm	**5.0 mm**	**7.5 mm**	**10.0 mm**
Fracture angle (α)	0°	α0d2.5	α0d5.0	α0d7.5	α0d10.0
	15°	α15d2.5	α15d5.0	α15d7.5	α15d10.0
	30°	α30d2.5	α30d5.0	α30d7.5	α30d10.0
	45°	α45d2.5	α45d5.0	α45d7.5	α45d10.0

The dimensions and components of the analyzed geometries, which are presented in Fig. 3.2, were developed on the basis of literature reports [25]. Appropriate adjustment of the callus width to the type of fracture was also taken into account [20].

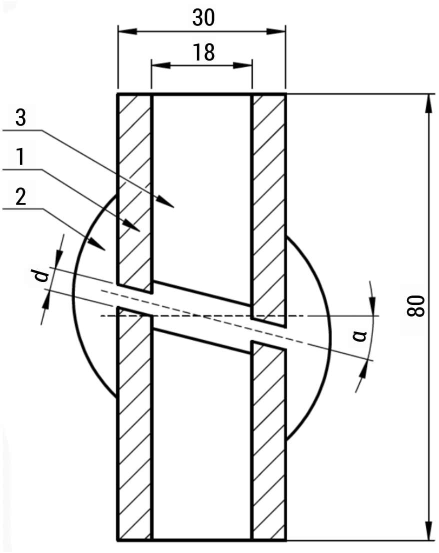


FIGURE 3.2. Scheme of the 2D model used for the analyses: 1 – bone shaft; 2 – callus; 3 – bone marrow

The generated model components were combined in order to consider node-sharing between them in the subsequent discretization process. This was to simulate the biological integrity between anatomical structures during the fracture healing process.

3.2.2. Phase 2 simulation – tissue differentiation

The previously prepared models were subjected to the discretization process, taking into account the 5% mesh refinement test for the analyzed values. Plane223 elements were used for discretization, which has the possibility to unlock diffusive degrees of freedom. The obtained models, an example of which is shown in Fig. 3.3, after optimization had 2,000 ± 1,000 finite elements.

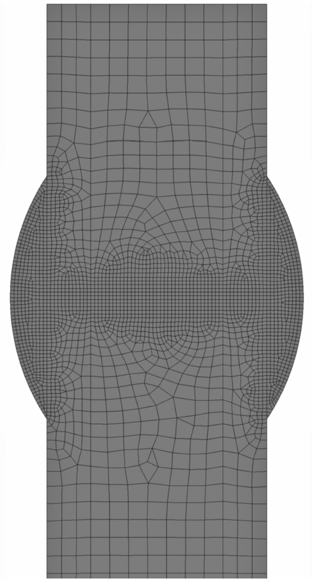


FIGURE 3.3. **Finite element mesh on the example of the α0d2.5 model**

The defined, initial material properties are presented in Tab. 3.2.

TABLE 3.2. **Initial material properties in phase II [26]**

		Material property		
		Density [g/cm^3]	**Young's Modulus [MPa]**	**Poisson's ratio**
Part	Bone	1.741	20000	0.30
	Bone marrow	1.100	2	0.17
	Callus	1.000	2	0.17

Phase II was divided into two processes: migration of mesenchymal cells to the fracture gap and differentiation of MSC according to the mechanoregulation theory.

For the first process, appropriate boundary conditions were defined in which the sources of mesenchymal cells were indicated, i.e. places where their concentration was 100%. In all of analyses, it was the periosteum and bone marrow (Fig. 3.4a).

Subsequently, using Fick's law (equation 3.1), the degree of callus filling over time with mesenchymal cells was determined [21]:

$$J = -D\, \partial c / \partial x, \tag{3.1}$$

where: J – diffusion flux [%/step], D – diffusion coefficient = 2.5 [cm^2/step], c – concentration of the component [%/cm^3], x – change in position [cm].

For the second process, again using the initial material properties and the generated mesh, different boundary conditions were defined, in which a force of 100 N was applied to the upper edge of the bone, while all degrees of freedom were blocked for the lower edge (Fig. 3.4b). A relatively low value of force was used because at the initial stage of fracture treatment, the bone is not able to transfer heavy loads without adversely affecting the bone healing process [21]. Then, the values of hydrostatic stress (σ) and strain (ε) were generated in each finite element of callus.

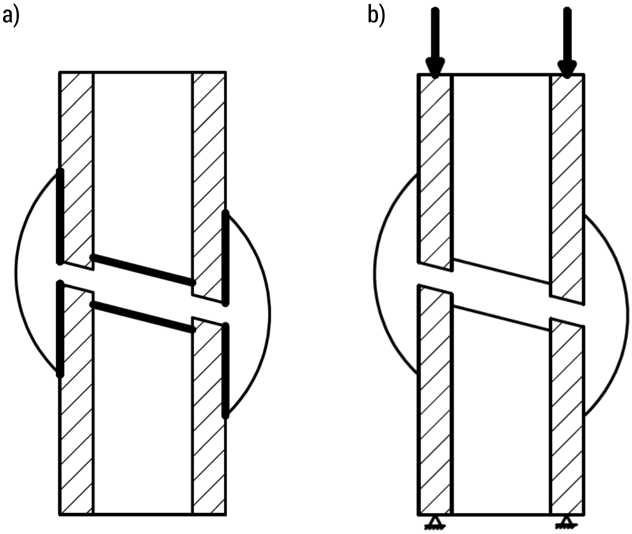


FIGURE 3.4. Boundary conditions for phase II: a) site of mesenchymal cell concentration; b) support and axial force

The concentration of MSC in individual areas of the fracture gap and the values of σ and ε allowed to change the Young's modulus of a specific finite element of callus, taking into account the theory of mechanoregulation fracture treatment according to Claes and Heigele [26]. In order to exclude sudden changes in the value of the Young's modulus, for each finite element the mean value from the previous 10 steps was assigned (equation 3.2):

$$E = MSC \cdot \frac{1}{10} \sum\nolimits_{i=n-9}^{n} E_i, \tag{3.2}$$

where: E – Young's modulus of a specific finite element [MPa], MSC – the percentage of mesenchymal cells [%], i – calculation step.

The values of the Young's modulus of fibrous, cartilage and primary bone tissue were 2 MPa, 10 MPa and 6000 MPa respectively [26]. After changing the properties of the whole callus, its mean Young's modulus (E_A) was determined. When it was greater than 2 GPa, the fracture was considered to be pre-stiffened and phase II was completed [26]. If the E_A was less than 2 GPa, the cycle of updating the callus properties was repeated, taking into account the overtime changes in MSC concentration in the fracture gap. When the fracture area was initially stiffened, the algorithm could go into phase III.

3.2.3. Phase 3 simulation – bone remodeling

In the third phase, the boundary conditions were defined in the same way as shown in Fig. 3.4b. However, the value of the axial force was increased to 600 N. This force was selected to enable the proper process of bone fracture healing [21]. Then, the deformation energy density (U) was determined and considering it as a stimulus stimulating bone tissue remodeling, the properties of callus were updated on the basis of the Huiskes et al. theory. That theory takes into account the so-called dead zone [27], which concept is represented by equation 3.3.

$$\frac{\partial \rho}{\partial t} = B\left(\frac{U}{\rho} - (1+s)K\right) if \frac{U}{\rho} > (1+s)K$$

$$\frac{\partial \rho}{\partial t} = 0 if (1-s)K \le \frac{U}{\rho} \le (1+s)K$$

$$\frac{\partial \rho}{\partial t} = B\left(\frac{U}{\rho} - (1-s)K\right) if \frac{U}{\rho} < (1-s)K, \tag{3.3}$$

where: ρ – bone tissue density [g/cm^3], U – strain energy density [J/cm^3], B – constant regulating the rate of bone remodeling = 1 [(g/cm^3)^2/(MPa/CTU)], s – constant regulating the size of the lazy zone = 0.1, k – constant regulating the threshold values = 0.004 [J/g].

After updating the density values of individual finite elements, the value of the callus Young's modulus was calculated. For this purpose, the relationship between the Young's modulus (E) and bone tissue density (ρ), developed in the studies by Carter and Hayes [28], was used.

3.3. Results

The analyses resulted in a numerical representation of the change in MSC concentration and the value of Young's modulus in the callus.

It has been established that a transverse fracture with a small fracture gap heals in about 12 weeks [29]. Therefore, 84 calculation steps were established, each representing one actual healing day. Having implemented the algorithm, the analysis was performed on the basis of the percentage change of mesenchymal cells in the callus and the change in the Young's modulus of the callus. The percentage change of mesenchymal cells was obtained by averaging all concentration values obtained for elements and dividing the mean by the area of the callus. In the case of Young's modulus of callus, a mean was calculated from all values obtained for each element, which was then divided by an initial value of 2.

3.3.1. Mesenchymal cells migration to fracture gap

From the obtained data, the mean percentage of MSC in the callus for each of the calculation steps was estimated. Fig. 3.5 shows the percentage change in the MSC content in the callus within 7 weeks for transverse and oblique fractures (with a fracture angle of 15°, 30° and 45°).

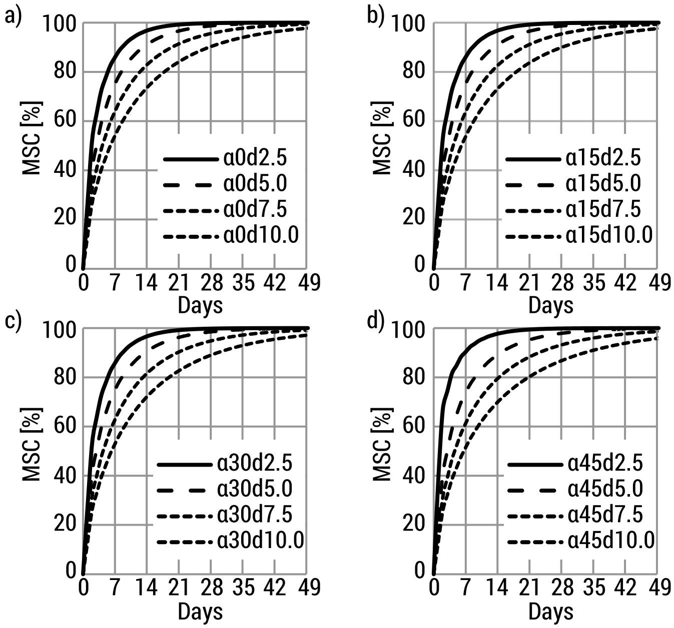


FIGURE 3.5. Change in the percentage of MSC in callus within 7 weeks for a fracture with an angle of: a) 0°; b) 15°; c) 30°; d) 45°

Increasing the width of the fracture gap increased the time for MSC to completely fill the callus. It can be seen that the curves representing the same fracture gap and another fracture angle coincide. This means that the fracture angle does not affect the rate of diffusion of cells to the fracture gap. Within 7 weeks, the complete filling of the callus is observed for the gap width of 2.5 mm, 5.0 mm and 7.5 mm. Fractures with a 10 mm gap are completely filled with MSC over a period of more than 49 days.

The obtained data was then used as input values to determine the change in Young's modulus of the callus during the tissue differentiation and bone remodeling phases.

3.3.2. Changes in Young's modulus of the callus

The results from the second part of the simulations were presented in the form of changes in the mean value of Young's modulus of the callus within 84 days. Fig. 3.6 shows the effect of the fracture gap width on the healing phases in transverse and oblique fractures.

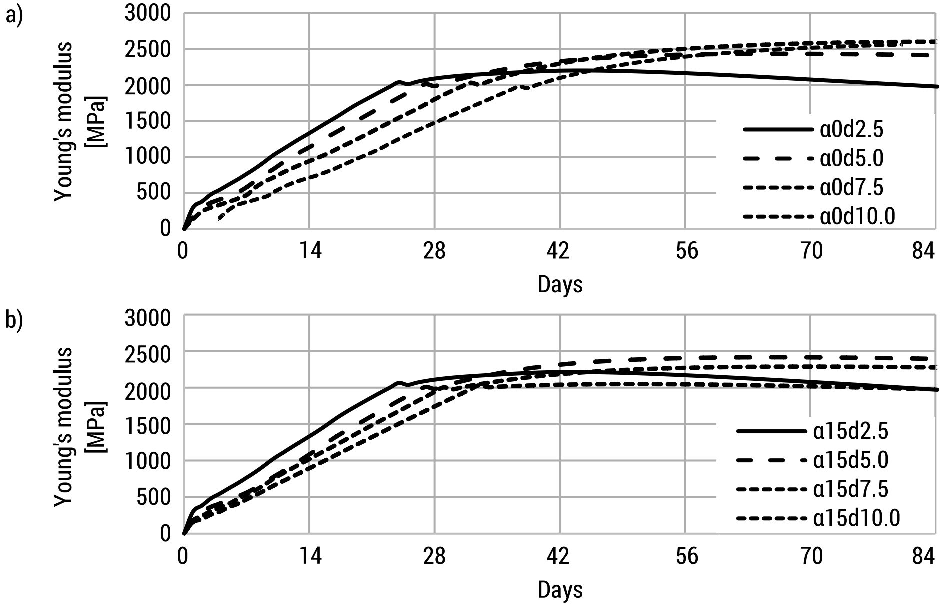


FIGURE 3.6. Change in Young's modulus of the callus during phase II and III of bone fracture healing for a fracture with an angle: a) 0°; b) 15°

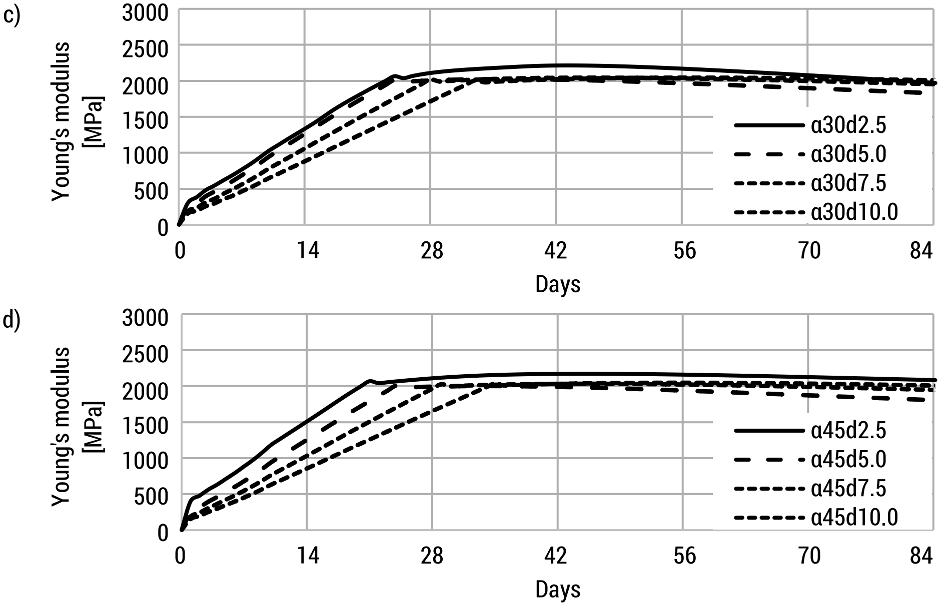


FIGURE 3.6. Change in Young's modulus of the callus during phase II and III of bone fracture healing for a fracture with an angle: c) 30°; d) 45°

The graphs presented in Fig. 3.6 show the process of both phases II and III of the bone healing process, while the transition from one to the other is characterized by a relatively sudden change in the Young's modulus change rate. During phase II, the mean Young's modulus of the callus increases nearly linearly, while in phase III its value changes non-linearly.

The analysis of the results showed that the width of the fracture gap influences the time that the healing process takes. As the distance of bone fragments increases, the time needed to fully heal the injury also increases. Moreover, it was noticed that an increase of a fracture gap also results in delaying a transition from phase II to phase III, which was presented in Fig. 3.7.

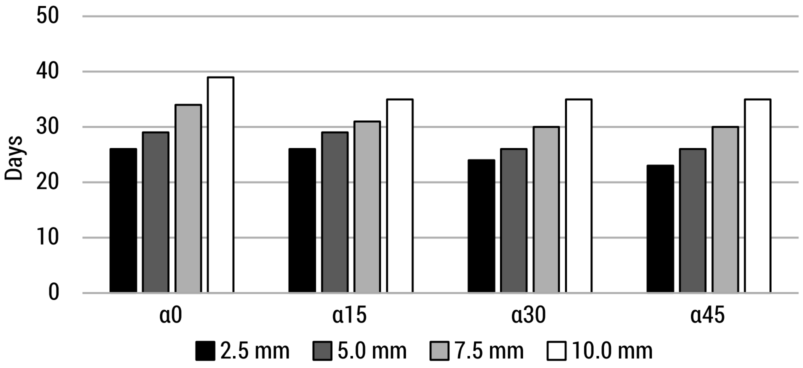


FIGURE 3.7. Days to transition from phase II to phase III for each type of fracture

In addition, irrespective of the fracture type, a larger fracture gap resulted in slower callus resorption. It was noticed on the basis of a more intensive increase in the mean Young's modulus at the beginning of phase III for the gap of 10 mm, compared to the gap of 2.5 mm.

3.4. Discussion

The algorithm used for the analyses was based on three basic assumptions: Fick's first law (diffusion law), the theory of mechanoregulation of fracture treatment according to Claes and Heigele and the theory of bone tissue remodeling by Huiskes et al. Fick's first law was used to simulate the diffusion of MSC into the callus. The mechanoregulation theory of bone healing described how callus evolved under the influence of mechanical stimulus. On the other hand, the theory of Huiskes et al. was applied to simulate bone tissue remodeling. Therefore, all phases of fracture treatment were taken into account in the conducted analyses, while the phase of callus formation was considered with the design of an appropriate geometry. A literature review has shown that scientists often focus only on the influence of mechanical stimuli on one of the described phases.

In the presented study, the operation of the designed algorithm for the assessment of bone fracture healing was verified. The functionality of the algorithm was confirmed by comparing the obtained results with the data available in the literature. The pattern of callus evolution obtained in 2D analyses coincides with the biological sequence of tissue creation. Similarities can also be found in the works of Isaksson and Lacroix [18,30]. In their research, they proved that the change of Young's modulus in the callus is related to the rate of migration of MSC to the fracture area.

Ghiasi in his research also discussed the effect of the distance between bone fragments on the time needed to fully restore the bone to its initial efficiency. According to the results, the increase in the fracture gap delays the healing of the fracture [20, 26]. In the presented research, the influence of the fracture gap on the bone healing process for the transverse and oblique fractures was also verified. According to the obtained data, irrespective of the type of fracture, the time necessary for the full healing of the injury increased with the increase of the fracture gap.

Interestingly, during the preparation of the geometry for analyses, it was also noticed that the width of the callus increases with the increase in the width of the fracture gap. This phenomenon is also observed in clinical and experimental analyses. Larger callus is needed with an enlarged fracture gap to achieve the appropriate healing process [30]. This allows to conclude that the elimination of the fracture fissure will reduce the resulting callus. This also confirms the assumptions necessary for primary healing.

The verification of the designed algorithm proved that the combination of the selected concepts and theories correctly predicts the course of bone fracture healing in numerical simulations.

3.5. Conclusions

The designed algorithm is an innovative tool for evaluating long bone fracture healing processes. It is one of the few that includes the three phases of healing and can be used not only for transverse, but also oblique fractures. The analyses carried out with the use of the above-mentioned algorithm filled a significant gap in the literature. The designed algorithm can be used to speed up the treatment process and enable designing a more effective rehabilitation for an individual patient.

3.6. Limitation of the study

The presented study is characterized by limitations typical of numerical research. Among them, there can be distinguished the use of a simplified bone model as well as neglecting other factors (such as chemical) that have a direct influence on fracture healing. However, the adapted simplifications are commonly used in numerical analyses and are crucial for evaluating the influence of selected variables on the analyzed process. For this reason, they are not negatively affecting the quality of the obtained and presented results.

References

1. Parziale J.R.: "Disability evaluation of extremity fractures", Physical Medicine and Rehabilitation Clinics of North America, 12(3), 647–657, 2001.
2. Harvey N., Dennison E., Cooper C.: "Osteoporosis: impact on health and economics", Nature Reviews Rheumatology, 6 (2), 99–105, 2010.
3. Mock C. et al.: "The global burden of musculoskeletal injuries: challenges and solutions", Clinical Orthopaedics and Related Research, 466 (10), 2306–2316, 2008.
4. Court-Brown CM. et al.: "The epidemiology of open long bone fractures", Injury, 29(7), 529–534, 1998.
5. Ishman S.L., Friedland D.R.: "Temporal bone fractures: traditional classification and clinical relevance", Laryngoscope, 1734–1741, 114 (10), 2004.
6. Garnavos C. et al.: "New classification system for long-bone fractures supplementing the AO/OTA classification", Orthopedics, 35 (5), e709–e719, 2012.
7. Meeson R. et al.: "The influence of gap size on the development of fracture union with a micro external fixator", Journal of the Mechanical Behavior of Biomedical Materials, 99, 161–168, 2019.
8. Cohen H. et al.: "The impact velocity and bone fracture pattern: Forensic perspective", Forensic Science International, 266, 54–62. 2016.
9. Marsell R., Einhorn T.A.: "The biology of fracture healing", Injury, 42 (6), 551–555, 2011.
10. Mirhadi S., Ashwood N., Karagkevrekis B.: "Factors influencing fracture healing", Trauma, 15 (2), 140–155, 2013.
11. Bahney C.S. et al.: "Cellular biology of fracture healing", Journal of Orthopaedic Research, 37 (1), 35–50, 2019.

12. Jagodzinski M., Krettek C.: "Effect of mechanical stability on fracture healing – an update", Injury, 38 (1), S3–S10, 2007.
13. Lienau J. et al.: "Initial vascularization and tissue differentiation are influenced by fixation stability", Journal of Orthopaedic Research, 23 (3), 639–645, 2005.
14. Perren S.M.: "Evolution of the internal fixation of long bone fractures: the scientific basis of biological internal fixation: choosing a new balance between stability and biology", The Journal of Bone and Joint Surgery. British Volume, 84 (8), 1093–1110, 2002.
15. Carter D.R., Blenman P.R., Beaupre G.S.: "Correlations between mechanical stress history and tissue differentiation in initial fracture healing", Journal of Orthopaedic Research, 6 (5), 736–748, 1988.
16. Claes L.E., Heigele C.A.: "Magnitudes of local stress and strain along bony surfaces predict the course and type of fracture healing", Journal of biomechanics, 32 (3), 255–266, 1999.
17. Huiskes R. et al.: "A biomechanical regulatory model for periprosthetic fibrous-tissue differentiation", Journal of Materials Science: Materials in Medicine, 8 (12), 785–788, 1997.
18. Isaksson H. et al.: "A mechano-regulatory bone-healing model incorporating cell-phenotype specific activity", Journal of theoretical biology, 252 (2), 230–246, 2008.
19. Byrne D.P., Lacroix D., Prendergast P.J.: "Simulation of fracture healing in the tibia: mechanoregulation of cell activity using a lattice modeling approach", Journal of orthopaedic research, 29 (10), 1496–1503, 2011.
20. Ghiasi M.S. et al.: "Computational modeling of human bone fracture healing affected by different conditions of initial healing stage", BMC Musculoskeletal Disorders, 20 (1), 1–14, 2019.
21. Lacroix D., Prendergast P.J.: "A mechano-regulation model for tissue differentiation during fracture healing: analysis of gap size and loading", Journal of biomechanics, 35 (9), 1163–1171, 2002.
22. Wehner T. et al.: "Prediction of the time course of callus stiffness as a function of mechanical parameters in experimental rat fracture healing studies-a numerical study", PLoS One, 9 (12), e115695, 2014.
23. Lipphaus A., Witzel U.: "Finite-element syntheses of callus and bone remodeling: Biomechanical study of fracture healing in long bones", The Anatomical Record, 301 (12), 2112–2121, 2018.
24. Mehboob A., Chang S.H.: "Biomechanical simulation of healing process of fractured femoral shaft applied by composite intramedullary nails according to fracture configuration", Composite Structures, 185, 81–93, 2018.
25. Lacroix D. et al.: "Biomechanical model to simulate tissue differentiation and bone regeneration: application to fracture healing", Medical and Biological Engineering and Computing, 40 (1), 14–21, 2002.
26. Ghiasi M.S. et al.: "Computational modeling of human bone fracture healing affected by different conditions of initial healing stage", BMC Musculoskeletal Disorders, 20 (1), 1–14, 2019.
27. Huiskes R. et al.: "Adaptive bone-remodeling theory applied to prosthetic-design analysis", Journal of Biomechanics, 20 (11–12), 1135–1150, 1987.
28. Carter D.R., Hayes W.C.: "The compressive behavior of bone as a two-phase porous structure", The Journal of Bone and Joint Surgery, 59 (7): 954–962, 1977.
29. Lacroix D., Prendergast P.J.: "A mechano-regulation model for tissue differentiation during fracture healing: analysis of gap size and loading", Journal of Biomechanics, 35 (9), 1163–1171, 2002.
30. Byrne D.P., Lacroix D., Prendergast P.J.: "Simulation of fracture healing in the tibia: mechanoregulation of cell activity using a lattice modeling approach", Journal of Orthopaedic Research, 29 (10), 1496–1503, 2011.

DOI: 10.24427/978-83-67185-49-3_4

Chapter 4
Comparative assessment of the use of the beighton score and the sachse test to diagnose hypermobility in the group of fitness instructors

Ewa Karolina Brzozowska[*], *Eugeniusz Sajewicz*[**]

[*] *Studio Rehabilitacji Funkcjonalnej, e-mail: ewakbrzozowska@gmail.com*
[**] *Bialystok University of Technology, Institute of Biomedical Engineering, e-mail: e.sajewicz@pb.edu.pl*

Abstract: Hypermobility is an increased range of joint mobility caused by the generalized laxity of the ligaments stabilizing the joints and joint capsules. Changes in the proportion of collagen are the basis of joint hypermobility. Hypermobile people suffer from a number of symptoms related to the musculoskeletal system and other systems due to the widespread presence of connective tissue in the body. The aim of the study was to compare the diagnosis of disorders from the hypermobility spectrum using the Beighton and Sachse tests. 60 women were included in the study. They were divided into two groups, similar in terms of age and BMI value. The Sachse test and the Beighton scale were used. The ambiguous, and in some cases contradictory, results obtained in both methods indicate the need for a more detailed analysis of diagnostic methods for disorders from the hypermobility spectrum.

Keywords: joint hypermobility, hypermobility spectrum disorders, Beighton score, Sachse test

4.1. Introduction

Hypermobility is an increased range of joint mobility caused by the generalized laxity of both the ligaments stabilizing the joints and the entire joint capsules. Each joint has a physiological, limited range of motion. In hypermobility, the physiological barrier of the joint is crossed without violating the anatomical barrier, which would result in damage to the joint elements [1]. Immobility may affect one or more joints, may be the result of an injury or the result of appropriately targeted training. It can also be a congenital feature and be associated with a number of symptoms related to the locomotor system and other organs and systems – in this case, one can suspect a set of symptoms referred to as disorders from the hypermobility spectrum (Hypermobility Spectrum Disorders, HSD).

Hypermobility is associated with disturbances in the proportions of type I and III collagen. The correct condition is 18–21%, while in hypermobile people – 28–46%. The effect is the flaccidity of the joint structures made of connective tissue [1, 2, 3].

The main symptom of hypermobility is pain. Its nature and frequency of occurrence vary depending on the case. The occurrence of pain is associated with a reduced tolerance to static loads, e.g. when standing or sitting at a desk for a long time. Most often it is located in the area of knee joints, spine, arms, feet and hips [2, 4]. The occurring hypermobility is often accompanied by posture defects, e.g. flat feet (flat foot and transverse flat foot), hallux valgus, deep thoracic kyphosis and lumbar lordosis or scoliosis [4, 5]. Hypermobile people are more likely to be injured. Repeated injuries lead to frequent pain and loss of function, and may predispose to earlier degenerative joint changes [4].

In addition to the symptoms related to the locomotor system, disorders of the hypermobility spectrum cause a number of other symptoms, such as: increased skin extensibility, stretch marks, large scars, difficult wound healing, instability of the masticatory organ, lowering of the reproductive organs, rectal prolapse, urinary incontinence, hernias, varicose veins, increased fragility of blood vessels, mitral valve prolapse. Symptoms such as depression, anxiety and neurosis are also a particularly important problem – patients are tired of pain, which translates into problems in everyday functioning [3, 4, 6].

Recent years have brought progress in research on hypermobility. Due to the discrepancies in the literature resulting from the use of heterogeneous nomenclature, the following classification of clinical cases has been proposed [4]:

1) people with "asymptomatic" hypermobility, which can be confirmed by a positive result on the Beighton score, but patients do not feel any discomfort from the musculoskeletal system;
2) people with hypermobility who meet the diagnostic criteria of the hypermobile type of Ehlers-Danlos syndrome;
3) people with hypermobility with pain, in whom the diagnosis of hEDS cannot be confirmed, it is proposed to define a disorder from the hypermobility spectrum (Hipermobility Spectrum Disorders, HSD).

Among patients with the hypermobility spectrum disorder (HSD), we can distinguish 4 basic types, described below (tab. 4.1) [4].

According to the available literature, hypermobility concerns 0.6 – 31.5% of the population, depending on the adopted study criteria. It is known to decrease with age and is more common in women (up to 3 to 5 times more often), as well as in Asian and African races [2, 3, 7].

Hypermobility is diagnosed on the basis of information obtained from an interview, physical examination and functional tests: Beighton's score, Sasche's test, Bulben's, Contompasis or the Brighton criteria [1]. The Beighton scale is most often used [8].

Initially developed as a screening test, it has become the primary diagnostic criteria for hypermobility [9]. The tests used are not precise – clinicians notice many discrepancies, which result in difficulties in making a diagnosis. Patients are considered hypochondriacal, are not adequately treated and their functioning worsens.

TABLE 4.1. Classification of cases of hypermobility spectrum disorders

Name	Specification
G-HSD (Generalized Hypermobility Spectrum Disorders)	Generalized joint hypermobility syndrome confirmed by a positive Beighton score. Additionally, there are symptoms (one or more) from the musculoskeletal system.
P-HSD (Peripheral Hypermobility Spectrum Disorders)	Peripheral joint hypermobility, which affects only hand and / or foot joints and one or more musculoskeletal symptoms, is usually negative.
L-HSD (Localized Hypermobility Spectrum Disorders)	Hypermobility limited to only one joint with musculoskeletal symptoms but only in a given location, the Beighton score is negative.
H-HSD (Historical Hypermobility Spectrum Disorders)	History of hypermobility, Beighton score is negative, but there may be musculoskeletal symptoms (one or more).

The aim of the study was to compare the indications of the Beighton scale and the Sachse test, used to diagnose disorders from the hypermobility spectrum. Such a comparison is important insofar as there is no reference in the literature to which the results of the above-mentioned tests could be compared, it may give an indirect assessment of their usefulness in practice.

4.2. Materials and methods

Two groups of people with various physical activity were created for the purpose of the research. The first group consisted of 30 women aged 23–45, working as fitness instructors. The second group, treated as a control group, also consisted of 30 women in the same age group, who were characterized by a lack of systematic physical activity.

The selection of people for the control group was made on the basis of two parameters – age and BMI. The homogeneity of the features of both groups was verified using the Student's t-test for independent samples, checking the normality of the distribution of the analyzed quantities (for all variables, the indications of the Kolmogorov-Smirnov test $p > 0.2$) and homogeneity of variance (Levene's test indications: $p = 0.35$ for age and $p = 0.26$ for BMI). The results of the Student's t-test are shown in Fig. 4.1. On their basis, it can be stated that there are no statistically significant differences in the tested parameters between the groups. Both groups were interviewed with questions about age, weight and height. The research group additionally defined the seniority of the instructor's work, which was on average 4.7 years (from 0.5 to 20 years).

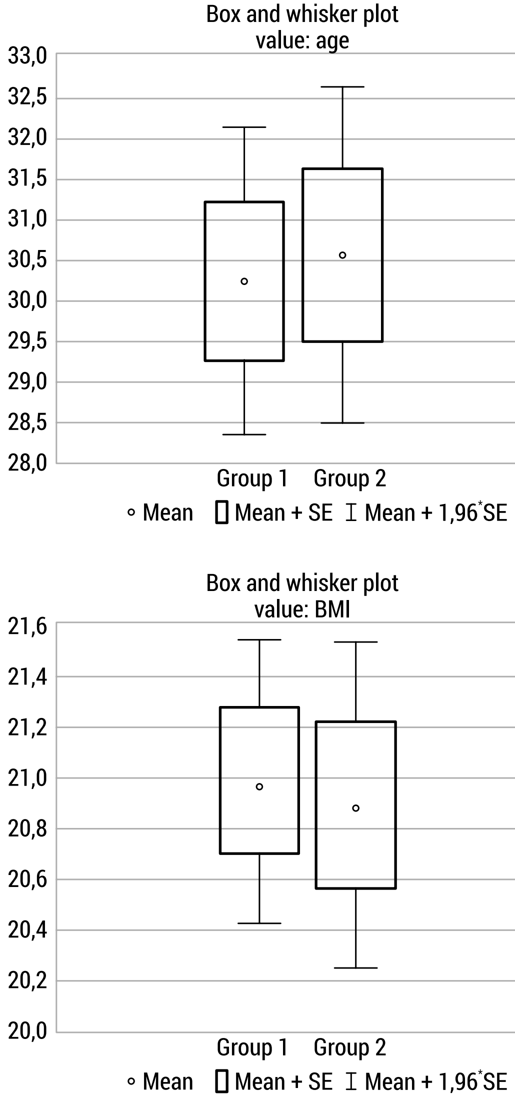


FIGURE 4.1. Box and whisker plot for age values and BMI for groups 1 and 2

The two tests were selected to assess the occurrence of hypermobility among the respondents: the Sachse test and the Beighton score. The Sachse test consisted in assessing the possibility of making the following movements:

1) flexion of the lumbar spine,
2) rotation of the hip joints,

3) lateral flexion of the thoracic-lumbar spine,
4) twist of the thoracic-lumbar spine,
5) opposing joining of hands behind the back,
6) horizontal flexion shoulder joint,
7) dorsiflexion of the fingers in the metacarpophalangeal joints,
8) valgus elbows,
9) head rotation.

1 point was awarded for the correct performance of each test. The result is the number of points scored. Hypermobility is found when obtaining at least 7 points.

The Beighton score assesses the ability to make the following movements:

1) touch thumb to forearm;
2) fifth finger metocarpalphalageal joint extension >90°,
3) touch palms flat to floor with knees straight;
4) elbow extension >10°,
5) knee extension >10°.

1 point was awarded for each movement made as described. All trials, except for the forward bend, were repeated and scored separately for the right and left limbs. Obtaining at least 4 points proved the presence of hypermobility.

Measurements were made using a goniometer in the appropriate positions.

4.3. Results

In the Sachse test, the first group averaged 5.4 points. Hypermobility can be found in 9 people (30%). The full results of all Sachse test trials in the first test group are presented in Table 4.2. In the second group, 3 people (10%) obtained at least 7 points. The average score in this group is 4.13 points. The results of the individual trials are presented in Table 4.3.

TABLE 4.2. **Analysis of individual movements of the Sachse test in group 1**

Movement	Yes (can do)		No (can't do)	
Flexion of the lumbar spine	29	97%	1	3%
Rotation of the hip joints	10	33%	20	67%
Lateral flexion of the thoracic-lumbar spine	27	90%	3	10%
Twist of the thoracic-lumbar spine	16	53%	14	47%
Opposing joining of hands behind the back	29	97%	1	3%
Horizontal flexion shoulder joint	17	57%	13	43%

Movement	Yes (can do)		No (can't do)	
Dorsiflexion of the fingers in the metacarpophalangeal joints, passive movement above 60°	10	33%	20	67%
Valgus elbows	14	47%	16	53%
Head rotation over 90° in each side	10	33%	20	67%

TABLE 4.3. Analysis of individual movements of the Sachse test in group 2

Movement	Yes (can do)		No (can't do)	
Flexion of the lumbar spine	28	93%	2	7%
Rotation of the hip joints	4	13%	26	87%
Lateral flexion of the thoracic-lumbar spine	11	37%	19	63%
Twist of the thoracic-lumbar spine	8	27%	22	73%
Opposing joining of hands behind the back	25	83%	5	17%
Horizontal flexion shoulder joint	8	27%	22	73%
Dorsiflexion of the fingers in the metacarpophalangeal joints, passive movement above 60°	16	53%	14	47%
Valgus elbows	21	70%	7	30%
Head rotation over 90° in each side	4	13%	26	87%

On the Beighton score, group 1 scored 2.47 points on average. The features of hypermobility were demonstrated by 7 people (23.33%). The results of individual trials are presented in Table 4.4.

In group 2, the result confirming hypermobility was obtained by 7 people (23.33%). The average score was 2.7 points. The results of all tests are presented in Table 4.5.

TABLE 4.4. Analysis of individual movements of the Beighton score in group 1

Movement	Yes (can do)		No (can't do)	
Ability to touch thumb to forearm – right hand	6	20%	24	80%
Ability to touch thumb to forearm – left hand	7	23%	23	77%
Fifth finger metocarpalphalageal joint extension >90° – right hand	5	17%	25	83%
Fifth finger metocarpalphalageal joint extension >90° – left hand	5	17%	25	83%
Ability to touch palms flat to floor with knees straight	29	97%	1	3%
Elbow extension >10° – right upper limb	7	23%	23	77%
Elbow extension >10° – left upper limb	7	23%	23	77%
Knee extension >10° – right lower limb	3	10%	27	90%
Knee extension >10° – left lower limb	3	10%	27	90%

TABLE 4.5. **Analysis of the individual movements of the Beighton score in group 2**

Movement	Yes (can do)		No (can't do)	
Ability to touch thumb to forearm – right hand	2	7%	28	93%
Ability to touch thumb to forearm – left hand	4	13%	26	87%
Fifth finger metocarpalphalageal joint extension >90° – right hand	5	17%	25	83%
Fifth finger metocarpalphalageal joint extension >90° – left hand	5	17%	25	83%
Ability to touch palms flat to floor with knees straight	28	93%	2	7%
Elbow extension >10° – right upper limb	15	50%	15	50%
Elbow extension >10° – left upper limb	15	50%	15	50%
Knee extension >10° – right lower limb	3	10%	27	90%
Knee extension >10° – left lower limb	3	10%	27	90%

Comparing the results of the Sachse test and the Beighton score, it was noticed that in the first group only in 3 people the hypermobility was simultaneously confirmed by both tests, in 10 cases the results were inconsistent. In the second group, the positive result of both tests was noted in 2 people, in 6 the result was inconsistent.

4.4. Discussion

The basis for effective treatment of the consequences of hypermobility syndrome is precise and reliable diagnostics of the disease in question. Obtaining contradictory results in the presented studies leads to the need to develop a more detailed analysis of the reliability of the functional tests performed. Theoretically, both methods used should give similar results, but the data obtained during the research differs significantly. The Sachse test assessed the mobility of almost the entire musculoskeletal system, starting with the spine (flexion, lateral flexion, rotations in the thoracic-lumbar and cervical spine), through the mobility of the hip joints (rotations), shoulder joints (flexion, abduction, horizontal flexion, rotations), elbow joints (degree of valgus, hyperextension), to the mobility of small metacarpophalangeal joints. The Beighton scale narrows down to the assessment of spine flexion in the forward tilt movement, mobility of the small joints of the hands (metacarpophalangeal joint of the thumb and V finger) and hyperextension over 10° in the knee and elbow joints. Therefore, the question arises as to which of the applied research methods is more accurate and more precisely measures the immobility of the joints. The fact that the vast majority of respondents (over 90% in both studied groups) are able to bend forward with straightened knee joints, and a large proportion of people who do not regularly train any sport have significant hyperextension in the knee and elbow joints, may explain the achievement of a high result on the Beighton score in group 2, which does not necessarily have to be associated with the hypermobility of the entire musculoskeletal system. Thus, the conclusion is that the diagnosis of disorders from the hypermobility spectrum cannot be based on one functional test assessing the mobility of selected joints.

Authors dealing with the issue of hypermobility concluded that the mere use of functional tests is insufficient [9]. The research trials described above also confirm this statement. It cannot be forgotten that hypermobility is a disease with a wide clinical picture – due to the common presence of collagen of connective tissue, symptoms appear in various, often distant organs. Therefore, the diagnostic scope was extended to include symptoms from various systems and organs, creating the Brighton criteria, which are based on the Beighton score and take into account extra-articular symptoms [2, 10]. The Brighton criteria were divided into major criteria and minor criteria (Tab. 4.6). Hypermobility spectrum disorders are defined when two major criteria, one major and two minor criteria, four minor or two minor criteria are met, and a first degree relative is diagnosed with hypermobility.

TABLE 4.6. Brighton Criteria

MAJOR CRITERIA
1. Beighton score ≥ 4.
2. Arthralgia (joint pain) in ≥ 4 joints for ≥ 3 months.
MINOR CRITERIA
1. Beighton scores 1–3.
2. Arthralgia in 1–3 joints for ≥ three months or back pain, spondylosis (spinal arthritis) or spondylolisthesis (spinal subluxation) for ≥ three months.
3. Dislocating/Subluxation (partial dislocation) > one joint, or the same joint more than once.
4. More than three soft tissue injuries, tenosynovitis (inflammation of sheath around tendons) or bursitis (inflammation of the fluid-filled sac in a joint).
5. Marfanoid Habitus; a characteristic appearance including being tall and slim and having long, thin fingers.
6. Thin, stretchy skin or abnormal scarring (cigarette paper scarring or easily scarring).
7. Droopy eyelids, short-sightedness, double vision.
8. Varicose Veins (particularly at a young age).
9. Rectal or uterine prolapse.

It is also worth paying attention to the research by Bravo and Wollf. They developed a characteristic phenotype that may be useful in the diagnosis of hypermobility [11]. Conducting observations on the population of women diagnosed with disorders from the hypermobility spectrum in Chile, they found that there are common features in the appearance of the patients' faces. They included:

- blue color of the sclera, which is the result of a lack of transparency in the diaphragm;

- changes in ear cartilage: small, round ear lobe adhered to the head, kidney-shaped ears or “dumbo ears”;
- changes in nasal cartilage: lump connected to bone or cartilage, deviations in the nasal septum, asymmetrical nose;
- triangular face, sometimes with a prominent chin.

The authors also included the characteristic features of the phenomenon of “hand holding the head sign” – supporting the head on the palmar side of the hand with maximum flexion in the metacarpophalangeal and carpal-radial joints or on the dorsal side of the hand with hyperextension in the finger joints. According to the authors, taking into account the above-mentioned phenotypic features could significantly simplify the diagnosis of hypermobility.

The diagnosis of hypermobility spectrum disorders in clinical practice is relatively rare. Perhaps it is related to a selective view of the patient’s ailments, e.g. concerning only one painful joint and its vicinity, which limits the scope of the physical examination. In addition, the diagnostic criteria ignore quite important issues, e.g. more frequent occurrence of hypermobility in women – therefore it is worth considering a classification that also takes into account gender [12]. There is also no possibility of a uniform, objective assessment by means of laboratory tests, as no biochemical or genetic markers confirming the diagnosis of hypermobility have been found [11].

Another important issue is that the phenomenon of hypermobility is well described in the literature on rheumatology and orthopedics, but there are no relevant studies addressed to family medicine physicians, to whom patients with symptoms of hypermobility most often report. Therefore, according to British studies, only about 10% of cases are diagnosed by the primary care physician [10].

Disorders from the hypermobility spectrum are a set of symptoms that significantly affect the daily functioning of patients. Due to the presence of chronic pain and a number of comorbid symptoms, better diagnostic tools are needed, taking into account the entire clinical picture of the disease in question [13, 9]. This will improve the comfort of everyday life of people with hypermobility, showing the possibilities and limitations related to disorders from the hypermobility spectrum.

4.5. Conclusion

Based on the results of the research, the authors of this study concluded that both of the presented tests used in the diagnosis of disorders from the hypermobility spectrum do not precisely assess the disease under study. Both the Beighton scale and the Sachse test do not guarantee an unambiguous and reliable method of diagnosis. In addition, efforts were made to prove that only functional tests, which only assess the musculoskeletal system, cannot be used. Hypermobility spectrum disorders affect the whole organism, due to the common presence of collagen of connective tissue, which should be taken into account in both diagnosis and treatment. Only taking into account the various, multi-organ consequences of the disease in question guarantees an accurate and complete diagnosis. Due to the discrepancy in the assessments obtained with the use of both tests, it should be stated that there is an urgent need to develop new or modify the previously used tests.

References

1. Stodolna J., Stodolny J., Marczyński W.: "Hipermobilność konstytucjonalna jako przyczyna zespołów bólowych narządu ruchu", Kwartalnik Ortopedyczny, 2, 100–104, 2006.
2. Kopff B., Raczkowski J.W.: "Zespół hipermobilności stawów – rzadko rozpoznawana patologia w obrębie narządu ruchu", Kwartalnik Ortopedyczny, 2, 80–92, 2011.
3. Zweers M.C., Hakim A.J., Grahame R., Schalkwijk J.: "Joint hypermobility syndrome: the pathophysiologic role of tenascin-X gene defects", Arthritis & Rheumatism, 50(9), 2742–2749, 2004.
4. Castori M., Tinkle B., Levy H., Grahame R., Malfait F., Hakim A.: "A framework for the classification of joint hypermobility and related conditions", American Journal of Medical Genetics, 175(1), 148–157, 2017.
5. Mirska A., Kalinowska A.K., Topór E., Okulczyk K., Kułak W.: "Łagodny zespół hipermobilności stawów", Neurologia dziecięca, 20(41), 135–140, 2011.
6. Engelbert R., Juul-Kristensen B., Pacey V., De Wandele I., Smeenk S., Woinarosky N., Sabo S., Scheper M.C., Russek L., Simmonds J.V.: "The evidence-based related for physical therapy treatment of children, adolescents, and adults diagnosed with joint hypermobility syndrome/hypermobile Ehlers Danlose Syndrome", American Journal of Medical Genetics, 175(1), 158–167, 2017.
7. Hakim A.J., Cherkas L.F., Grahame R., Spector T.D., MacGregor A.J.: "The genetic epidemiology of joint hypermobility", Arthritis & Rheumatism, 50(8), 2640–2644, 2004.
8. Bockhorn L.N., Vera A.M., Dong D., Delgado D.A., Varner K.E., Harris J.D.: "Interrater and Intrarater Reliability of the Beighton Score: A Systematic Review", The Orthopaedic of Sports Medicine, 9(1), 2021.
9. Malek S., Reinhold E.J., Pearce G.S.: "The Beighton Score as a measure of generalised joint hypermobility", Rheumatology International, 2021.
10. Simpson M.R.: "Benign joint hypermobility syndrome: evaluation, diagnosis, and management", The Journal of the American Osteopathic Association, 106(9), 531–536, 2005.
11. Bravo J.F., Wolff C.: "Clinical study of hereditary disorders of connective in a Chilean population: joint hypermobility syndrome and vascular Ehlers-Danlos syndrome", Arthritis & Rheumatism, 54(2), 515–523, 2006.
12. Armstrong R.: "Relative joint contribution to joint hypermobility in rugby players, netballers and dancers: the need for careful consideration of lumbar flexion", International Journal of Sports Physical Therapy, 13(4), 676–686, 2018.
13. Molander P., Novo M., Hallstam A., Lofgren M., Stalnacke B.-M., Gerdle B.: "Ehlers-Danlos Syndrome and Hypermobility Syndrome Compared with Other Common Chronic Pain Diagnoses – A Study from the Swedish Quality Registry for Pain Rehabilitation", Journal of Clinical Medicine, 9(7), 2020.

DOI: 10.24427/978-83-67185-49-3_5

Chapter 5

Kinematic analysis of instantaneous centre of rotation of prosthetic knee mechanisms

Kinga Grodzka[*], *Eugeniusz Sajewicz*[*], *Marcin Dziemianowicz*[**]

[*]*Bialystok University of Technology, Faculty of Mechanical Engineering,*
e-mail: niemyjskakinga@gmail.com,
e.sajewicz@pb.edu.pl
[**]*Design Pro Technology Ltd., e-mail: m.dziemianowicz@designprotechnology.com*

Abstract: The purpose of this work was to determine the influence of kinematic evidence, based on the trajectories of the instantaneous center of rotation (ICR) on the optimization of the functionality of selected prosthetic knee mechanisms. For the research 11 commercially used knees based on the articulated quadrilateral mechanism were used to visualize the operation of mechanisms with two methods: analytical and graphic. The generated paths were compared in terms of position changeability and shape. It was found that depending on the position of the ICR related to the axes of the mechanism, prosthetic knees obtained different level of stability and energy expenditure. The results have also shown that ICR point position and shapes of its trajectories are variable for particular classes of mechanisms. An analysis of the obtained results suggests that manipulation of the position of the instantaneous center of rotation by a suitable design of the linkages length of the mechanism enable to achieve the desired features, corresponding to the real needs of patients.

Keywords: knee prosthesis, instantaneous center of rotation, four-bar linkage

5.1. Introduction

Nowadays, choosing the suitable knee prosthesis for people after amputation is a complex and complicated task, which depends, among others, on the user's activity, required gait stability, as a sense of security and has an individual character. The literature presents that these features have a strong relation with instantaneous centre of rotation (ICR) and its trajectory, which also allows to determine different functional properties of the whole prostheses [1, 17]. There are a number of knee solutions for four-bar mechanisms on the market, but rarely at the construction stage they are subject to in-depth kinematic analysis, including the one based on the description of the trajectory of the instantaneous centre of rotation (centrode).

A kinematic analysis of one of the basic parameters, namely ICR, enables to visualize the operation of mechanisms, so that the gait is more cosmetic and controlling the limb becomes easier and less tiring. The articulated quadrilateral concept is widely used in knee prostheses due to its simplicity and a low price as well as high functionality, and determination of the length of the linkages allows to obtain various types of mechanisms. In the four-bar knee mechanisms an ICR is created similar to the human knee, allows the combination of rotation and translation motion of the knee joint in sagittal plane, so it has been used in the prosthetic knees to imitate the complex function of the knee joint for the creation of orthotic and rehabilitation devices [24]. Well-documented research shows that the achievement of stability is possible through appropriate prosthesis structure [3, 12], containing mechanical braking mechanisms [2], blocking the knee under load or through geometric design [15]. Radcliffe [20, 21, 22] developed a stability diagram, which in combination with the instantaneous centre of rotation trajectory is used to classify mechanisms to adapt to the needs and capabilities of different groups of amputees. The design of mechanism linkages and the position of the prosthetic knee axis (approximate to the anatomical position) affect ICR manipulation, and thus can obtain beneficial features [14, 23, 25].

The instantaneous centre of rotation is the point where, with a small change in the knee bending angle, the femoral part rotates around this point, while at the same time straightening relative to the part of the shin, which appears temporarily stationary. During a natural gait cycle with larger changes in the bend angle, the ICR changes its position, thereby generates a specific trajectory of the moving centrode [13]. The ICR path provides valuable information about the kinematic properties of the knee. This means that a well-designed prosthesis recreates the ICR trajectories of the knee mechanism, and the more the ICR path is similar to a natural knee joint, the better is the biomechanical performance of the prosthesis [3, 16, 24] Therefore, for path-generation purposes, the optimization of knee mechanisms is required.

There are many methods for the generation and optimization of the design of four-bar prosthetic knee mechanisms based on ICR, such as techniques using mathematical algorithms [9, 18], a computer simulation connected with solid modelling [5] and random dimensions generation [25]. According to the abovementioned studies, it is essential to determine appropriate dimensions of the linkages to obtain the trajectory of ICR, which follows the precision points. However, research on trajectory changes and comparative characteristics in relation to different mechanisms available on the market are not widely developed [24]. Catalogues of knee prostheses provide no accurate information that could be used to evaluate the functionality of the mechanism and to fit the proper prosthetic knee to an amputee's needs.

The problem of kinematics of prosthetic knee mechanisms is currently rarely undertaken, and if it is, rather based on classical analyses from the late 20th century, e.g. on the basis of the father of prosthetic biomechanics – Radcliffe. Hence, there exists a modest number of articles on this subject, despite its significant importance, which the authors want to show in this publication. Taking the above aspects into account, the authors – by generating ICR trajectories with analytical and graphic methods – have made a comparison of selected knee mechanisms in terms of their usability tailored to the needs of a particular amputee and the possibility to apply the results to design new mechanisms or to optimize the existing solutions.

5.2. Materials and methods

In the analysis, eleven types of four-bar knee mechanisms were considered, differentiated, among others, in terms of mobility (values of the maximum bending angle), dimensions and construction masses, as well as limits to the user's weight and activity level (Tab. 5.1). The selection of the mechanisms was made on the basis of appropriate catalogue photos (side views of models), then structural diagrams were created to be used for determining basic structural dimensions (length of linkages and their mutual location), necessary for further analysis. An example of one of the analysed models with its created structural model is shown in Fig. 5.1. A detailed description of the given prosthetic knee is available in the 2020 Össur catalogue [19].

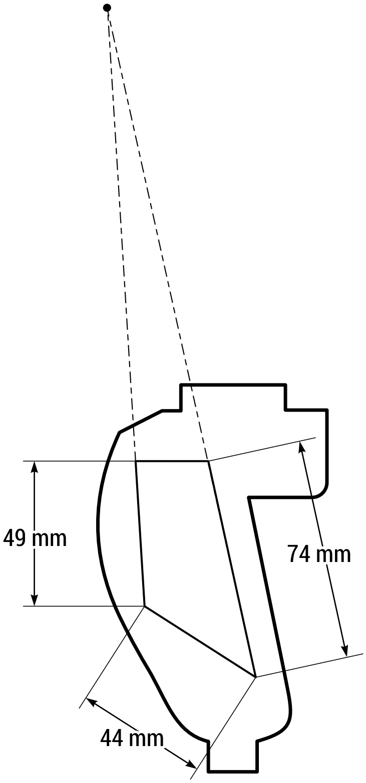


FIGURE 5.1. Example of the structural model adapted on the Össur Cheetah knee mechanism [19]

TABLE 5.1. A list of analysed knee mechanisms with basic technical data

No.	Manufacturer, Country	Name	Weight [g]	Height [mm]	Max. bend angle [°]	Max. user's weight [kg]	Patient activity level
1	Össur, Iceland	Balance Knee	590	160	154	125	K1, K2
2	Össur, Iceland	OP2	650	160	160	125	K2, K3
3	Össur, Iceland	Cheetah Knee	620	110	108	100	K4
4	Össur, Iceland	Balance OFM1	640	150	107	136	K1, K2
5	Össur, Iceland	Balance OM8	450	160	76	136	K1, K2
6	ST&G, USA	1321	652	135	108	125	K1, K2
7	Endolite, India	KX06	1230	160	230	150	K3, K4
8	Endolite, India	PSPC E4BKD	888	125	230	100	K1, K2, K3
9	D-Rev, USA	Remotion	618	160	154	80	K1
10	Proteor, Burgundy	1P200KD	743	170	195	125	K2, K3
11	Roadrunnerfoot, Italy	RRF Rodilla	520	165	195	130	K1, K2

5.2.1. Analytical method

To determine the trajectory of the instantaneous centre of rotation by the analytical method, SAM Designer (Artas Engineering Software, the Netherlands) and Excel (Microsoft Office) were used, where calculations based on the articulated quadrilateral model were the key assumption. The scheme of the analytical model of the articulated quadrilateral mechanism, which was used to determine the ICR coordinates, is presented in Fig. 5.2.

The L1, L2, L3 and L4 refer to the length of the connecting elements of the mechanism, θ2 and θ4 indicate the angles formed between the links L2, L4 and the X axis, and θ3 is the output angle of the link. The L1 link is treated as fixed, where the coordinates of the end points assume the values: A- (0.0), B- (L1.0). The alpha (α) is the angle formed between the link and the perpendicular line in its original position [4, 7, 23].

The coordinates of point C and D were determined by the following formulas [8]:

$$X_C = L_2 cosQ_2 , \tag{5.1}$$

$$Y_C = L_2 sinQ_2 , \tag{5.2}$$

$$X_D = L_2 cosQ_2 + L_3 cos\left(\alpha + Q_3\right), \tag{5.3}$$

$$Y_D = L_2 cosQ_2 + L_3 sin\left(\alpha + Q_2\right). \tag{5.4}$$

Alternatively, the following formulas can be used for point D:

$$X_D = L_1 + L_4 cosQ_4, \tag{5.5}$$

$$Y_D = L_4 sinQ_4. \tag{5.6}$$

After aligning and simplifying both forms of the X, Y coordinate formulas for point D, the following relationships are obtained:

$$L_4 cosQ_4 = L_2 cosQ_2 + L_3 \cos\left(\alpha + Q_3\right) - L_1, \tag{5.7}$$

$$L_4 cosQ_4 = L_2 cosQ_2 + L_3 \sin\left(\alpha + Q_3\right). \tag{5.8}$$

Then, using trigonometric functions, the equation is reduced to the following form:

$$AsinQ_2 + BcosQ_2 = C, \tag{5.9}$$

wherein

$$A = 2L_2L_3 \sin\left(\alpha + Q_3\right), \tag{5.10}$$

$$B = 2L_2L_3 \cos\left(\alpha + Q_3\right) - L_1, \tag{5.11}$$

$$C = L_4^2 - L_2^2 - \left(L_3 \cos\left(\alpha + Q_3\right) - L_1\right)^2 - \left(L_3 \sin\left(\alpha + Q_3\right)\right)^2. \tag{5.12}$$

The solution of the above equations refers to the tangent semi-angular trigonometric functions:

$$Q_2 = 2tan^{-1} \frac{A \pm \sqrt{A^2 + B^2 - C^2}}{B + C}, \tag{5.13}$$

$$Q_4 = tan^{-1} \frac{Y_D}{X_D - L_1}. \tag{5.14}$$

IC position coordinates are described by the following formulations:

$$Y_{IC} = X_{IC} tanQ_2, \tag{5.15}$$

$$Y_{IC} = X_{IC} tanQ_4 - L_1 tanQ_4. \tag{5.16}$$

Solving the above equations results in obtaining the final coordinate equations of the IC point:

$$X_{IC} = \frac{L_1 tanQ_4}{(tanQ_4 - tanQ_2)}, \tag{5.17}$$

$$Y_{IC} = X_{IC} tanQ_2. \tag{5.18}$$

The SAM Designer program was used to determine the changes in the angular values of the linkages (tan θ2 and tan θ4) necessary to generate the ICR trajectory. Based on the constructed structural models, the movement path of the mechanism was visualized during 90° bending. Then, a list with changes in angular values of linkages during bending and extension of the mechanism was generated by this program. A further calculation was conducted using the Excel-Microsoft Office program to define the X_{ic} and Y_{ic} coordinate values during the bending movement by subsequent degrees.

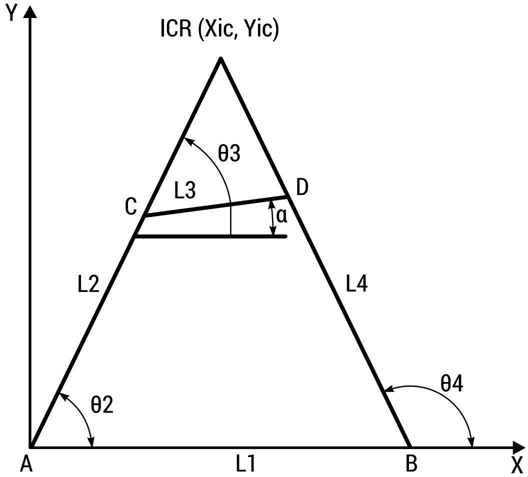


FIGURE 5.2. Scheme of the analytical model of the articulated quadrilateral mechanism [21]

5.2.2. Graphical method

To determine the trajectory of the instantaneous centre of rotation by the graphical method (diagrammatic), a student version of SolidWorks 2020 was used. The created structural models perform the functions of equivalents of real mechanisms, which are used for descriptive operations, thanks to which it was possible to generate the ICR path. Anchoring the appropriate segments allowed to simulate the rotation – i.e. bending the mechanism by an angle of 90 degrees, thereby to generate motion paths and, consequently, to draw the ICR trajectory.

5.3. Results

5.3.1. Analysis and comparison of analytical and graphical methods

In In the first part of the research the charts with changes in angular values of linkages (θ2 and θ4) during bending and extension (from 0° to 90°) of the analysed knee mechanisms were generated by SAM Designer Artras software together with a course of mechanism movement paths. The results obtained for the Cheetah mechanism are presented in Fig 5.3. Also, the values of the angular changes presented in the charts were converted into a number format in the form of lists.

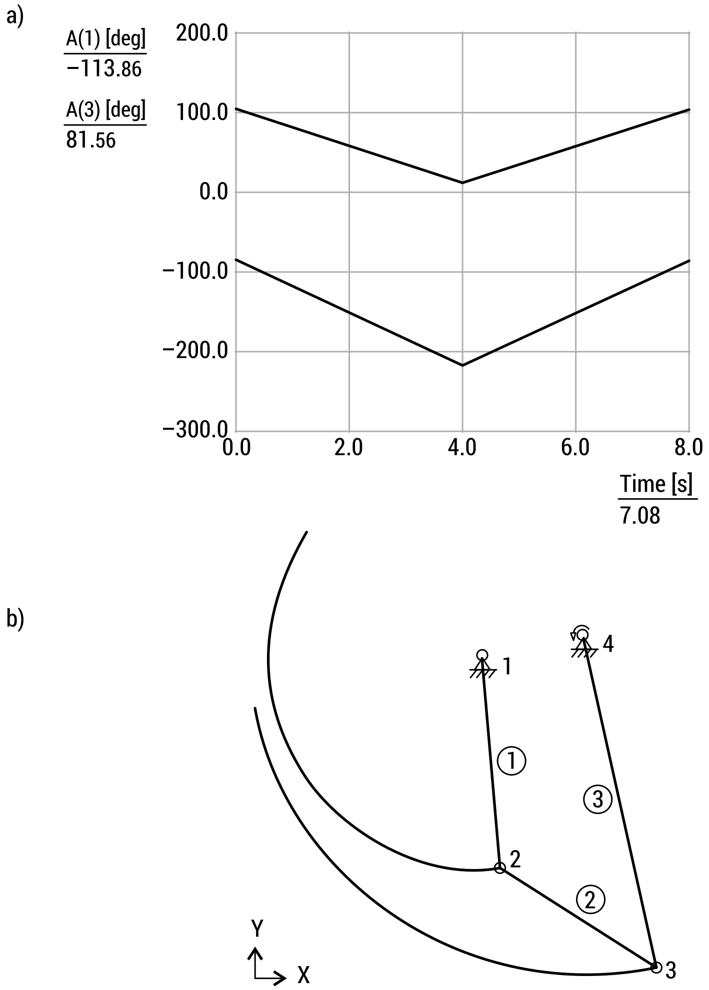


FIGURE 5.3. Angular changes chart (a) together with the trajectory of Cheetah mechanism segments (b) [19]

TABLE 5.2. Angular values of linkages and coordinates of the ICR

No.	**Time [s]**	*Q*4	*Q*2	*tanQ*4	*tanQ*2	X_{ic}	Y_{ic}
0	0	102.25	95.13	−4.60572	−11.1389	−30.3138	337.6631
1	0.11	99.75	91.59	−5.81966	−36.0258	−8.28457	298.4586
2	0.22	97.25	88.07	−7.86064	29.6757	9.004809	267.224
3	0.33	94.75	84.56	−12.0346	10.50064	22.96351	241.1316
4	0.44	92.25	81.07	−25.4517	6.364064	34.39877	218.916
5	0.56	89.75	77.57	229.1817	4.536933	43.86843	199.0281
6	0.67	87.25	74.08	20.81883	3.505882	51.70753	181.2805
7	0.78	84.75	70.59	10.88292	2.838073	58.1696	165.0896
8	0.89	82.25	67.1	7.347861	2.367332	63.43864	150.1803
9	1	79.75	63.6	5.530072	2.014487	67.63969	136.2593
10	1.11	77.25	60.09	4.419364	1.738351	70.88092	123.2159

In order to determine the ICR trajectory, numerical values were imported to the Microsoft Office Excel program, and then the X_{ic} and Y_{ic} coordinates were calculated, showing the changes in the position of the ICR. Table 5.2 shows exemplary final results. On this basis a graph showing the trajectory of a moving centrode was created (Fig. 5.4).

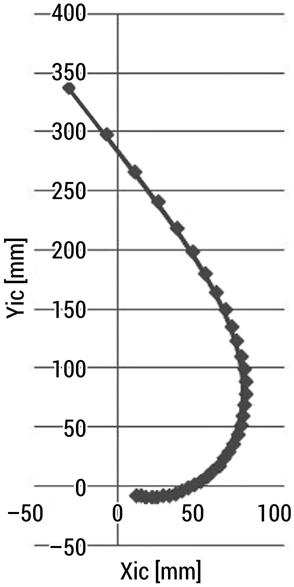


FIGURE 5.4. Trajectory of ICR – analytical method

The graphical method was also based on the created structural models, where after importing into the SolidWorks program, the ICR path with an immediate visualization of the results was generated (Fig. 5.5).

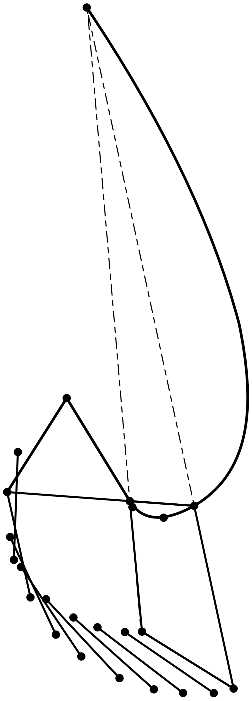


FIGURE 5.5. Trajectory of ICR – graphical method

In order to verify the accuracy of the results, the trajectories of the analysed mechanism obtained from the both methods were compared by superimposing (Fig. 5.6).

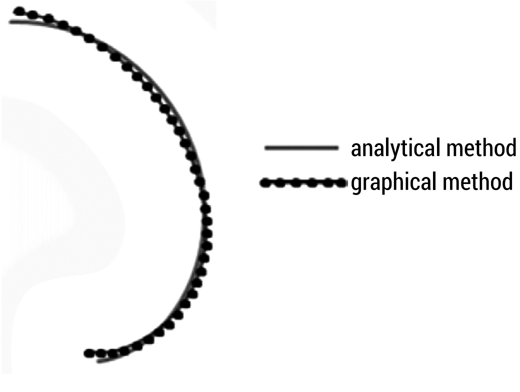


FIGURE 5.6. Comparison of ICR trajectories of all mechanisms – analytical method

For preliminary interpretation, the ICR trajectories of all mechanisms were also compiled in a graph to determine differences in stability and energy expenditure based on the shapes of the ICR trajectory and an initial point position (Fig. 5.7).

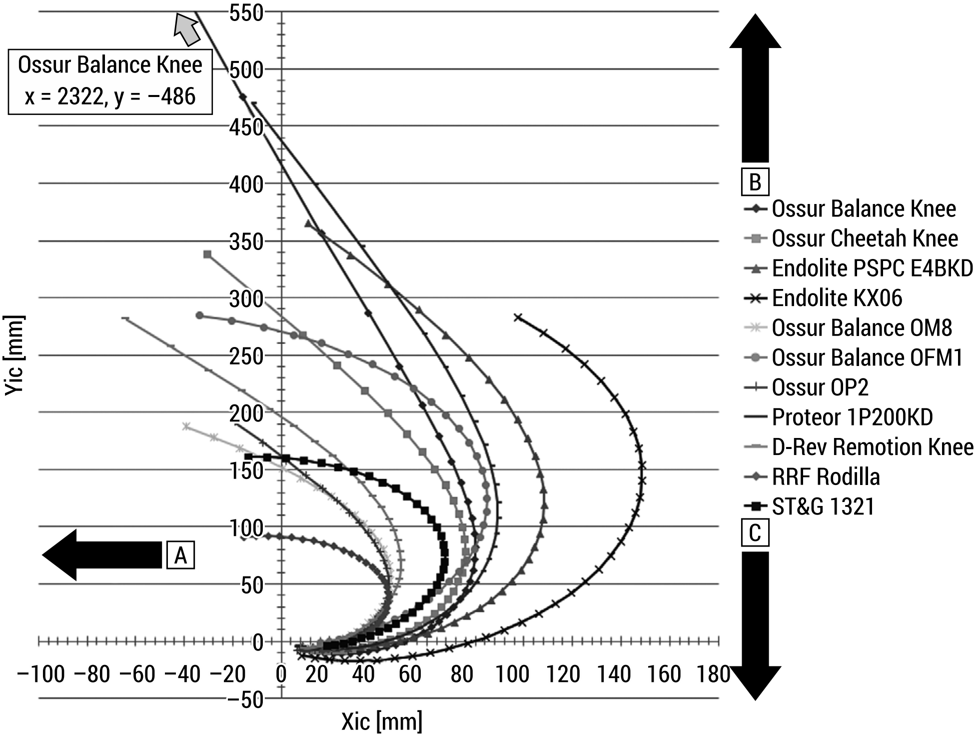


FIGURE 5.7 Comparison of the trajectory of the instantaneous centre of rotation (ICR) of analysed knee mechanisms, using the analytical method. A – increased support stability; B – decreased energy consumption when having to keep the knee straight while placing the foot on the ground, increased stump lever – greater control over the movement of the prosthesis; C – decreased energy consumption during flexion at the start of the gait cycle

5.3.2. Comparative analysis of stability and energy expenditure of knee mechanisms

At the end of this research, to evaluate the mechanisms with respect to stability and energy expenditure, a comparative list with scores (1–10, where 10 points is the best result) of both criteria was adopted. Tab. 5.3 contains a list of obtained points, considered individually for each mechanism. There is some additional information about the maximum bend angle and construction weight, which can be used as further aspects for assessment. For summary, the sum of points scored for each mechanism was calculated, according to which the mechanisms were ranked.

TABLE 5.3. Summary of mechanisms with comparative scoring

No.	Name	Patient activity level (K1–K4)	Stability (1–10)	Energy consumption (1–10)	Total points (1–20)
1	**Össur Balance Knee**	K1, K2	10	10	**20**
2	**Össur Cheetah** Knee	K4	7	7	**14**
3	D-Rev Remotion Knee	K1	9	5	**14**
4	**Össur Balance OFM1**	K1, K2	7	6	**13**
5	Proteor 1P200KD	K2, K3	4	9	**13**
6	**Össur Balance OM8**	K1, K2	8	4	**12**
7	Endolite PSPC E4BKD	K1, K2, K3	2	8	**10**
8	**Össur OP2**	K2, K3	4	4	**8**
9	Endolite KX06	K3, K4	1	6	**7**
10	ST&G 1321	K1, K2	4	2	**6**
11	RRF Rodilla	K1, K2	4	1	**5**

Fig. 5.8 presents a comparison between stability and energy consumption in the form of bar charts, generated in the Excel program. The charts are based on values related to the adopted scoring.

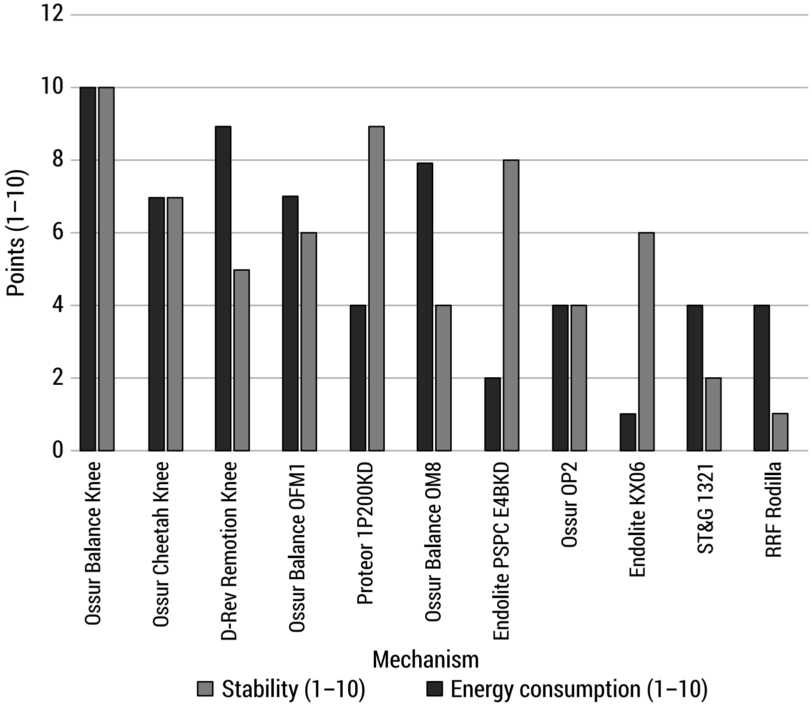


FIGURE 5.8. Comparison of stability – energy consumption

5.4. Discussion

The following article presents a comparison of ICR positions and the trajectories of the four-bar prosthetic knee mechanisms in relation to the stability and energy expenditure. These trajectories are a valuable source of information about the kinematic properties of the knee; therefore it is true that a correctly constructed prosthesis should reproduce the ICR trajectory design for the knee mechanism, and the greater the similarity to the natural knee, the better the performance of the prosthesis [3, 11]. However, it should be remembered, that the authors were guided by only two of the many important aspects of lower limb prosthesis.

Currently, there are no studies that clearly determine the optimal prosthetic knee mechanism for a specific user in terms of the above criteria, which would facilitate the selection for amputees. One of the closest studies is the analysis conducted by Munoz-Cezar and Hernandez-Gomez [16], where the authors carried out the optimization of the designed mechanism with the Taboo Search Algorithm. This work consisted in the exploration of random combinations of the length of the linkages within the range of stability, and based – on this information – the prosthesis was manufactured and adapted to a patient. Unlike the above mentioned research, the authors of this article proposed a visualization of ICR trajectories and an assessment of the suitability by graphical and analytical methods, not only in terms of stability but also of energy consumption. Energy expenditure is a measure of walking that reflects the functionality imparted by the prosthesis to the user. The measurement of energy consumption is sufficiently sensitive to changes in the designs of prosthetic components [6]. The paths from both methods are characterized by considerable similarity, which confirms the correctness and accuracy of the results. Further examples of research, such as a design of a linkage gait trainer [23] has been developed with the use of a four-bar mechanism to generate normal gait trajectories, where the linkage parameters are considered on the basis of the N-precision points to solve a loop-closure equation. The real trajectories are taken with respect to a global reference frame. However, the obtained dimensions should not exceed the design values proposed in this study and there is a need of the usage of control technology.

The variability of ICR path shapes testifies to a different functional property of mechanisms [4, 10]. The shapes and location of the ICR is characteristic of each mechanism, some paths had a gentle course, while others fell sharply at the initial stage of bending, where with each increment by subsequent degrees, the ICR moved in the direction: down and forward. Paths, which quickly fall close to the knee axis with the initial bend movement also provide good cosmetics while sitting. However, a too rapid fall is not safe as it leads to an increase in the moment of muscle forces responsible for control of the leg [6, 18]. The greater the elevation of the instantaneous centre of rotation towards the hip, the lower the value of the force required to manipulate the prosthesis, which facilitates the recognition of the ground before the transfer of weight to the artificial limb [3]. In order to design an efficiently working mechanism, the ICR should be placed as high as possible in the rearward direction when it is in a fully extended position, but also the appearance of the natural knee when bent by 90° should be preserved [22, 23]. For example, the Balance Knee mechanism obtained the greatest stability – its trajectory in the initial bending phase was the farthest from the axis. However, Remotion Knee, Balance OMF1, Cheetah Knee and Balance OM8 mechanisms also achieved good stability, while the lowest values were obtained by the Endolite PSPC E4BKD and Kx06 – knee joints for the most active users who have the strongest stump muscles and the greatest control over the prosthesis, so they do not need so much stability built into the prosthesis. It could even limit their activity. The knee for active users (K3, K4) should provide stabilization at the first 10° of flexion, for less active users (K1, K2) the stabilization should cover the range of up to 30° of flexion.

From the point of view of the energy expenditure, Össur Balance Knee, 1P200KD, PSPC E4BKD and Cheetah Knee were characterized by the lowest energy consumption necessary to initiate the mechanism bending – they had the highest elevated ICR points, while RRF Rodilla, 1321, Balance OM8 and OP2 mechanisms required higher energy input during usage. This is due to the fact that these are safe knee mechanisms intended for users with a low activity class, in this case a large amount of energy needed to bend the joint is a kind of brake and provides safety for a less active user of the prosthesis.

Additionally, by visualization of the ICR trajectories it was possible to determine mechanism classes. In this study, most of them were classified as a type with elevated and retracted initial ICR position during full extension, while the knees for which this point was located in the area in the front of the axis were defined as a mechanism with controlled flexion movement combined with low energy expenditure. The right selection of a prosthetic knee depends, among others, on the age, health condition, expected level of activity and lifestyle of the amputee [4]. The achieved results showed that the latest or advanced technologies (1P200KD, PSPC E4BKD, Kx06) will not be an ultimate solution for every person, they are mainly suitable for active people who expect a higher function level and the possibility of smooth movement at variable speeds and terrain. Surprisingly, good results were obtained by the Cheetah Knee prosthetic knee, which is intended for runners – despite the fact that it scored 7 points on the scale of stability and 7 on the scale of energy consumption, it is worth taking a closer look at the kinematics of this mechanism. The prosthetic knees like OP2 can be classified as four-axis mechanisms for typical, moderately active users, with a wide application in clinical prosthetics, which – due to increased support stability – allow for a natural, efficient movement. Mechanisms with very high stability (Balance Knee, Remotion Knee, Balance OM8, Balance OFM1) are intended for the elderly, less mobile people, for whom the sense of security is more important than functional performance. Prostheses of this design also help in the initial stages after the amputation procedure, while learning how to use a new knee, which is crucial to avoid discomfort and fear of tripping or falling [6].

A collective comparison of the mechanisms (Tab. 5.3) with the adopted scoring shows that the Balance Knee obtained the highest total points and the highest results, at the same time regarding the two analysed criteria. However, it should be remembered, that the authors were guided by only two of the many important aspects of the lower limb prosthesis. It should be also borne in mind that the instantaneous centre of rotation for the Balance Knee, at a 90° bend, is distant from the anatomical axis of the knee joint – which results in an unnatural position of the prosthesis, because these points should coincide. Due to the fact that such a large distance has to be covered, the ICR must be quickly moved towards the X axis during the bending movement in the joint. At the moment of an unpredictable event, when the burdened knee bends by several degrees, there is a sudden change in characteristics, which may be unintuitive and also dangerous for the user. One parameter is always obtained at the expense of the other, hence it is worthwhile to conduct a comprehensive evaluation of the mechanism to obtain its full characteristics.

It can be concluded that the most optimal solutions are Cheetah Knee and Balance OFM1 knees, which had similar values for both considered criteria. The least favorable parameter values were obtained by 1321 and Rodilla RRF knees.

5.5. Conclusions

This article has demonstrated how manipulation of the position of the ICR by appropriate length and arrangement of the linkages and the location of the axis contributes to the achievement of the desired features and help to understand how four-axis mechanisms affect stability, energy expenditure, comfort, and the cosmetics of gait of the amputee. This analysis, along with the indicated dependencies, may also lead to further research on the design method for the specific user of the prosthetic knee mechanism, as well as may be used to assess and compare the suitability of models available on the market.

Nevertheless, the presented study, which is necessarily based on a limited number of examples, clearly shows that the analysed mechanisms differ greatly in terms of the position and shape of the centrode. This means that the mechanisms have probably not been optimized in order to minimize the patient’s energy expenditure and also to ensure maximum stability during gait. This also means that the classic considerations of Radcliff and Green are still not reflected in the design process of modern knee mechanisms.

References

1. Andrysek J.: “Lower limb prosthetic technologies in the developing world: A review of literature from 1994–2010”, Prosthetics and Orthotics International, 34.4, 378–393, 2010.
2. Bulea T.C. Kobetic R.: “A variable impedance knee mechanism for controlled stance flexion during pathological gait”, IEEE/ASME Transactions on Mechatronics, 17, 822–832, 2012.

3. Cardenas A.M., Uribe J, Hernandez A.M.: “Parametric modelling of kinetic-kinematic polycentric mechanical knee”, VII Latin American Congress on Biomedical Engineering CLAIB, 60, 597–600, 2017.
4. Ceri C.N.: “Design analysis of the four-bar Stanford prosthetic knee for developing countries”, Department of Mechanical Engineering, Massachusetts Institute of Technology, 2013.
5. Copilusi C.P., Marin M.: “Design of a new human knee prosthesis based on cam mechanism”, DAAAM International, 21, 1726–9679, 2010.
6. Gard S.A.: “The influence of prosthetic knee joints on gait”, Handbook of Human Motion, 1359–1382, 2018.
7. Greene M.P.: “Four bar linkage knee analysis”, Orthotics and Prosthetics, 37, 15–24, 1983.
8. Hobson D.A., Torfason L.E.: “Computer optimization of polycentric prosthetic knee mechanisms”, Bull. Prosthet. Res., 10, 187–201,1975.
9. Hobson D.A., Torfason L.E.: “Optimization of four-bar knee mechanisms – a computerized approach”, J. Biomech. 7(4), 371–376, 1974.
10. Huiqun Fu, Zhang X.: “A novel prosthetic knee joint with a parallel spring and damping mechanism”, International Journal of Advanced Robotic Systems, 1–9, 2016.
11. Khorshidi M.: “Optimal design of four-bar mechanisms using a hybrid multi-objective GA with adaptive local search”, Mech. Mach. Theory; 46(10), 1453–1456, 2011.
12. Kramer S.N.: “Ortotyka. UT – the result of the corporation between university and industry”, 1997.
13. Lianfa Y., Jiating Ch., Zhifang Z.: “Design of a novel knee prosthesis mechanism with good stability”, International Robotics & Automation Journal, 4, 278–284, 2018.
14. Michael J.W.: ‘MEd.” Clinical Orthopaedics and Related Research, 361, 39–47, 1999.
15. Mohanty R.K., Sabut S.K.: “Polycentric prosthetic knee joint: A review”, International Journal of Engineering Science and Mathematics, 6, 503–510, 2017.
16. Munoz-Cesar J.J., Hernandez-Gomez L.H.: “Optimization of the design of a four bar mechanism for a lower limb prosthesis using the taboo algorithm”, Advances in Bio-Mechanical Systems and Materials, 40, 107–126, 2016.
17. Poliakov O., Lazarev V.: “Transfemoral prosthesis with polycentric knee mechanism: Design, kinematics, dynamics and control strategy”, Journal of Rehabilitation Robotics, 1, 109–113, 2013.
18. Poliakov O., Chepenyuk O.: “Multicriteria synthesis of a polycentric knee prosthesis for transfemoral amputees”, World Academy of Science, 6, 257–262, 2012.
19. Össur Master Prosthetic Catalog. Accessed: 24.02.2020.
20. https://res.cloudinary.com/ossur/image/upload/v1573570077/product-documents/en-us/PN20168/catalogs/PN20168_Cheetah_Knee.pdf.
21. Radcliffe C.W.: “Biomechanics of knee stability control with four-bar prosthetic knees”, ISPO Australian Annual Meeting, Melbourne, 2003.
22. Radcliffe C.W.: “Four-bar linkage prosthetic knee mechanisms: kinematics, alignment and prescription criteria”, Prosthetics and Orthotics International, 18, 159–173, 1994.
23. Radcliffe C.W.: “Prosthetic knee mechanism for above knee amputees”, Prosthetics and Orthotics Practice, 225–249, 1970.
24. Romo Duane H.: “Prosthetic knees, physical medicine and rehabilitation clinics of North America”, 11.3, 595–607, 2000.
25. Singh R., Chaudhary H., Singh A.K.: “A novel gait-based synthesis procedure for the design of 4-bar exoskeleton with natural trajectories”, Journal of Orthopaedic Translation, 12, 6–15, 2018.

26. Sun Y., Ge W., Zheng J., Dong D.: “Design and evaluation of a prosthetic knee joint using the geared five-bar mechanism”, Transactions on neural systems and rehabilitation engineering, 23, 1031–1038, 2015.

DOI: 10.24427/978-83-67185-49-3_6

Chapter 6

The impact of physiotherapy treatment in patients with balance disorders and dizziness

Robert Terlikowski, Anna Kuryliszyn-Moskal*, Katarzyna Kaniewska*, Andrzej Niewiński*, Paweł Konarzewski*, Robert Milewski**, Leszek Garbowski***, Jolanta Pauk****, Zofia Dzięcioł-Anikiej**

Medical University of Bialystok, Department of Rehabilitation, e-mail: robert.terlikowski@umb.edu.pl, a.kuryl@umb.edu.pl, katarzyna.kaniewska@umb.edu.pl, andrzej.niewinski@umb.edu.pl, pawel.konarzewski@umb.edu.pl, zofia.dzieciol-anikiej@umb.edu.pl

***Medical University of Bialystok, Statistics and Medical Informatics Institute, e-mail: robert.milewski@umb.edu.pl*

****Medical University of Bialystok, Department of Human Anatomy, e-mail: leszek.garbowski@umb.edu.pl*

*****Bialystok University of Technology, Institute of Biomedical Engineering, e-mail: j.pauk@pb.edu.pl*

Abstract: Balance disorders and dizziness, which are the results of the impairment function of the sensorimotor control system, are often symptoms mentioned by patients in GP's surgery. In regard to the fact that there are no effective pharmacological or surgical treatments for cervical dizziness, the purpose of this article is to verify if pharmacological and surgical treatment complemented with physiotherapy has a stimulating effect on the circuits responsible for posture control. Ambulatory patients – 47 patients with cervicogenic dizziness attended in the physiotherapeutic program, which lasted 10 days in a system of 5 days of therapy intertwined with 2 days off. In order to objectify the therapy results the efficiency of the balance system was evaluated, using the examination which registered forces and foot pressure moments on the ground during unperturbed stance. A statistical analysis of the stabilogram length in planta sagitalia (forward – backward) F-B and planta frontalia (right – left) R-L proved: with open eyes – statistically significant reduction of stabilogram length in planta frontalia (R-L) $p<0.043$; biofeedback – on the border of statistic significance $p<0.068$, a reduction of stabilogram length in planta frontalia (R-L). A statistical analysis of the speed of body stability stabilogram proved: with open eyes – statistically significant reduction of stabilogram speed in planta frontalia (R-L), $p<0.018$. A statistical analysis of the path length of the center of foot pressure proved: with open eyes – a statistically significant reduction ($p<0.017$) of the path length of the center of foot pressure. An analysis concerning the moment of foot pressure and forces pressured on the ground during unperturbed stance confirms that the application of physiotherapy had a stimulating influence on the systems responsible for postural control.

Keywords: vertigo, stabilography, vestibular rehabilitation, proprioception, biofeedback

6.1. Introduction

Balance disorders and dizziness resulting from the impairment of the sensorimotor system are common symptoms reported by patients at GP surgeries [1]. Spatial orientation is dependent on information coming from the sensorimotor control, which includes vestibular and ocular receptors as well as proprioceptors.

The processing and interpretation of impulses arriving from the peripheral sensorimotor system by the central nervous system ensure proper static and dynamic performance. Due to the fact that dizziness generates an increased number of falls in people of working age, thereby leading to a greater cost of treatment, absenteeism from work and deterioration of quality of life, it becomes one of the most important medical problems [1]. Dizziness has been categorized as a subjective symptom; therefore, there are many difficulties in identifying its cause. In about one quarter of the cases, the origin of dizziness cannot be accurately determined during the first diagnosis [2]. According to Drachman and Hart, there are three clinical conditions that, apart from the real vertigo, may be associated with the emergence of certain symptoms [3]. These include: presyncope (a feeling close to fainting with causes to be sought in the reduction of blood supply to the brain), disequilibrium (a feeling of unsteadiness associated with impairment of movement, or impaired proprioception) and lightheadedness (blackout sensations and confusion, a condition most likely of a psychological origin). Unfortunately, in Polish otolaryngologic terminology there are no equivalents specifying the aforementioned clinical conditions [4]. The term "dizziness" refers to illusions of motion (a sensation of environment spinning, spinning of one's own body, a sensation of rocking or swaying), an illusion of instability (unsteadiness of posture and gait, a sensation of chaos) and other ailments that do not correspond with the clinical meaning of "dizziness" (agoraphobia, fear of heights, flickering spots in vision). Whether the dizziness of cervical origin exists is still under discussion. The term "cervical dizziness" refers only to the fact that the disease is localized within the cervical spine [5]. Abnormalities within the muscular system of the cervical spine cause many diseases that have been described as idiopathic [6]. Flexion-extension neck injuries and a degenerative disease of the cervical spine are among most common causes of the cervical dizziness [7]. The consequence of the above mentioned conditions are abnormalities in muscle tension, muscle functions and even abnormalities in the structure of the suboccipital muscles, particularly the rectus capitis muscles, which undergo atrophy and fatty degeneration [8]. The relationship between prioperception of the cervical region and posture control has been confirmed by numerous studies in Europe and the United States [5]. A change from a vestibular sensation to a deep sensation structure reception is particularly evident in the elderly [8]. Disorders of the proprioceptive impulsation may arise in the pathologically tight muscles that manifest themselves in a lack of normal cervical reflexes compensation. The diagnosis of cervical dizziness rests in excluding other causes rooted in the vascular, vestibular or central nervous system [7]. Stabilometry is the most frequently used diagnostic tool that reveals, in an objective manner, deficits within the sensorimotor system. The test allows for the observation of changes in symptoms that result from postural control disorders by means of measuring and monitoring the deflection of the body [1]. Rehabilitation is a valuable supplementary treatment that goes with the traditional therapy for dizziness. Physiotherapy that involves stimulation of the deep sensation should not be limited only to the exercise of the cervical spine. Due to its comprehensive approach, the treatment activates the repair mechanisms of cervical proprioception and integrates afferent information coming from the vestibular system [9]. The aim of physiotherapy is to pave the neurotransmission of the deep sensation receptors and impact the sympathetic nervous system [10]. Subjective and objective symptoms associated with balance disorders are reduced within the process of compensation arising from the plastic possibilities of the central nervous system; the compensation may be accelerated and amplified by an appropriately selected physical therapy [11]. The facilitation and acceleration of the adaptation and compensation mechanisms that take place primarily at the level of the brainstem and cerebellum reduce feelings of dizziness and unsteadiness [12].

Treatment programs are aimed at improving the function of vestibular-oculomotor reflex, the perception of motion, postural stability and general physical condition [11, 13]. With regard to the fact that there are no effective pharmacological or surgical treatments for cervical dizziness, the purpose of this article is to verify if pharmacological and surgical treatment complemented with physiotherapy has a stimulating effect on the circuits responsible for posture control.

6.2. Methods and patients

6.2.1. Patients

The study involved 47 patients (24 F, 23 M) with systemic dizziness of cervical origin. All patients underwent diagnostic testing in the Department of Otolaryngology at the Medical University in Bialystok. As part of the diagnostic process, ENT (ear, nose and throat) ophthalmic and neurological, static, and dynamic tests were performed together with electronystagmography (ENG) with caloric stimulation, optokinetic nystagmus and Pendulum Test. In clinically difficult cases, radiological examinations were additionally performed, including CT or MRI of the head. Ambulatory patients with cervical dizziness were offered to participate in the physiotherapy program performed at the MUB Department of Rehabilitation (Resolutions of the Bioethics Committee of the R-I-002/405/2011). The program lasted 10 working days in a system of a five-day therapy intertwined with a two-day break.

6.2.2. Measurement protocol

In order to objectify the results of the therapy, the efficiency of the balance system was rated by testing the force and moment of foot pressure applied to the substrate (center-of-foot pressure, COP) during standing still. The study was performed twice: before and after the treatment. In order to objectivize the results of the therapy, the functioning of the balance system was tested by examining the horizontal shift of the body in the function of time by measuring the displacement of the point of application of the resultant reaction force of the substrate.

The study was performed twice: before and after the treatment. Balance platform PK, model 254 by Tecno Body was used for stabilometry test. The PRO KIN Line Software (included) enabled the recording and analysis of stabilometry test results [30]. Patients were tested twice, before and after the treatment, each test lasted 1 minute 30 seconds. During the first testing, the patient was instructed to stand still for 60 seconds and to focus his/her attention on a point that was 4 meters ahead in front of him/her; for another 60 seconds the patient had eyes covered with a band. In the second test the patient was able to control deflection of his/her body for the first 60 seconds in a function of time, watching the monitor screen located 0.5 m ahead at his/her eye level (biofeedback), the displacement of the point of application of the resultant reaction force of the substrate. For the next 60 seconds the patient was blindfolded. The test was done in a standing position, according to the manufacturer's instructions. The parameters to be recorded were: average deflection from axis X and Y, average deflection in the sagittal (F-B) and frontal plane (L-R), average swing speed (mm / sec), the distance traveled by the projection of the center of gravity (mm) field, ellipses (mm^2).

6.2.3. Therapy program

The therapy included a performance of a single set of rehabilitation exercises. A set of exercises was developed in the Department of Otolaryngology at the Medical University of Bialystok, with a recommendation to perform the mat home. The intensity was adjusted to the patient's condition – initially twice a day, and during the resolution of the symptoms the exercises were performed five times a day. The set of exercises allowed for practicing while lying down, sitting and standing. Another element of the program was physiotherapy sensorimotor training. The aim was to properly align the body (especially the head and the neck), which lasted up to 30 minutes, for 10 days (5 days of therapy and a 2-day break). The training was applied in accordance with the gradation of the difficulty principle: at first in static positions and with activation of smaller muscle groups, and later more intense, the so-called dynamic exercises on a shaky ground. The last elements of the treatment program were based on manual therapy techniques using: post-isometric relaxation and therapy of trigger points on the trapezius, levator scapulae, suboccipital muscles, scalene muscles, sternocleidomastoid muscles, the infraspinatus muscle as well as splenic, temporal and the platysma muscles. The mobilization to specific targeted sites was used on the articular joints C0-C1, C2-C3, and C6 to Th3 [14,15].

6.2.4. Statistical analysis

In order to compare the qualitative variables without normality, the Wilcoxon test for two dependent variables was used. Statistically significant results were considered at $p < 0.05$. The Statistica 10.0 package by StatSoft was used in the calculation.

6.3. Results

The statistical analysis of stabilometry in the sagittal (PT) and the front (EN) plane with eyes opened showed that there was a statistically significant reduction in the length of stabilometry in the frontal plane (LP), where p <0.043, but a statistically insignificant reduction in the length of stabilometry in the sagittal plane (PT), (Fig. 6.1).

With conscious control of visual inspection (biofeedback) there was a reduction in the length of stabilometry in the frontal plane (LP), where $p < 0.068$, but a statistically insignificant reduction in the length of stabilometry in the sagittal plane (P-T), (Fig. 6.2).

The statistical analysis of the velocity of deflection in stabilometry showed that: with eyes opened there was a statistically significant reduction of the deflection speed of stabilometry in the frontal plane (LP), where $p < 0.018$, but a statistically insignificant decrease in velocity in the sagittal plane. The statistical analysis of the distance traveled by the projection of the center of gravity showed that with eyes opened there was a significant reduction in the distance traveled by the projection of the center of gravity $p < 0.017$ and a statistically insignificant reduction with a conscious controlled visual (biofeedback) (Fig. 6.3).

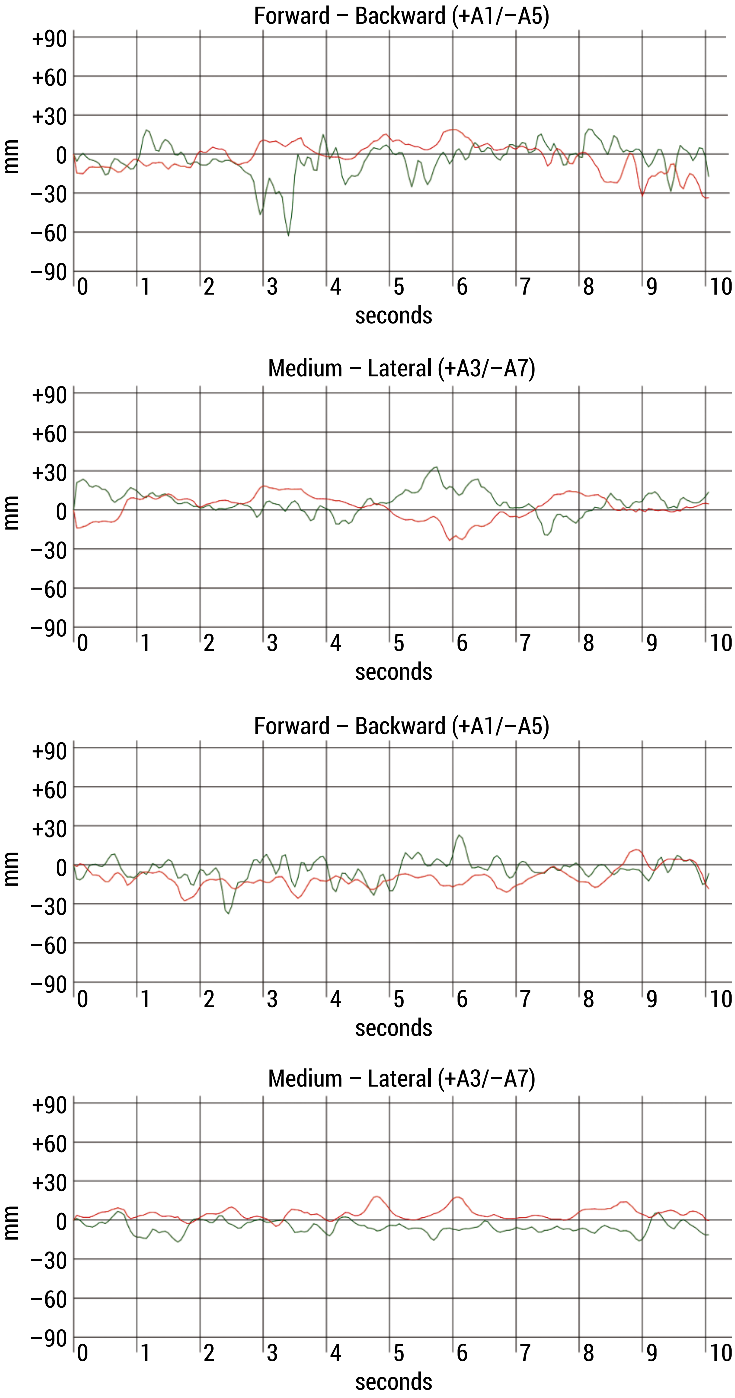


FIGURE 6.1. Example of examining the patient's deviation in directions forward-backward and medium-lateral without biofeedback before and after the therapy

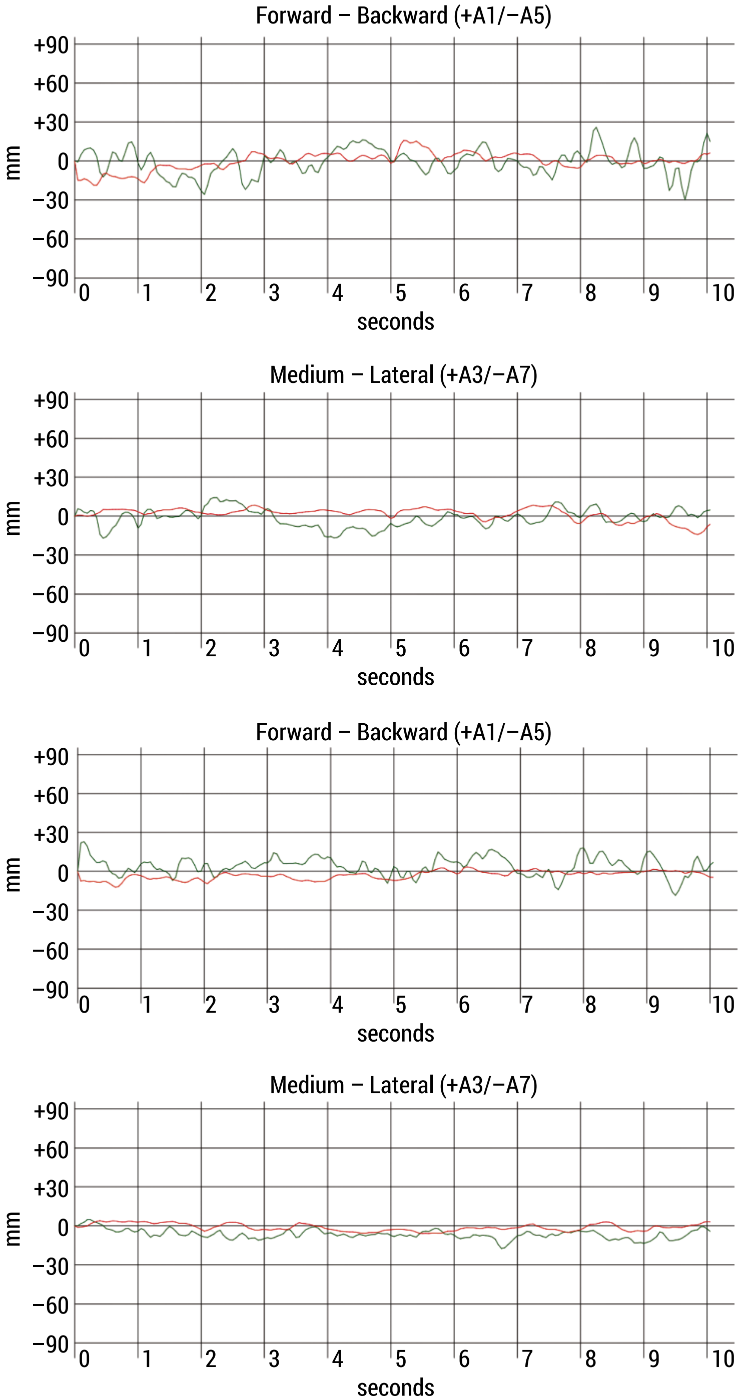


FIGURE 6.2. Example of examining patient’s deviation in directions forward-backward and medium-lateral with biofeedback before and after the therapy

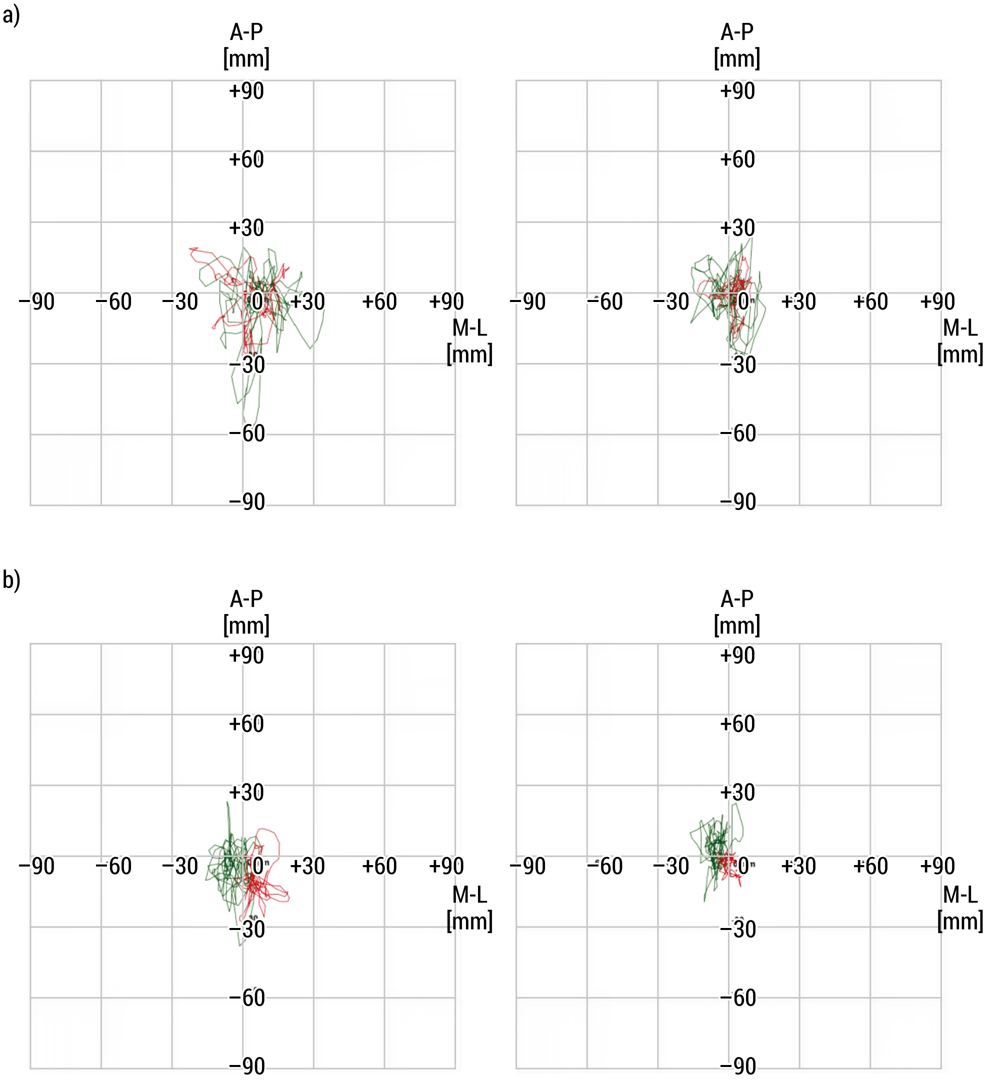


FIGURE 6.3. Example of stabilometry test: a) without biofeedback (red color – open eyes, green color – closed eyes); b) with biofeedback

The greatest reduction of the length of stabilograph in the frontal plane (LP) was obtained in the study in which the patient focuses on the point that is 4 meters ahead. There was a statistically significant difference in results of the examination before the treatment and after it, where $p < 0.043$ (Fig. 6.4).

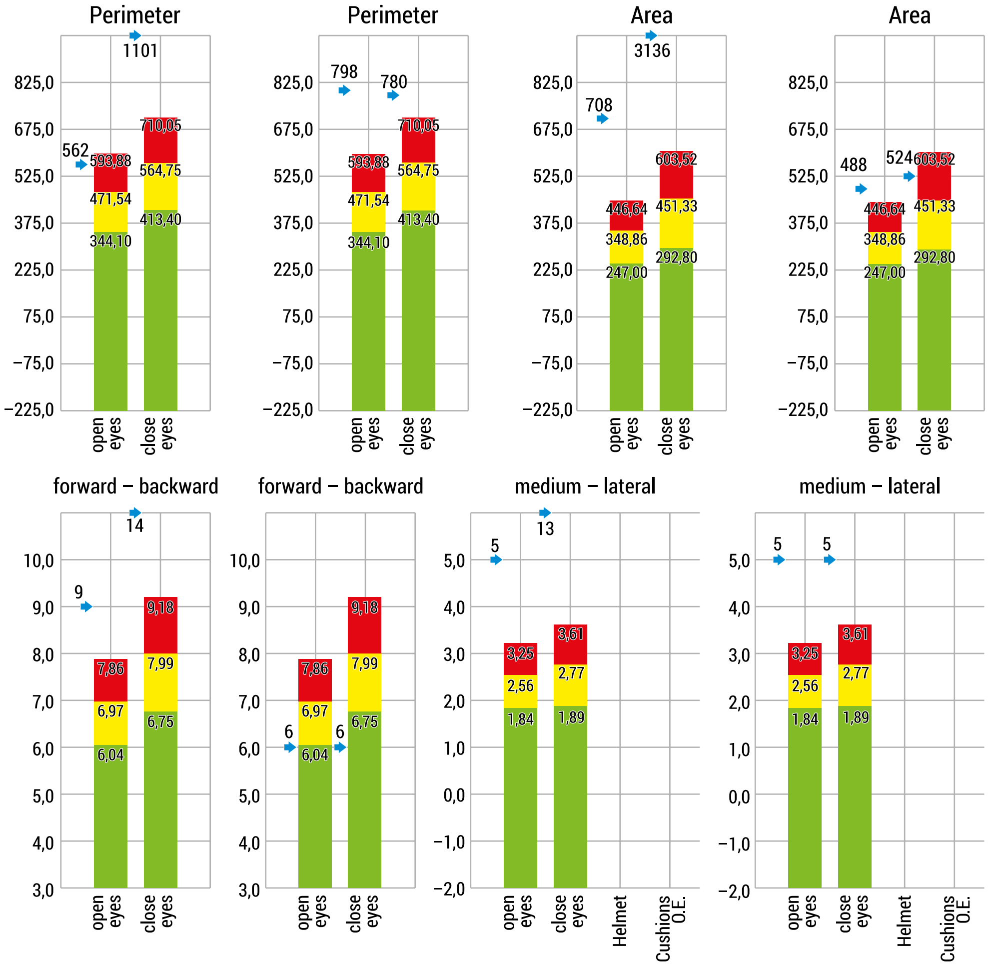


FIGURE 6.4. **Results of exemplary examination collective deviations in directions forward-backward and medium-lateral, perimeter of the trace and ellipse area achieved during Centre of Foot Pressure measurement without biofeedback before and after the therapy**

During the biofeedback test the patient observes the movement of the point of application to the resultant ground reaction force in the frontal plane (LP). An improvement within the range of statistical significance of $p < 0.068$ has been obtained (Fig. 6.5).

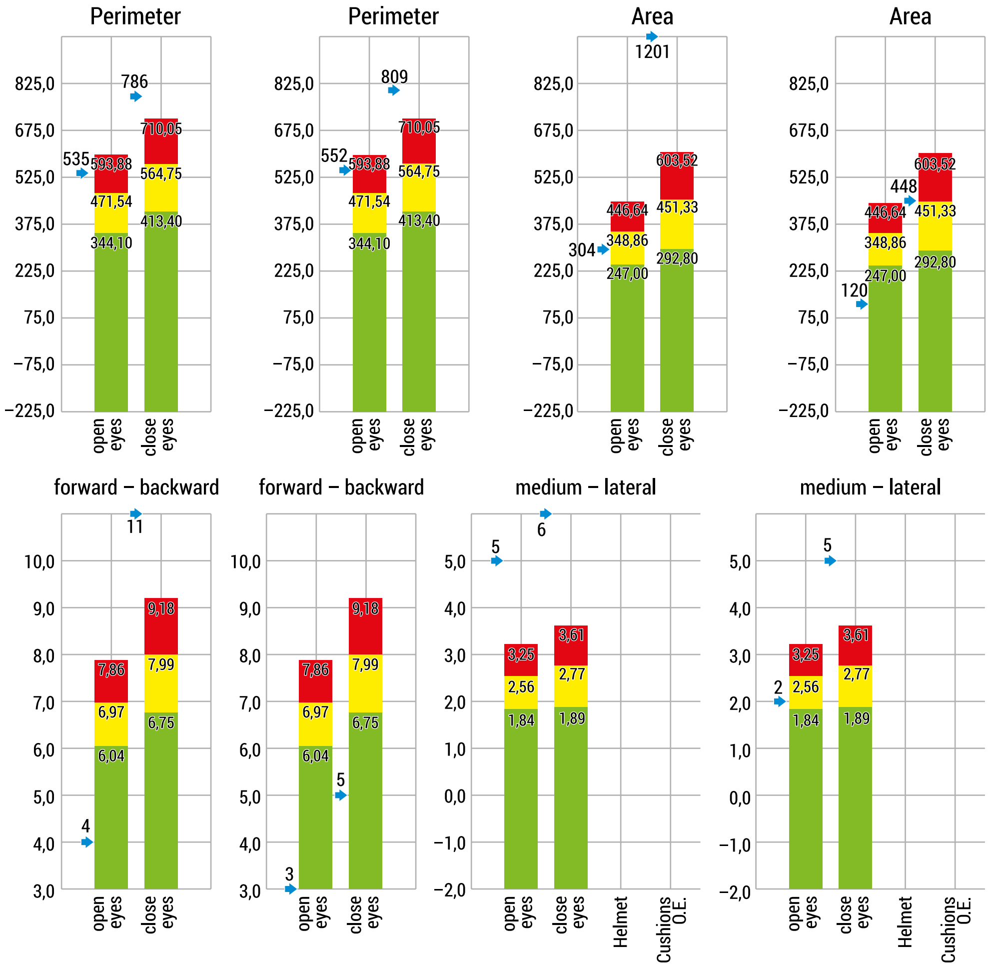


FIGURE 6.5. **Results of exemplary examination collective deviations in directions forward-backward and medium-lateral, perimeter of the trace and ellipse area achieved during Centre of Foot Pressure measurement with biofeedback before and after the therapy**

During the examination with eyes opened, there was a significant reduction in the velocity of deflection in the frontal plane (LP), where $p < 0.018$, and in the distance traveled by the projection of the center of gravity of $p < 0.017$. After physiotherapy the length of the stabilograph, the velocity of deflection in the frontal plane and the distance traveled by the projection of the center of gravity were shortened, which proves the normalization of deflection while maintaining body balance in a standing position.

6.4. Discussion

The terms "dizziness" and "balance disorders" in Europe and the United States do not mean the same; in some countries they are treated interchangeably, while in others one concept involves another. Therefore, there is a need for establishing diagnostic and therapeutic standards for dizziness and balance disorders. The decision to initiate an international debate over the classification of the vestibular dysfunction was undertaken during the 24th session of the Meeting of the Barany Society in Uppsala (Sweden) in 2006, as a result of which, in 2009, there was published a classification of disorders of the vestibular. That has given rise to the establishment of diagnostic and therapeutic standards for dizziness and balance disorders [20]. Since 2008, in Poland, within the framework of Vertigo forum meetings, a group of experts in otolaryngology, audiology, neurology and phoniatrics has been making efforts to obtain a common stance concerning the algorithms in dizziness and otolaryngological nomenclature [21]. Vestibular rehabilitation has been developing since the 1940s. It has been based on the studies that presented beneficial effects of physical exercise in the treatment of dizziness. Initially, physical exercises were used in the patients with persistent symptoms of unilateral vestibular damage. To the present day this type of rehabilitation has been significantly modified, and the indications for its use extended [10]. Rehabilitation uses brain plasticity, synaptic remodeling, natural phenomena, compensation vestibular adaptation and habituation. In a further phase of rehabilitation, the vestibular-cerebellar and vestibular core connections have become activated. The success of the vestibular compensation process depends on the timing and the extent of damage to the vestibular receptor as well as the patients' age and their physical activity. In the modern model of physiotherapy, a number of methods used in the treatment of disorders with cervical origin can be distinguished. These include: physiotherapy, prioperception training, exercise in eye-hand coordination on a posturographic platform, using a visual stimulation balance system in the feedback loop (biofeedback), manual therapy, post-isometric relaxation, trigger point therapy, physical therapy, massage, Kinesio Taping. This paper presents the process of rehabilitation of people with dizziness of cervical origin and proposes a set of rehabilitation exercises, as well as sensorimotor training, a selection of manual therapy techniques using: post-isometric relaxation, trigger point therapy of selected muscles and mobilization of targeted sites on joints C0-C1, C2 C3 and C6 to Th3. The treatment of patients with impaired postural stability is consistent with other researchers' observations which suggest that the proper selection of exercises using visual, proprioceptive and somatosensory stimulation allows the patient to learn how to skillfully use other vestibular information in order to maintain body balance [11, 22, 23, 24, 25]. The test results of the stability of posture (with eyes opened) showed a statistically significant ($p < 0.043$) reduction in the length of stabilograph in the frontal plane (LP). During the study, with a conscious visual inspection (biofeedback), the results were of borderline significance of $p < 0.068$ (LP). A statistical analysis of velocity of deflection in stabilometry showed a statistically significant decrease of stabilograph's deflection velocity in the frontal plane (LP), $p < 0.018$, with eyes opened. A statistical analysis of the distance traveled by the projection of the center of gravity showed a statistically significant ($p < 0.017$) reduction in the distance traveled by the projection of the center of gravity in the course of the study with eyes opened. The results confirm that the standing posture of the subject, with support on both lower extremities, is characterized by a greater deflection of posture in the sagittal plane than in the frontal plane. Proper education of patients plays a key role in the effectiveness of the treatment of dizziness and balance disorders [26, 27, 28]. Patients who understand the nature of their balance disorders better tolerate treatment and rehabilitation methods. Education contributes to the active participation of the patient in the therapeutic process. This improves the effectiveness of treatment and, consequently, brings satisfaction to both patients and therapists. ENT specialists and GPs appreciate the role of physical therapy in the treatment of dizziness and balance disorders caused by abnormalities of the vestibular system, and more often refer patients to this form of treatment [29].

6.5. Conclusions

The results of the study encourage further analyses over the supporting role of physiotherapy (in addition to pharmacological and surgical treatment) in cervical dizziness. With regard to the fact that there are no effective pharmacological or surgical treatments of cervical dizziness, the results of this study might inspire further detailed research.

References

1. Hansson E.E., Mansson O.N., Hakansson A.: "Effects of specific rehabilitation for dizziness among patients in primary health care, A randomized controlled trial", Clinical Rehabilitation, 18, 558–565, 2004.
2. Dommaraju S., Perera E.: "An approach to vertigo in general practice", AFP, 45, 4, 190–194, 2016.
3. Drachman D.A., Hart C.W.: "An approach to the dizzy patient", Neurology, 22, 323–324, 1972.
4. Narożny W., Siebert J., Wojtczak R.: "Epidemiologia zawrotów głowy i zaburzeń równowagi", Forum Medycyny Rodzinnej 4, 5, 356–365, 2010.
5. Wrisley D.M. et al.: "Cervicogenic dizziness: A review of diagnosis and treatment", Journal of Orthopaedic & Sports Physical Therapy, 30, 755–765, 2010.
6. Matsui T. et al.: "Cervical neuro-muscular syndrome: Discovery of a new disease group caused by abnormalieties in the cervical muscles", Neurologia Medico-Chirurgica (Tokyo), 52, 75–80, 2012.
7. Rubin W.: "Whiplash with vestibular involvement", Archives of Otolaryngology, 97, 85–87, 1973.

8. Chaitow L., DeLany J.: “Clinical application of neuromuscular techniques: The upper body”, Elsevier Health Sciences, London, 2008.
9. Włodarczy K., Pawlak-Osińska K.: “Rehabilitacja w zespole zawrotu szyjnego”, Otolaryngolia Polska, 9, 2, 55–61, 2012.
10. Rzewnicki I., Rogowski M.: „Rehabilitacja ruchowa w leczeniu zawrotów głowy i zaburzeń równowagi”, Polski Merkuriusz Lekarski, 24, 141, 244–246, 2008.
11. Mraz M. et al.: “Body balance in patient with systemic vertigo after rehabilitation exercise”, Jour of Physiology and pharmacol, 58, 5, 427–436, 2007.
12. Brandt T.: “Vestibular nerve and labyrinthine disorders” [in:] “Vertigo, ITS multisensory syndromes”, Springer 2000, 175–195, 2000.
13. Tacikowska G.: “Rehabilitacja ruchowa (rehabilitacja przedsionkowa) w zawrotach głowy i zaburzeniach równowagi”, Nowa Audiofonologia, 1, 1, 107–111, 2012.
14. Emary P.: “Use of post-isometric relaxation in the chiropractic management of a 55-year--old man with cervical radiculopathy”, The Journal of the Canadian Chiropractic Association, 56, 1, 9–17, 2012.
15. Abrahams VC., Richmond F.J.: “Specialization of sensorimotor organization in the neck muscle system”, Progress in Brain Research, 76, 125–135, 1988.
16. Ptaszkowski K. et al.: “Comparison of the short-term outcomes after postisometric muscle relaxation or kinesio taping application for normalization of the upper trapezius muscle tone and the pain relief: A preliminary study”, Evidence-Based Complementary and Alternative Medicine, doi: 10.1155/2015/721938, 2015.
17. Moraska A.F. et al.: “Myofascial trigger point-focused head and neck massage for recurrent tension-type headache: a randomized, placebo-controlled clinical trial”, The Clinical Journal of Pain., 31, 2, 159–168, 2015.
18. Fernández-de-Las-Peñas C., Courtney C.A.: “Clinical reasoning for manual therapy management of tension type and cervicogenic headache”, Journal of Manual and Manipulative Therapy, 22, 1, 44–50, 2014.
19. Bron C. et al.: “Treatment of myofascial trigger points in patients with chronic shoulder pain: A randomized, controlled trial”, BMC Medicine, 9:8, 2011.
20. Bisdorffa A. et al.: “Classification of vestibular symptoms: Towards an international classification of vestibular disorders”, Journal of Vestibular Research, 19, 1–13, 2009.
21. Narożny W. et al.: „Mianownictwo anatomiczno-histologiczne ucha wewnętrznego – przegląd piśmiennictwa neurootologicznego”, Vertigo Forum, 2, 1, 19–30, 2010.
22. Crutchfield C., Shumway-Cook A., Horak FB.: Balance and coordination: Physical therapy, New York, NY: Lippincott, 825–843, 1989.
23. Cass S.P., Borello-France D., Furman J.M.: “Functional outcome of vestibular rehabilitation in patients with abnormal sensory-organization testing”, The American Journal of Otology, 17, 581–594, 1996.
24. Horak F.B. et al.: “Effects of vestibular rehabilitation on dizziness and imbalance”, Otolaryngology – Head and Neck Surgery, 106, 2, 175–180, 1992.
25. Badke M.B. et al.: “Effects of vestibular and balance rehabilitation on sensory organization and dizziness handicap”, The Annals of Otology, Rhinology and Laryngology, 114, 48–54, 2005.
26. Pośpiech L.: “Praktyczne podejście do rehabilitacji zawrotów głowy”, Solvay Pharma, Biblioteczka ProsperaMénière’a, Warszawa 4, 1–65, 1997.
27. Serafini G., Palmierei AMR., Simoncelli C.: “Benign paroxysmal positional vertigo of posterior semicircular canal: Results in 160 cases treated with Semont’s manoeuvre”, The Annals of Otology, Rhinology and Laryngology, l105, 770–5, 1996.

28. Bhattacharyya N., Baugh R.F., Orvidas L.: "Clinical practice guideline benign paroxysmal positional vertigo", Otolaryngology Head and Neck Surgery, 139, 5, 47–81, 2008.
29. Józefowicz-Korczyńska M.: "Vestibular and disability rehabilitation therapy", Otolaryngologia Polska, 9, 1, 7–1, 2010.
30. Zhang M., You H., Zhang H., Zhao W., Han T., Liu J., Jiang S., Feng X.: "Effects of visual feedback balance training with the Pro-kin system on walking and self-care abilities in stroke patients", Medicine (Baltimore), 2020 Sep, 25, 99 (39).

DOI: 10.24427/978-83-67185-49-3_7

Chapter 7

Comparative analysis of the vertical component of ground reaction force during gait of healthy people and people with Parkinson’s disease

Karolina Tymińska, Marcin Derlatka

Bialystok University of Technology, Institute of Biomedical Engineering,
e-mail: tyminska.karolina@o2.pl, m.derlatka@pb.edu.pl

Abstract: One of the symptoms of Parkinson’s disease is an abnormal gait cycle. Locomotion becomes slower, less sure, and gait very often leads to falls. This paper analyses the vertical component of the ground reaction force and pressure force for healthy subjects and those with Parkinson’s disease. The data used in the paper comes from the physionet.org website and includes the results of tests that were conducted with Ultraflex Computer Dyno Graphy system Infronic insoles. The dataset contains recorded ground reaction forces during gait of 164 people, including 71 healthy people and 93 people with Parkinson’s disease at varying degrees of disease severity as determined by the Hoehn-Yahr scale. The analysis showed statistically significant differences ($\alpha \leq 0.05$) between the highest value and the local minimum of the vertical component of the ground reaction force. The course of the vertical component of the ground reaction force changed with increasing disease progression, characteristic gait phases disappeared, and the foot contacted the ground in a different way during the heel strike phase. This is due to the development of an abnormal gait pattern, which protects against falls and helps to maintain independence for longer. Gait testing and analysis of the vertical component of the ground reaction force can help to detect the disease earlier and therefore to implement treatment more quickly, resulting in a slower onset of symptoms.

Keywords: Parkinson’s disease, ground reaction forces, human gait, statistical analysis

7.1. Introduction

Parkinson’s disease is most commonly associated with a resting tremor of one hand. However, a number of other motor symptoms can be distinguished in this condition, such as slowing down and adopting a characteristic posture [7]. The patient is hunched over, bent forward, and has bent elbow and knee joints even at rest. Among the symptoms, gait disturbances can also be distinguished. Locomotion becomes slower, less sure, and very often leads to falls. Individuals also have impaired automaticity and movement control [5].

During locomotion of people with Parkinson's disease there are permanent and episodic disorders [7]. The former are manifested by shuffling, moving with small steps, and disabling the work of the upper limbs. Episodic disorders are more dangerous because they appear unexpectedly. Among these are walking, freezing and falling [2]. Trotting occurs when the patient makes small movements with the lower limbs to stabilize the center of gravity. Freezes of gait, on the other hand, include shuffling, trembling and even akinesia, i.e. a complete lack of movement [2]. Increasing symptoms as the disease progresses can lead to loss of independence in daily life. Although Parkinson's disease is not fully curable, early detection allows symptoms to slow down and independence to be maintained for longer [3, 6].

At the first stage of Parkinson's disease, symptoms are not easy to observe. Only at an advanced stage are they visible to the naked eye. Diagnosis is clinical, based on observation, interview, and examination. Unfortunately, there is no test that can uniquely identify Parkinson's disease [9]. A person with suspected Parkinson's disease is usually sent for neuroimaging examinations, such as computed tomography and magnetic resonance imaging [9]. This is followed by a levodopa test, which involves giving the patient this drug and observing if there is an improvement in the condition [9]. The next option is to use preclinical diagnostics – genetic, olfactory, neuropathological and neuropsychiatric. The former involves performing genetic diagnostics to see if there is a familial occurrence of the disease. The second is olfactory testing to rule out hypotonia (weakness of the sense of smell, which is one of the most common symptoms). The latter consists of a psychological examination to identify motor and cognitive abnormalities [9].

Early diagnosis of Parkinson's disease is very important as it allows for slower progression of symptoms and longer independence in daily life. One way of early diagnosis may be to assess individual gait parameters. The aim of this article is to distinguish between the vertical component of the ground reaction force in healthy people and those with Parkinson's disease, which may serve as a potential tool for the early diagnosis of this disease.

7.2. Materials

The data for the analysis came from the physionet.org website. The database contains records of the vertical ground reaction forces of the subject people and the contact forces on the different parts of the foot, conducted by Jeffrey Hausdorff.

The study involved 164 people, including 71 healthy people and 93 people with Parkinson's disease at different stages of the disease. All people were of similar age, height and body weight (Tab. 7.1).

TABLE 7.1. Data of the people analyzed

Study group	Number of people	Stage of advancement [Hoehn – Yahr scale]	Body weight [kg]	Body height [m]	Age [years]
Parkinson's	55	2	72 ± 12	1.68 ± 0.09	69 ± 9
	27	2.5	74 ± 11	1.66 ± 0.06	61 ± 9
	9	3	70 ± 12	1.67 ± 0.07	65 ± 5
Healthy	71	–	73 ± 12	1.68 ± 0.08	64 ± 8

The study was conducted using Ultraflex Computer Dyno Graphy insoles, an Infronic system. The insoles have eight sensors that record the output signal as the pressure force of a given part of the foot on the ground every 0.01s (Fig. 7.1). In addition, in the recording of the whole test, there are two more signals summing the pressure forces from all sensors, hence it is possible to approximate the course of the vertical component of the ground reaction force.

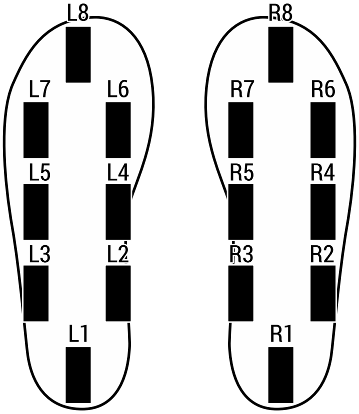


FIGURE 7.1 Location of sensors

People walked at a leisurely, self-selected pace for approximately two minutes. The walk was on flat ground. The subjects' footwear included insoles to record pressure.

The anthropometric data included parameters such as age, weight, body height and stage of the disease. It should be noted that in the degrees of the disease analyzed in this study, symptoms occur bilaterally. Therefore, right and left feet were compared for particular groups.

7.3. Method

Among the results obtained from the phisionet.org free database containing the studies conducted by Jeffrey Hausdorff [1], the results of three subjects were randomly selected and the vertical component of ground reaction force of a healthy person and two subjects with Parkinson's disease (with the second and third stages of the disease according to the Hoehn-Yahr scale) were analyzed. One middle gait cycle was selected from each study to compare the course of the vertical component of the ground reaction force for all patients. The obtained values from a given cycle were then normalized to body weight, and the studied parameter was presented on a graph. Statistical analysis was then performed, setting the level of statistical significance at $\alpha \leq 0.05$.

7.4. Analysis of results

The graph below (Fig. 7.2) shows the vertical component of the ground reaction force for a healthy person. It shows two local maxima and one minimum. The first major peak is related to heel shock absorption during impact.

The first peak is related to heel shock absorption during impact, whereas the second peak indicates toe lift off the ground. The local minimum indicates when there is a transfer of the body over the limb in contact with the ground and toe-off of the opposite foot.

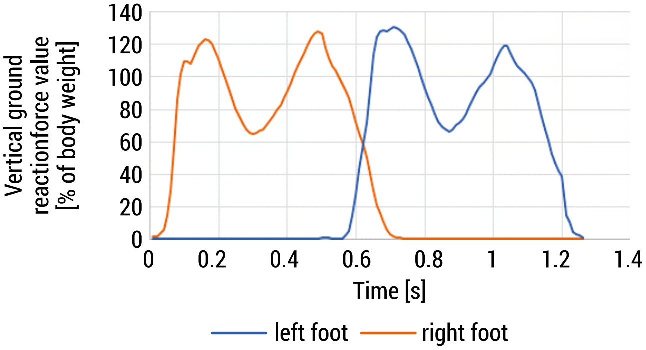


FIGURE 7.2. Vertical component of the ground reaction force for a healthy person

The course of the vertical component of the ground reaction force for a healthy person is similar to the data described in [10]. The gait is normal and all characteristic phases of the gait cycle are preserved.

The next graph (Fig. 7.3) shows the course of the vertical ground reaction force for a person with the second stage of the disease according to the Hoehn-Yahr scale. Only one local maximum and minimum can be observed here. There is no clear increase in the force values at the time of toe-off.

When comparing the vertical component of the ground reaction force of healthy people and patients with the second stage of the disease, large differences are apparent. There are smaller differences in the values for the different phases of the gait cycle in the diseased people. The second maximum also disappears.

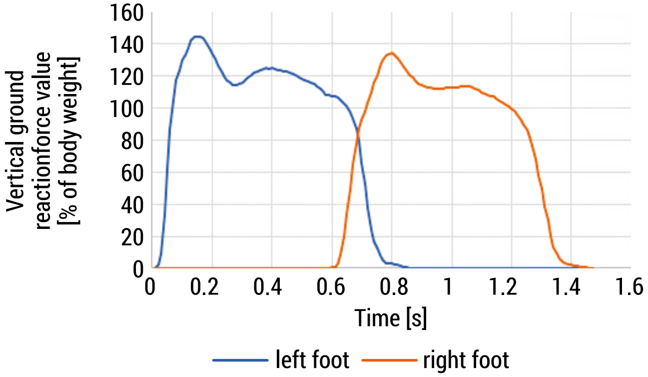


FIGURE 7.3. The vertical component of the ground reaction force for an ill person with second-degree Hoehn-Yahr scale

Next, the vertical component of the ground reaction force was analyzed for a patient with stage 3 of the disease. Here, a much smoother pattern of values can be seen, with local minima and maxima almost disappearing (Fig. 7.4).

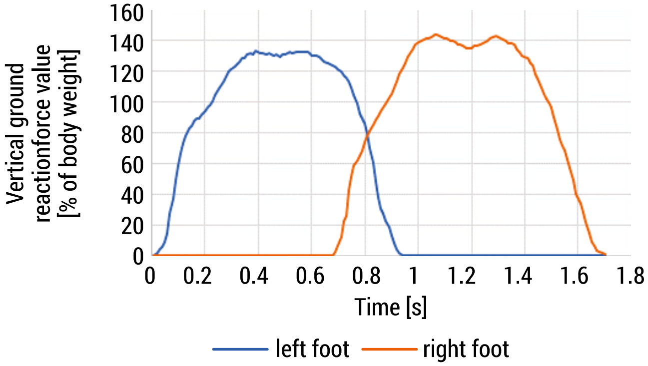


FIGURE 7.4. The vertical component of the ground reaction force for an ill person with third-degree Hoehn-Yahr scale

Analyzing the graphs (Fig. 7.2–7.4), it can be seen that the vertical component changes with the stage of the disease. Individual characteristic gait phases disappear, which may be related to the different derailment of the foot in people with Parkinson's disease. In healthy people, first there is a heel strike, then a transfer of the body over the foot, and then an explosion from the toes. In sick people, on the other hand, this division gradually disappears. People with a more advanced stage of the disease make contact with the ground with the whole foot at once, which results in a characteristic thud.

The greatest differences are seen in local minima and maxima. Therefore, the difference between the highest value and the local minimum was calculated, and then a statistical analysis was performed, to point out the potential statistically significant differences between both groups.

TABLE 7.2. Analysis for maximum and local minimum values of the vertical component of the ground reaction force

People	Shapiro – Wilk test	Type of distribution	Levene's test	Test	p
Left foot					
ILL	0.007	deviate from normal	–	U Mann – Whitney test	**<0.01**
Healthy	0.209	normal			
Right foot					
ILL	0.143	normal	–	U Mann – Whitney test	**<0.01**
Healthy	0.026	deviate from normal			

The statistical analysis (Tab. 7.2) has shown that the differences seen in the graphs in local minima and maxima are statistically significant. Due to the presence of Parkinson's disease symptoms bilaterally and the lack of a clear indication in the data for the affected side, the analysis was performed for both feet, but for the two limbs the statistical differences are the same.

Subsequently, a statistical analysis was performed for healthy people and those suffering from Parkinson's disease, taking into account the severity of the disease according to Hoehn-Yahr scale. In the table (Tab. 7.3), the nomenclature "ILL 2" was used, which means a patient with the second stage of the disease. This applies similarly to the other people.

TABLE 7.3. Statistical analysis for the maximum and local minimum of the vertical component of the ground reaction force for healthy and ill people with different degrees of disease – left foot

People	Shapiro – Wilk test	Type of distribution	Levene's test	Test	p
Healthy	0.209	normal	**<0.001**	Kruskal – Wallis test	**<0.001**
ILL 2	0.094	normal			
ILL 2.5	0.490	normal			
ILL 3	0.666	normal			

Table 7.3 shows that there are statistically significant differences for the studied patients. In view of this, the POST-HOC (Scheffe) test was performed to see between which groups statistically significant differences occur.

TABLE 7.4. Scheffe test – left foot

	Healthy	**ILL 2**	**ILL 2.5**	**ILL 3**
Healthy		**<0.001**	**<0.001**	**<0.001**
ILL 2	**<0.001**		0.8202	0.4049
ILL 2.5	**<0.001**	0.8202		0.8684
ILL 3	**<0.001**	0.4049	0.8684	

Table 7.4 shows that statistically significant differences occur between healthy and sick people regardless of the severity of the condition in the case of the left foot.

TABLE 7.5. Statistical analysis for the maximum and local minimum of the vertical component of the ground reaction force for healthy and ill people with different degrees of disease – right foot

People	**Shapiro – Wilk test**	**Type of distribution**	**Test**	**p**
Healthy	0.027	deviate from normal	Kruskal – Wallis test	**<0.001**
ILL 2	0.496	normal		
ILL 2.5	0.404	normal		
ILL 3	0.165	normal		

Table 7.6 shows that there are statistically significant differences, as in the case of the left foot. Again, the POST-HOC test was performed to test statistical significance between groups.

Analyzing table 7.5 it can be seen that, again, statistically significant differences exist between healthy people and sick people with each degree of the condition.

TABLE 7.6. Scheffe test – right foot

	Healthy	**ILL 2**	**ILL 2.5**	**ILL 3**
Healthy		**<0.001**	**<0.001**	**<0.001**
ILL 2	**<0.001**		0.7263	0.6512
ILL 2.5	**<0.001**	0.7263		0.9992
ILL 3	**<0.001**	0.6512	0.9992	

The figures 7.5 below show the recorded pressure force values on each sensor. The results are shown for a selection of three people: a healthy person and two people with different degrees of disease.

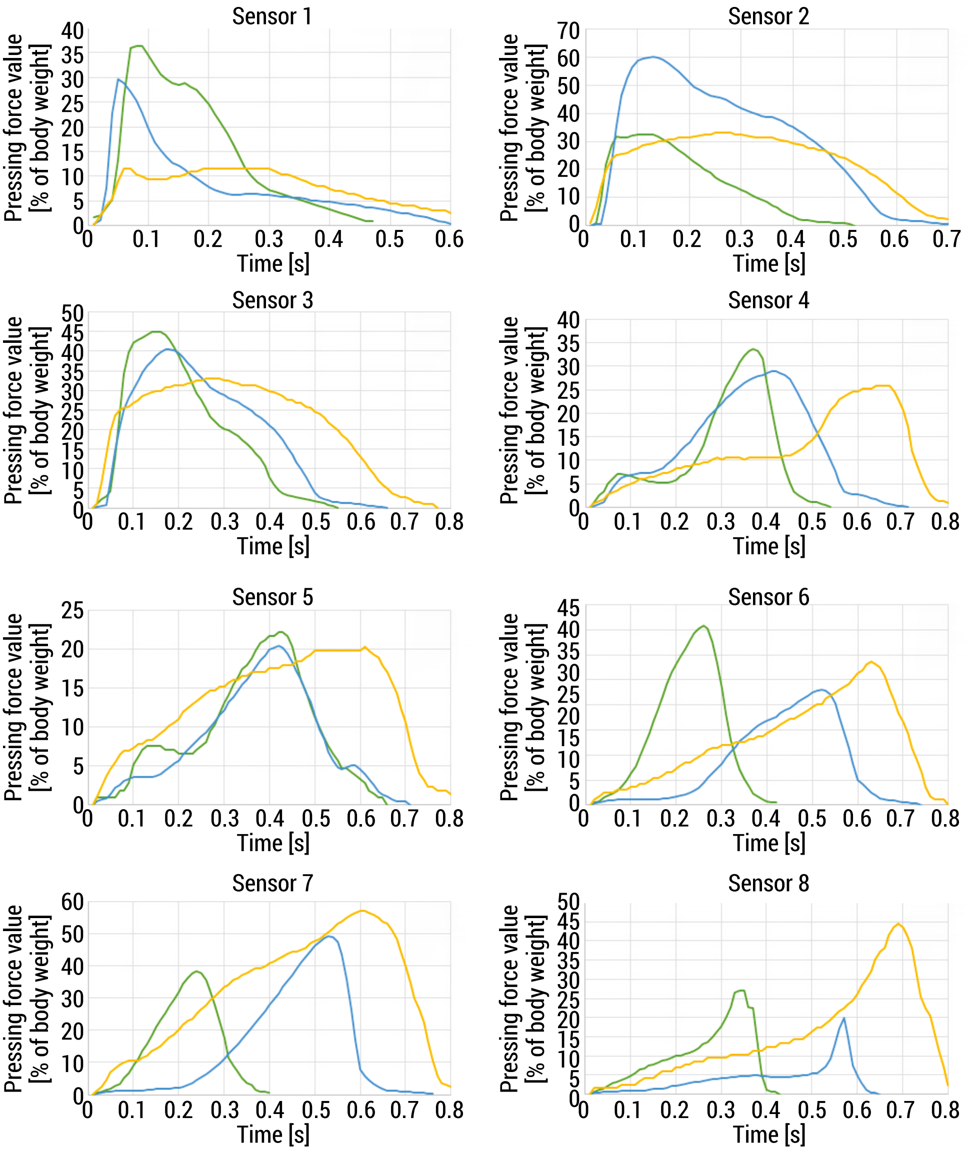


FIGURE 7.5. Course of pressure forces on each sensor for three analyzed people (green – healthy person, blue – person with second stage of disease, yellow – person with third stage of disease according to Hoehn-Yahr scale)

It can be observed (Fig. 7.5) that in the initial phase of gait (sensor 1) higher thrust forces are achieved by healthy people, whereas in the final phase of the gait cycle (sensor 8) thrust forces are higher sick for people with a third-degree disease.

Noting that the waveforms of values for each person are different, a statistical analysis was carried out to test whether the differences visible to the naked eye are statistically significant.

The highest and the mean value of the contact force were determined for statistical analysis. Initially, the normality of the distribution was checked and an appropriate statistical test was selected. The value of p indicated whether the differences in contact force values were statistically significant. Statistical significance is highlighted in red in the table below. The level of statistical significance was set at $\alpha \leq 0.05$.

TABLE. 7.7. Statistical significance of pressure force values for individual sensors

Number of sensor	**Statistical significance**			
	Highest value of contact force		**Average value of contact force**	
	Type of test	**Statistical significance**	**Type of test**	**Statistical significance**
1	ANOVA	**p < 0.001**	ANOVA	**p < 0.001**
2	ANOVA	**p = 0.0021**	ANOVA	**p = 0.0014**
3	Kruskal – Wallis test	p = 0.1356	Kruskal – Wallis test	p = 0.8890
4	ANOVA	p = 0.1065	Kruskal – Wallis test	p = 0.1042
5	ANOVA	**p < 0.001**	ANOVA	p = 0.1075
6	ANOVA	**p = 0.0019**	ANOVA	p = 0.5753
7	ANOVA	p = 0.3234	ANOVA	p = 0.5147
8	Kruskal – Wallis test	p = 0.7317	Kruskal – Wallis test	p = 0.1931

Table 7.7 shows that statistically significant differences occur for sensors 1 and 2. Additionally, they also occur for sensor 5 and 6 for the highest values of contact forces. In other cases, the tested differences were not statistically significant. For sensor 8, a difference can be seen in the graph, but the POST-HOC test showed that it is not statistically significant.

The POST-HOC test (Scheffe) was then performed for the sensors that recorded statistically significant differences.

TABLE 7.8. First sensor – highest values

	Healthy	**ILL 2**	**ILL 2.5**	**ILL 3**
Healthy		**<0.001**	**< 0.001**	0.3231
ILL 2	**<0.001**		0.3501	0.8898
ILL 2.5	**< 0.001**	0.3501		0.3321
ILL 3	0.3231	0.8898	0.3321	

TABLE 7.9. First sensor – average values

	Healthy	**ILL 2**	**ILL 2.5**	**ILL 3**
Healthy		**<0.001**	**0.0024**	**0.0079**
ILL 2	**<0.001**		0.9854	0.6487
ILL 2.5	**0.0024**	0.9854		0.8174
ILL 3	**0.0079**	0.6487	0.8174	

TABLE 7.10. Second sensor – highest values

	Healthy	**ILL 2**	**ILL 2.5**	**ILL 3**
Healthy		**0.0107**	0.2564	0.8217
ILL 2	**0.0107**		0.9308	0.0761
ILL 2.5	0.2564	0.9308		0.2378
ILL 3	0.8217	0.0761	0.2378	

TABLE 7.11. Second sensor – average values

	Healthy	**ILL 2**	**ILL 2.5**	**ILL 3**
Healthy		0.3727	**0.0069**	0.1762
ILL 2	0.3727		0.2476	0.6350
ILL 2.5	**0.0069**	0.2476		1.0000
ILL 3	0.1762	0.6350	1.0000	

TABLE 7.12. Fifth sensor – highest values

	Healthy	**ILL 2**	**ILL 2.5**	**ILL 3**
Healthy		0.8055	**< 0.001**	**<0.001**
ILL 2	0.8055		**< 0.001**	**0.0014**
ILL 2.5	**< 0.001**	**< 0.001**		0.9994
ILL 3	**<0.002**	**0.0014**	0.9994	

TABLE 7.13. Sixth sensor – highest values

	Healthy	**ILL 2**	**ILL 2.5**	**ILL 3**
Healthy		**0.0056**	0.0831	0.3266
ILL 2	**0.0056**		0.9950	1.0000
ILL 2.5	0.0831	0.9950		0.9983
ILL 3	0.3266	1.0000	0.9983	

From the tables 7.8–7.13 it can be read that the largest number of groups between which there are statistically significant differences is for the mean values for sensor 1 and the largest values for sensor 5. In the first case, there are healthy people and people with all stages of the condition, while in the second case, there are healthy people and people with Parkinson's stage 2.5 and 3. On the other hand, in the largest values for sensor 1 statistically significant differences are for the groups "healthy", "ILL 2" and "ILL 2.5". For sensor 2 and 6, there are two groups of people – healthy people and people with stage 2 disease.

Low values of p-level could indicate that the sample has been overpowered. This paper is based on research conducted by Jeffrey Hausdorff and made available at physionet.org. Due to the use of a pre-prepared database, it is not possible to make a deeper analysis to show what exactly influenced the results. The obtained values should therefore be treated with caution [16, 17].

7.4. Discussion

Previous work analyzing Parkinsonian gait has mostly focused on stride length and width, locomotor speed and changes in joint angle. The data were taken from the free database physionet.org, conducted by Jeffrey Hausdorff. It has been shown that a reduction in stride speed is associated with the condition in question [11, 12]. This can also be observed by comparing figures 7.2–7.4, as the duration of the support phase increases and is approximately 1.2 sec, 1.4 sec and 1.7 sec respectively. It was noted in [14] that a reduction in walking speed affects the control of body stability, resulting in a more confident gait and a lower risk of falling. A slightly prolonged double support phase can also be observed in Parkinson's disease patients. This allows for the control of the center of gravity and, consequently, protection from falling. This has also been noted by Sharifmoradi and Farahpour [14].

The vertical component of the ground reaction force changes with the stage of the disease. In case of healthy people, two local maxima and one minimum can be clearly identified, representing characteristic phases of the gait cycle. In case of sick people, on the other hand, the course of the component is completely different. In case of a patient with the second stage of the disease, only one local maximum and one minimum are visible. In the patient with the highest analyzed degree of the disease, it is very difficult to distinguish the characteristic phases of the gait cycle and the course is more benign [13].

Analyzing the vertical component of the ground reaction force for the right and left foot of affected individuals, it can be concluded that the course of values is very similar for both feet. It cannot be shown exactly which limb is affected. This may be due to the fact that in the analyzed stages of the disease the symptoms occur bilaterally, as also noted in [14].

To date, there have been very few studies analyzing the vertical component of the ground reaction force in the gait of people with Parkinson's disease. In [4], Kistler platforms and the Vicon Motion Analysis System were used for the study. All people were in the age range of 50–70 years. The analysis of the results allowed for the observation that the vertical component of the ground reaction force in people with Parkinson's disease has a similar course for both feet. This observation is confirmed by the results presented in this paper despite the use of a different data acquisition method and slightly different test thickness.

Analyzing the pressure forces on the individual sensors of Ulfraflex insoles, reduced values at the heel were noted for Parkinson's disease patients in the third stage of the disease. A similar relationship was obtained in [15]. There, it was found that lower loads transmitted through the heel were associated with the production of a characteristic gait cycle by people with Parkinson's. It was noted that a weaker heel strike on the ground affects the uncertainty and instability of locomotion. When analyzing forefoot forces, it is noted that they are increased in people with Parkinson's disease. This was also noted in the study [15]. During heel strike while walking, the toes of the forefoot are still on the ground. It was concluded that this could be explained by a compensatory mechanism of possible loss of balance. It is worth noting that in [15] both insoles from a different manufacturer (Pedar NOVEL) and the division of the foot into ten sub-areas were used when analyzing the pressure forces.

7.5. Conclusions

Gait testing and analysis of the vertical component of the ground reaction force based on a free database from phisionet.org conducted by Jeffrey Hausdorff, may allow for an earlier detection of the disease and implementation of therapy, which may alleviate and slow down the onset of symptoms.

Patients with Parkinson's disease try to protect themselves from falling by slowing down their locomotion and by being more assertive in their stopping. As the disease progresses, the characteristic phases of gait disappear and the way the feet bounce off the ground changes. In a person with the second stage of the disease, according to the Hoehn-Yahr scale, the way of kicking changes, and the second maximum of the vertical component of the ground reaction force disappears. In a person with the third stage of the disease, however, it is difficult to determine the characteristic phases of the gait cycle, because the limb almost immediately contacts the ground with the whole foot, resulting in a characteristic "treadmill", which allows for the control of the center of gravity and maintaining balance.

Patients with Parkinson's disease develop their own characteristic gait pattern to protect themselves from falling. This is also influenced by an increase in the double-support phase and a prolongation of one step time.

When analyzing the pressure forces, there appear increased values in the forefoot and decreased values in the heel. In the case of the heel, this is due to a hesitant step and a different foot position during the initial contact with the ground. On the other hand, changes in the forefoot are influenced by a longer contact with the ground and thus increased values.

References

1. https://physionet.org/content/gaitpdb/1.0.0/ [accessed: 16.11.2020].
2. Morris M., Huxham F., McGinley J., Dodd K., Iansek R.: "The biomechanics and motor control of gait in Parkinson Disease", Clinical Biomechanics, 16, 459–470, 2001.
3. Lim S., Fox S.H., Lang A.E.: "Overview of the extranigral aspects of Parkinson Disease", Archives of Neurology, 66 (2) 167–172, 2009.
4. Savica R., Rocca W.A., Ahlskog J.E.: "When does Parkinson Disease start?", Archives of Neurology, 67(7), 798–801, 2010.
5. Dobosiewicz A. M., Kwiatkowski Sz., Miętkowska P., Marszałek A. Wendland S.: "Motor skills problems in Parkinson Disease", Collegium Medicum, 7(6), 244–264, 2017.
6. Perry J.: GAIT ANALYISIS Normal and Pathological Function, SLACK Incorporated, United States of America, 1992.
7. Skalska-Dudlińska B., Witkiewicz B., Ptasznik I.: "Freezing of gait rehabilitation in Parkinson's disease", Current Neurology, 14 (2), 140–148, 2014.
8. Zeng W., Yuan Ch., Wang Q., Liu F., Wang Y.: "Classification of gait patterns between patients with Parkinson's disease and healthy controls using phase space reconstruction (PSR), empirical mode decomposition (EMD) and neural networks", Neural Networks, 111, 64–76, 2019.
9. https://www.mp.pl [accessed: 09.07.2021].
10. Whittle M.W., Richards J., Levine D.: "Whittle Analiza chodu", Elsevier Urban & Partner, Wrocław, 2014.
11. Nardello F., Bertoli E., Bombieri F., Bertucco M., Monte A.: "The effect of a secondary task on kinematics during turning in Parkinson's disease with mild to moderate impairment", Symmetry (Basel), 12(8), 2020.
12. Soufa O., Nieuwboer A., Desloovere K., Willems A., Chavert F., Jonkers I.: "Quantitative gait analysis in Parkinson's disease: Comparison with a healthy Control group" Archives of Physical Medicine Rehabilitation, 86(4), 1007–1013, 2005.
13. Pistacchi M, Gioulis M., Sanson F.: "Gait analysis and clinical correlations in early Parkinson's disease" Functional Neurology, 32(1), 28–34, 2017.
14. Sharifmoradi K., Farahpour N.: "An assessment of gait spatiotemporal and GRF of Parkinson patients", Health and Rehabilitation, Vol. 1, 29–34, 2016.
15. Nardello F., Bertoli E., Bombieri F., Bertucco M., Monte A.: "The effect of a secondary task on kinematics during turning in Parkinson's disease with mild to moderate impairment", Symmetry (Basel), 12(8), 2020.
16. Andrade C.: "The P value and statistical significance: misunderstandings, explanations, challenges, and alternatives", Indian journal of psychological medicine, 41(3), 210–215, 2019.
17. Zhu W.: "p< 0.05,< 0.01,< 0.001,< 0.0001,< 0.00001,< 0.000001, or< 0.0000001….", Journal of sport and health science, 5(1), 77, 2016.

DOI: 10.24427/978-83-67185-49-3_8

Chapter 8

Comparative analysis of computer typing of healthy people and people with Parkinson's disease

Anna Anchimowicz, Marcin Derlatka

Bialystok University of Technology, Institute of Biomedical Engineering,
e-mail: ania.anchimowicz@wp.pl, m.derlatka@pb.edu.pl

Abstract: Nowadays, methods are being searched in order to detect Parkinson's disease at an early stage. This would aid faster diagnosis and a more objective assessment of the severity of the condition. This paper focuses on examining whether typing differs between healthy people and people with varying degrees of Parkinson's disease. The analysis used publicly available data from the physionet.org database [1], which contains results from 227 people over several experimental days. Data from only 108 subjects were used, with a maximum of 3 days of use of the Tappy application by one participant with at least 200 clicks considered. Computer keyboard typing parameters, such as button hold time, latency and flight time were taken into account. Statistical analysis of the data obtained using non-parametric Kruskal-Wallis tests showed that there were no statistically significant differences ($\alpha < 0.05$) between the study groups (healthy people and people with varying degrees of Parkinson's disease). However, the reason for this may be the small study group and the lack of consideration of medications taken when matching to groups.

Keywords: Parkinson's disease, keyboard typing, statistical analysis

8.1. Introduction

Parkinson's disease is a degenerative and progressive disorder of the nervous system. It occurs most commonly in people aged between 55 and 65 years, and the average age of onset is 58 years. Men are more commonly affected [2]. The preclinical period of the disease, in which typical symptoms, such as rest tremor or motor retardation are not yet apparent, may last up to 10 years [3]. During this period, autonomic symptoms (cardiovascular, gastrointestinal, genitourinary, or thermoregulatory disorders) are the most common [4]. By the time the visible symptoms of Parkinson's disease are present, the atrophy of nerve cells in the black matter of the brain is already significant at around 60%. The concentration of dopamine (an organic compound that plays a key role in the transmission of nerve impulses, responsible for the execution of human will-dependent movement) in the striatum is around 20–40% [5]. Parkinson's patients have difficulty moving, dressing, eating or brushing their teeth. Alternate movements and precision activities, such as buttoning buttons and screwing in screws are difficult. They may have problems with washing themselves and combing their hair. In addition, their handwriting becomes less legible and letters smaller. Activity itself is slower and the hand on the affected side is contracted [6].

An early detection of Parkinson's disease is important, as dopamine deficiencies can be corrected with medication. This extends the patient's period of function and independence. Appropriately implemented treatment allows patients to maintain their current job without having to give up their career or suffer a significant decrease in quality of life [7].

Various scales are used to determine the stage of the disease, e.g. Hoehn-Yahr scale and UPDRS. They are not very objective, because they are based on the feelings of the examiner and the patient, and thus they do not allow for an objective assessment of the patient's condition [6, 8].

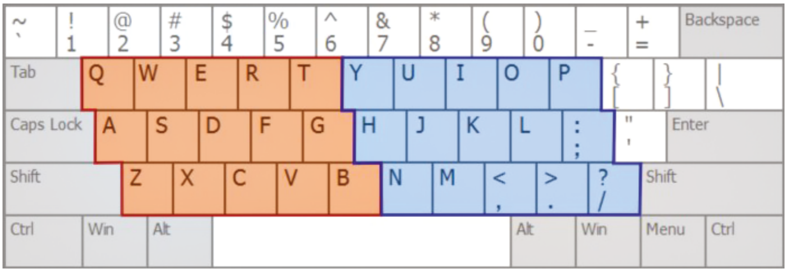
Currently, there is ongoing research into methods for the early detection of Parkinson's disease. Among others, gait characteristics [9], handwriting [10] or computer keyboarding [11, 12, 13] are used for this purpose. Skin tests [14] and retinal thickness [15] are also performed. Parkinson's disease may also be indicated by disorders of the REM sleep phase [16].

The aim of this article is to investigate whether there is a relationship between the way people type on a computer keyboard and the severity of Parkinson's disease. This would allow a more objective diagnosis to be made. No information on the statistical analysis of the data obtained in this study alone was found in the available literature.

8.2. Materials

Publicly available data from the physionet.org database [1] were used to perform the analysis. They concerned the typing characteristics of healthy people and people with Parkinson's disease at different stages of its progression [1, 11]. Data were collected using the Tappy application, which was downloaded by people willing to take part in the experiment. The whole process was preceded by the registration of these people on a platform where information about the study was presented. People aged 50–80 could take part in the experiment. When registering, people indicated, among other things, the stage of the disease, the side affected by the disease, whether they had tremors and whether they were taking medication (if so, which). To ensure anonymity, a random ten-character code was automatically assigned. The Tappy application allowed for the time of pressing the corresponding keys to be read out. Respondents used their computers according to their needs – the application recorded movements while searching websites, creating text documents, or writing emails [1, 11].

Figures 8.1 and 8.2 show the assignment of keys to the hand they were pressed with and time dependencies of individual parameters. Subjects were unaware of the key assignment to a particular hand and used the keyboard the way they were comfortable with, which may have affected the recorded results [1, 11].



🟧 Left hand keys captured 🟦 Right hand keys captured

All other keys were excluded from capture, except the Spacebar.

FIGURE 8.1. Assignment of keys to the hand the key was pressed with [11]

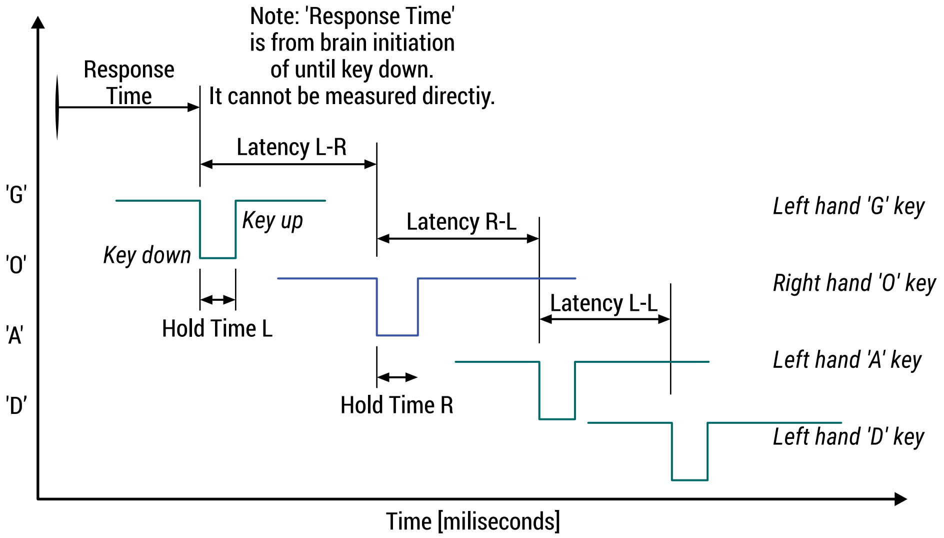


FIGURE 8.2. Time dependencies of parameters describing the way of typing on a computer keyboard [11]

The database contains data from 227 individuals. The present study used data collected from only 108 individuals. This limitation was due to the ability to create groups clearly according to the severity of a participant's disease and to assign the results to the affected site accordingly. In addition, the analysis used data from only the first three days of app use, in which the subjects made at least 200 clicks per hand.

This decision was dictated by minimizing the risk of the results being dominated by people who had been using the app for longer. A higher number of clicks allowed for a more accurate representation of how people typed on a computer keyboard.

The group breakdown included people who clearly indicated the severity of the disease and the side affected during registration. Factors, such as gender or prescribed medication, were not taken into account. The detailed characteristics of the individuals included in the analysis by group are presented in Table 8.1.

TABLE 8.1. Number of participants and sheets (one sheet consisted of the results obtained by one participant on one survey day) used for the analysis

	Healthy subjects	Subjects without tremors		Subjects with tremors		Total
		Mild degree	Medium degree	Mild degree	Medium degree	
Number of participants	38	10	13	27	20	108
Number of sheets	70	24	22	57	43	216

8.3. Method

Statistical analysis was performed to compare individual parameters according to the severity of Parkinson's disease, divided into groups with and without tremors. First, it was examined whether the distributions obtained from the mean values of the studied parameters obtained by the individuals were normal using the Shapiro-Wilk test. It was assumed that a p value of less than 0.05 obtained in the test indicates a distribution deviating from normal. The results showed that most of the distributions were deviating from normal, so further analyses were carried out using non-parametric tests. The Kruskal-Wallis test was used to examine whether there were statistically significant differences at a significance level of $\alpha<0.05$.

8.4. Analysis of results

The results are shown in Table 8.2. It can be seen that statistically significant differences were found only for the time the key was held down with the hand affected by the disease.

In order to check between which groups there were differences, a multiple comparison test was performed. Its results are shown in Table 8.3. However, this test indicated that there were no statistically significant differences between the groups examined.

TABLE 8.2. Results of the Kruskal-Wallis test for group comparisons

	Results of the Kruskal-Wallis test	
	Non-affected side	**Affected side**
Hold time	0.2516	**0.0091**
Latency	0.4384	0.5192
Flight time	0.205	0.2212

TABLE 8.3. Results of multiple comparisons test for button hold time for the affected hand (Zdr – healthy person, LBD – person with mild Parkinson's disease without tremors, ŚBD – person with medium Parkinson's disease without tremors, LD – person with mild Parkinson's disease with tremors, CD – person with medium Parkinson's disease with tremors)

	Zdr	LBD	ŚBD	LD	CD
Zdr		0.1785	1.0000	0.1662	0.1061
LBD	0.1785		0.2880	1.0000	1.0000
ŚBD	1.0000	0.2880		0.3620	0.3116
LD	0.1662	1.0000	0.3620		1.0000
CD	0.1061	1.0000	0.3116	1.0000	

The degree of dependence between the groups is illustrated in the figures below (Fig. 8.3–8.5).

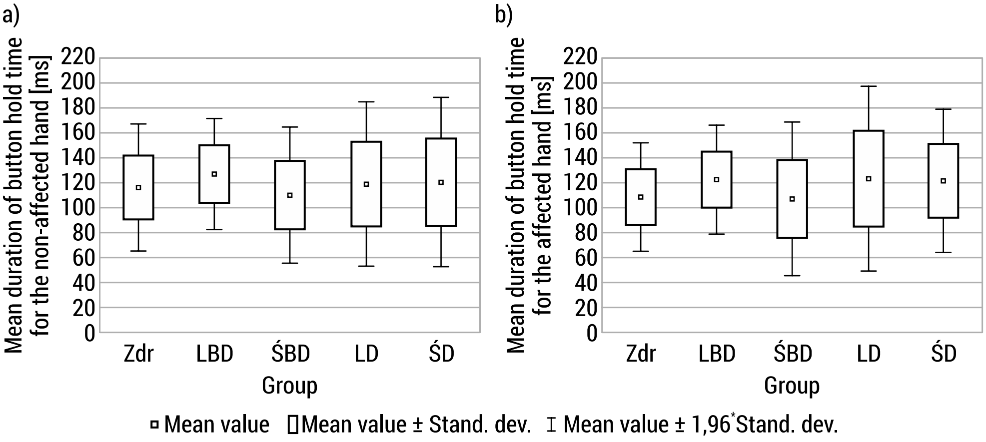


FIGURE 8.3. Plots of the mean duration of button hold time for the a) non-affected hand b) affected hand (in the case of healthy people, the left and the right hand respectively)

It can be seen that the plots (Fig. 8.3) confirm the results of the Kruskal-Wallis test – the frames overlap. Some differences occur in the case of the mean value, but these values are also similar. For the hold time of both hands the mean value is about 120 ms. Mild and medium condition people with tremors have a higher variability of results than healthy and without tremors people.

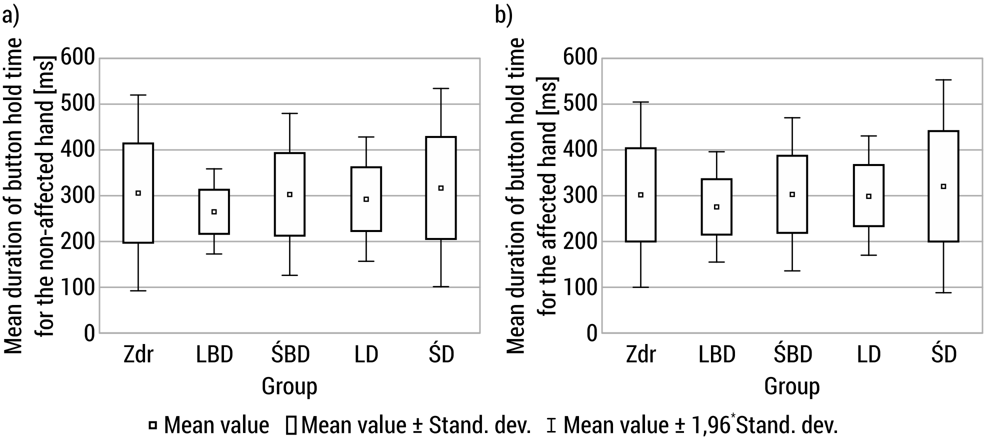


FIGURE 8.4. Plots of the mean duration of latency for the a) non-affected hand b) affected hand (in the case of healthy people, the left and the right hand respectively)

For the latency time, again the correctness of the Kruskal-Wallis test results can be observed as the frames overlap. The average latency was about 300 ms for both hands. The smallest variability of the results can be observed for people in the mild condition without tremors. The highest variability is observed for people in the medium condition with tremors.

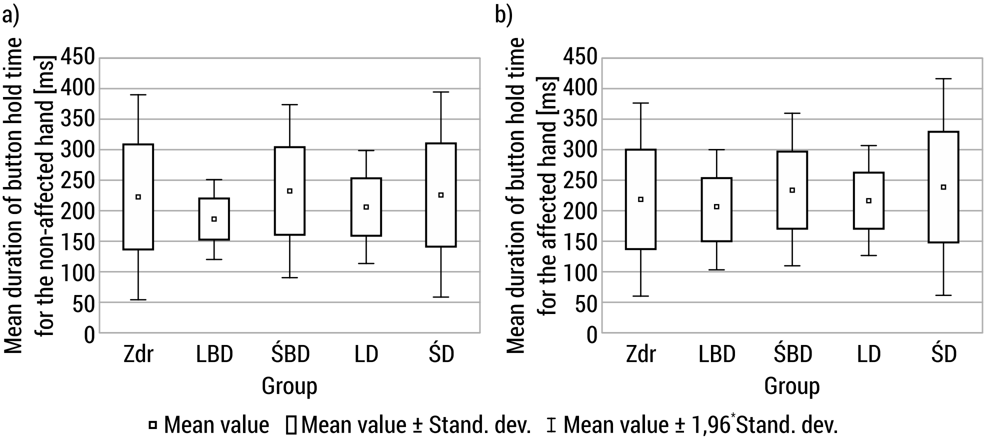


FIGURE 8.5. Plots of the mean duration of flight time for the a) non-affected hand b) affected hand (in the case of healthy people, the left and the right hand respectively)

A similar situation to that for latency occurs for the flight time, where the mean value is around 200 ms, and in addition, a lower variability of the results can also be observed for mild condition people with tremors.

8.5. Discussion

Often, data from healthy people and people with Parkinson's disease at different stages of progression are used to create classifiers that automatically detect Parkinson's disease and its severity.

This analysis showed that there were no significant differences between the results obtained by healthy people and people with different stages of Parkinson's disease when typing on a computer keyboard. This is in contrast to the results obtained in [11], which analysed data obtained from the same database as in the present study. However, it only used results from healthy individuals and people with mild Parkinson's disease who were not taking medication to reduce symptoms of the condition. Differences may be due to different criteria for selecting the study subgroup – in this study the effect of medication on the results obtained by individual participants was not taken into account.

However, the correctness of the analyses in the present study is indicated by studies performed in [12]. Their main aim was to check the possibility of differentiating people with tremors. Healthy people and people with medium Parkinson's disease were taken into account with and without tremors. Again, the database used for the analyses in this paper was used, but expanded to include new consecutive study participants not available at physionet.org. The results showed that the different groups of people could be distinguished. However, it was noted that those taking levodopa-containing medication achieve hand tremor much less than those not taking such medication, comparable to that of healthy people. For this reason, an analysis of the fit of individuals to each group was also performed using classifiers without taking into account those taking levodopa. This analysis showed a better fit of the individuals to the different groups. This confirms the significant influence of the medication taken on the course of the disease.

In contrast to the results obtained in the present study, another study [13] also indicated that it was possible to distinguish between healthy individuals and individuals at different stages of Parkinson's disease. In this case, however, the location of the computer mouse pointer during the click and the interaction with the computer keyboard during the process of searching for keywords in an Internet search engine were used. This study took into account different parameters than those described so far – the time elapsed between the presentation of results to the user and the first interaction event with the website; the average time elapsed between each two mouse positions (mean dwell time); one of the automatically learned functions. The results showed that it was possible to distinguish the actual diagnostic status of the subjects on this basis. In this case, the effect of medication was not taken into account due to the lack of such information about the users, but different parameters of computer use were investigated than in the case of this article and the previously described work, and in addition, the study groups were larger than in the present analysis.

Summing up, the results of the study were heavily influenced by grouping without taking into account the medication taken to reduce the symptoms of Parkinson's disease by individual participants, but taking this criterion into account would have reduced the size of the groups.

In addition, individuals used the computer unsupervised and on an as-needed basis. They were also unaware of the assignment of specific buttons to the left or right hand, which may also have influenced the results – the possibility that a button assigned to the right hand could be pressed by the left hand and vice versa was not eliminated. There was also no information about the dominant hand.

8.6. Conclusions

In the paper the statistical analysis of computer typing of healthy subjects and subjects with Parkinson's disease has been made. The analysis shows that there are no statistically significant differences between the results for each group for such parameters as button hold time, latency and flight time when typing on a computer keyboard. This makes it difficult to recognize whether the person studied is healthy or has Parkinson's disease and if so, to what degree. However, the reason for this may be a small study group and the lack of consideration for medications taken when matching to groups.

Future research should take into account larger group sizes and eliminate the influence of medications on results, for example by selecting groups of people taking specific medications or conducting studies with people who do not take such medications.

References

1. https://physionet.org/content/tappy/1.0.0/ [accessed: 16.11.2020r].
2. Reuter I.: Choroba Parkinsona, Elsevier Essentials, [eds.] Budrewicz S., Edra Urban & Partner, Wrocław, 2019.
3. Gaig C., Tolosa E.: "When does Parkinson's disease begin?", Movement Disorders, 24(S2), S656–S664, 2009.
4. Palma J.A., Kaufmann H.: "Autonomic disorders predicting Parkinson's disease", Parkinsonism & related disorders, 20, 94–98, 2014.
5. Krygowska-Wajs A.: "Przedkliniczny i wczesny okres choroby Parkinsona-diagnostyka i możliwości leczenia neuroprotekcyjnego", Polski Przegląd Neurologiczny, 2(4), 177–182, 2006.
6. Parkinson's disease diagnosis scale MDS – UPDRS, 1.07.2008 (https://www.movementdisorders.org/MDS-Files1/PDFs/Rating-Scales/MDS-UPDRS_Polish_Official_Translation_FINAL_MargaretSmith.pdf).
7. Prashanth R., Roy S.D.: "Early detection of Parkinson's disease through patient questionnaire and predictive modelling", International Journal of Medical Informatics, 119, 75–87, 2018.

8. Opara J., Małecki A., Małecka E., Socha T.: "Motor assessment in Parkinson's disease", Annals of Agricultural and Environmental Medicine, 24(3), 411–415, 2017.
9. Pistacchi M., Gioulis M., Sanson F.: "Gait analysis and clinical correlations in early Parkinson's Disease", Funct Neurol., 32(1), 28–34, 2017.
10. Zham P., Kumar D.K., Dabnichki P., Poosapadi Arjunan S., Raghav S.: "Distinguishing different stages of Parkinson's disease using composite index of speed and pen-pressure of sketching a spiral", Frontiers in Neurology, 8, 435, 2017.
11. Adams W. R.: "High-accuracy detection of early Parkinson's Disease using multiple characteristics of finger movement while typing", PloS one, 12(11), e0188226, 2017.
12. Adams W. R.: "The detection of hand tremor through the characteristics of finger movement while typing", bioRxiv, 385286, 2018.
13. Youngmann B., Allerhand L., Paltiel O., Yom-Tov E., Arkadir D.: "A machine learning algorithm successfully screens for Parkinson's in web users", Annals of Clinical and Translational Neurology, 6(12), 2503–2509, 2019.
14. Manne S., Kondru N., Jin H., Serrano G.E., Anantharam V., Kanthasamy A., Adler C.H., Beach T.G. Kanthasamy A.G.: "Blinded RT-QuIC Analysis of α-Synuclein biomarker in skin tissue from Parkinson's disease patients", Movement Disorders, 35(12), 2230–2239, 2020.
15. Ahn J., Lee J.Y., Kim T.W., Yoon E.J., Oh S., Kim Y.K., Kim J.M., Woo S.J., Kim K.W., Jeon B.: "Retinal thinning associates with nigral dopaminergic loss in de novo Parkinson disease", Neurology, 91(11), e1003–e1012, 2018.
16. Stiasny-Kolster K., Mayer G., Schäfer S., Möller J.C., Heinzel-Gutenbrunner M., Oertel W.H.: "The REM sleep behavior disorder screening questionnaire – a new diagnostic instrument", Movement Disorders, 22(16), 2386–2393, 2007.

DOI: 10.24427/978-83-67185-49-3_9

Chapter 9
Multifractal detrended fluctuation analysis of heart rate variability

Ewelina Bębas, Anna Kasperczuk*, Edward Oczeretko*,*
*Artur Bossowski**, Marta Borowska**

**Bialystok University of Technology, Institute of Biomedical Engineering,*
e-mail: e.bebas@doktoranci.pb.edu.pl, a.kasperczuk@pb.edu.pl,
e.oczeretko@pb.edu.pl, m.borowska@pb.edu.pl
***Medical University of Bialystok, Department of Pediatrics Endocrinology and Diabetology,*
e-mail: artur.bossowski@udsk.pl

Abstract: The analysis of heart rate variability can be used to predict cardiovascular state in diabetic children. Studies on HRV are mainly based on RR intervals of the ECG signal and include both linear indices and nonlinear characteristics. The current study aims at applying Multifractal Detrended Fluctuation Analysis to investigate signals of the heart rate variability (HRV) in 50 healthy children (group K), 27 patients suffering from diabetes type 1 without subsequent vascular complications (group A) and 35 patients suffering from diabetes type 1 with subsequent vascular complications (hypertension) (group B). Multifractal features were extracted to quantitatively compare signal complexity between the healthy group and the group of children with diabetes and diabetes with complications. The MFDFA analysis showed the multi-fractional nature of the signals in all groups. The characteristic MFDFA parameters were determined and subjected to statistical analysis. It showed differences in parameters h(1) and h(2) between the control group K and group A. This gives rise to the development of further research in an attempt to distinguish between these groups.

Keywords: fractal dimension, multifractal, multifractal detrended fluctuation analysis, heart rate variability, ECG, diabetes, children

9.1. Introduction

Multifractal detrended fluctuation analysis (MFDFA) is a widely used method to characterize the variability and uncertainty of empirical time series. It has its origins in detrended fluctuation analysis (DFA), which was developed by Peng et al. [1, 2] and later extended by Kandelhardt et al. to the MFDFA [3]. Multifractal analysis has found application in many fields. Its use can be observed in the time series analysis of sleep EEG data [4], heart rate [5], blood pressure [6], Parkinson's disease research [7], seismic activity [8], temperature variability [9], price electricity [10], epidemiology [11], finance [12] and many others.

Many studies indicate the effective use of multifractal analysis in cardiological diseases. MFDFA was used, inter alia, to analyze the RR series of patients with left ventricular systolic dysfunction, comparing them to a control group of healthy people [13]. Three scaling modifications were used in this study. This method was also used to assess the monofractal and multifractal properties of R-R intervals for the healthy group and with various changes in cardiac diseases [14]. The MFDFA analysis allowed for the determination of static differences in the group with pathological changes and distinguished congestive heart failure from others (atrial fibrillation and sudden death syndrome).

In their work Molina et al. [15] showed differences in the analysis of multifractal electrocardiogram (ECG) spectra of healthy people at rest and during exercise on a treadmill at different speeds. The multifractal spectra behaved differently at rest and in motion. In [16], a new method of asymmetric multiscale multifractal analysis was used for six groups of patients – healthy persons, with aortic valve stenosis, with hypertrophic cardiomyopathy, with atrial fibrillation, patients with coronary artery disease and with congestive heart failure. The authors indicate the effectiveness of the proposed method and the possibility of its development in medical screening.

MFDFA and its modifications are widely used in the analysis of heart rate variability as well as in other signals. This prompted the authors of this publication to undertake the analysis of heart rate variability in three groups of patients with 24-hour ECGs.

The aim of the study is to investigate the multifractality of ECG signals in children with type 1 diabetes mellitus with and without later vascular complications and healthy children, divided into the night (rest) and day (activity) part, and to determine the characteristic parameters of MFDFA in individual groups.

9.2. Materials and methods

In the article, the results of the 24-hour Holter ECG monitoring of children were used in accordance with the recommendations of the Task Force of the European Society of Cardiology. The results of 120 studies came from 3 groups of children. Group A (27 cases) included the results of studies in children with type 1 diabetes without subsequent vascular complications. The mean age in this group is 15 ± 2.2. Group B (35 cases) is the results of children with type 1 diabetes mellitus with subsequent vascular complications (hypertension) with a mean age of 15 ± 2.9. The third group K (50 cases) is a control group of healthy children, the results of which were matched with age and gender to the groups of sick children. The results were divided into a day and night part. The diagram in Figure 9.1 shows the analyzed example of the data exposure.

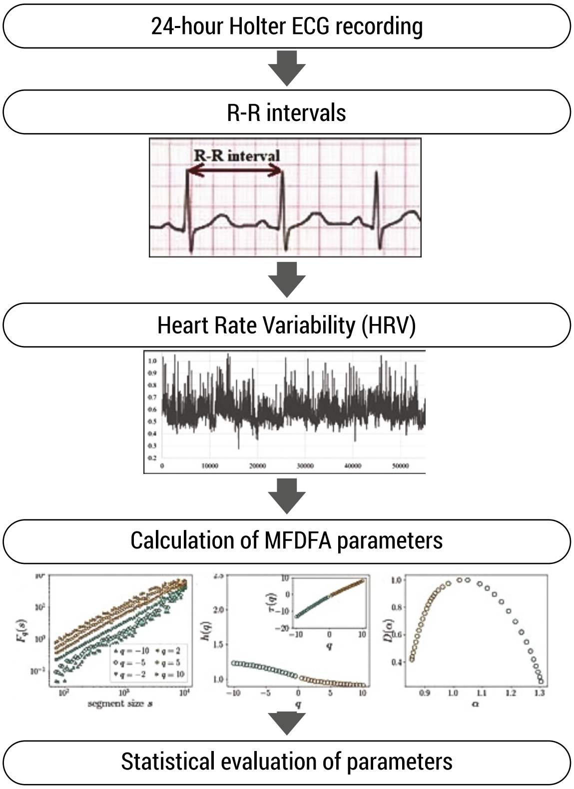


FIGURE 9.1. Log-log $F_q(s)$ plots as a function of segment size s for 6 exemplary q values

9.2.1. Multifractal detrended fluctuation analysis

The presented research results were carried out on the basis of the MFDFA algorithm described in the article [17]. The algorithm is based on a few steps presented below. It is described in detail by Kantelhardt et al. in the publication [48] in the following manner:

Step 1. Determine the profile of a time series of finite length N:

$$Y_i = \sum_{k=1}^{i} X_k - \bar{X}, i = 1,2,\ldots,N, \tag{9.1}$$

where: X_k – time series, $\bar{X}$ – mean of data

Step 2. Dividing the Y_i profile into smaller, non-overlapping segments s:

$$N_s = int \frac{N}{s}. \tag{9.2}$$

Due to the fact that the length of the series *N* is not always a multiple of the *s* scale, the endpoints of the data are discarded. Next, the operation is repeated, discarding the starting points, which results in 2Ns segments.

Step 3. Fitting the y_v polynomial of m for each of the obtained segments and determining the variance:

$$F(v,s)=\frac{1}{s}\sum_{i=1}^{s}\left[Y_{(v-1)s+i}-y_{((v-1)s+i)}\right]^{2}, \tag{9.3}$$

for $v = 1, 2,\ldots, N_s$ where $y_{((v-1)s+i)}$ is a polynomial fit of a segment of length *s* using the least squares method. Depending on the chosen degree *m* of the polynomial, the notations (MF) DFA1, (MF) DFA2, ..., (MF) DFAm are used.

Step 4. Define the q-order fluctuation function by averaging the segments of size *s*.

$$F_q(s)=\left\{\frac{1}{N}\sum_{v=1}^{N_s}\left[F(v,s)\right]^{\frac{q}{2}}\right\}^{\frac{1}{q}} \tag{9.4}$$

The above formula shows that the fluctuation function $F_q(s)$ depends on the indexing variable *q*, which can take any value, and on the size of the segment *s*.

After the jitter function is computed, its behavior is determined by analyzing the logarithmic dependence on the segment size *s* plot for different *q* values. An example of such a plot is shown in Fig. 9.2.

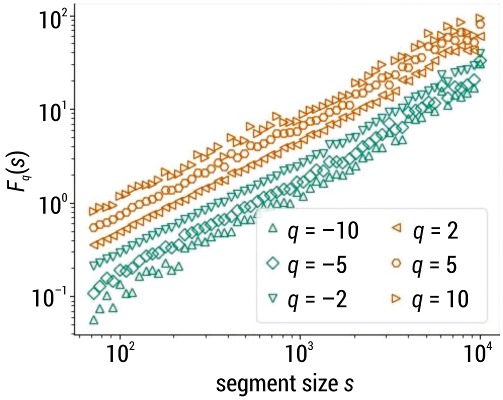


FIGURE 9.2. Log-log $F_q(s)$ plots as a function of segment size *s* for 6 exemplary *q* values

On the basis of the above dependence, it can be determined whether the data is multi-fractional. If the curves for different *q* values are parallel, it indicates monofractality of the process, and when they are not parallel – multifractality. Another indicator pointing to mono- or multifractality is the behavior of the plot of the dependence of the scaling exponent *h*(*q*) as a function of *q*. If it is parallel to the *x* axis, it means monofractality of the process, if not – multifractality.

If there is a correlation of the x_i series with the power law, $F_q(s)$ increases for large values of s according to the formula:

$$F_q(s) \sim s^{h(q)}, \tag{9.5}$$

where $h(q)$ is the generalized Hurst exponent depending on q. It is determined by finding the slope of the $F_q(s)$ curve on a logarithmic plot, as shown in Figure 9.2.

In the case of monofractal data, the Hurst exponent $h(q) = H$ does not depend on q. Then the generalized Hurst exponent is also the Hurst index H. While the data is multifractal, the dependence of $h(q)$ on q is examined the multifractal scaling exponent $t(q)$. The multifractal scaling exponent depends on the generalized Hurst exponent $h(q)$ as shown in the formula:

$$\tau(q) = qh(q) - 1. \tag{9.6}$$

In a similar way, the specter of the singularity $D(a)$ can be constructed as the Legendre transform. While $t(a)$ is smooth enough, the alpha singularity strength is determined by the formula:

$$\propto = \tau'(q) = h(q) + qh'(q). \tag{9.7}$$

On the basis of the above formula, the spectrum of the singularity $D(a)$ can be constructed:

$$D(\propto) = q \propto - \tau(q). \tag{9.8}$$

The singularity strength α characterizes the dimension of a subset of the time series described by $D(a)$. The strength of the multifractality of the time series is indicated by the width of the alpha singularity strength.

9.2.2. Research description

The study involved the use of MFDFA for 3 groups of children. Using the methods proposed at https://github.com/LRydin/MFDFA, MFDFA parameters were calculated and graphs developed for each case. This allowed for the assessment of the multifractality of individual results in each group. Then, statistically determined parameters were compared.

Comparisons between the three independent groups were made using ANOVA analysis. Normality of distribution was checked using the Shapiro Wilk test. When there were differences between groups, POST-HOC tests were performed to determine homogeneous groups. A series of Student's t-tests for dependent samples or Wilcoxon tests were performed to examine differences between day and night runs. The analysis was performed using Statistica 13.3. Differences were considered significant at $p < 0.05$.

9.3. Results

All three groups showed a multifractal nature. Graphs of sample analysis results for each group are shown on Figure 9.3 and Figure 9.4.

The parameters that were subjected to further statistical analysis were the dependencies of the scaling exponent $h(q)$ on q, for $q = -10, -5, -2, 1, 2, 5, 10$, the range of the width of the singularity force α (minimum value – α_{min}, value maximum – α_{max}, width Δα), which indicates the strength of multifractality of the data, the values of the $D(\alpha)$ singularity spectrum for minimum and maximum α and their difference ΔD.

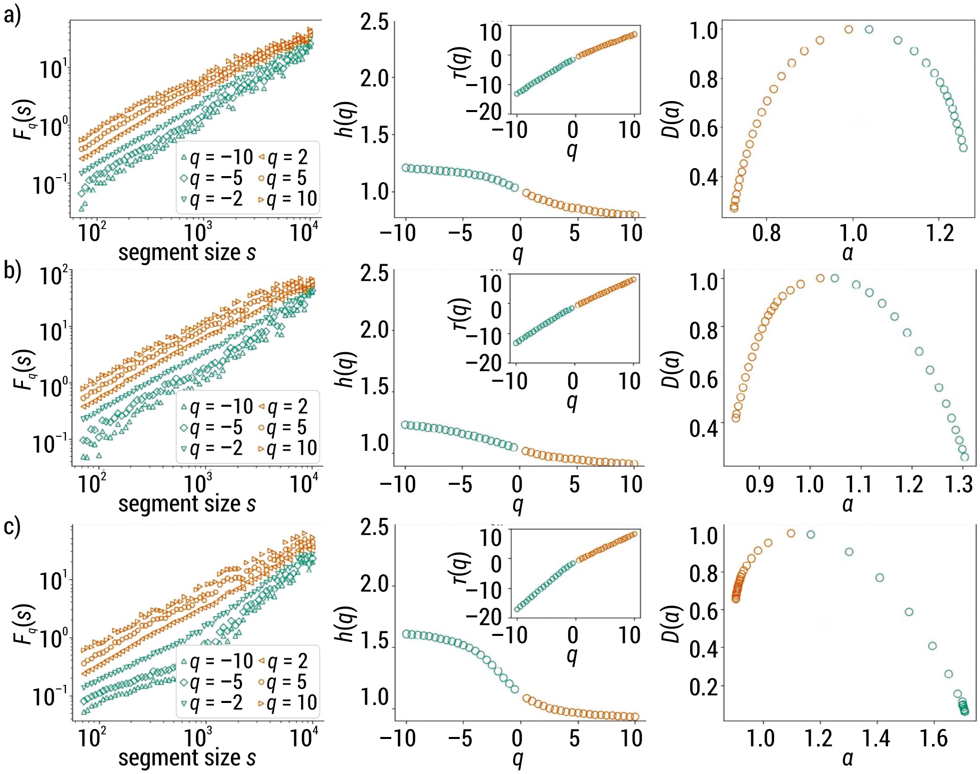


FIGURE 9.3. Graphs obtained during the MFDFA for a) group A, b) group B, c) group K by **day**. From the left there are log-log $F_q(s)$ plots as a function of the segment size s for 6 exemplary q values, the middle plot of the scaling exponent $h(q)$ as a function of q, and the inset displays the multifractal scaling exponent $\tau(q)$, on the right the spectrum the singularity $D(\alpha)$ over the singularity strength α

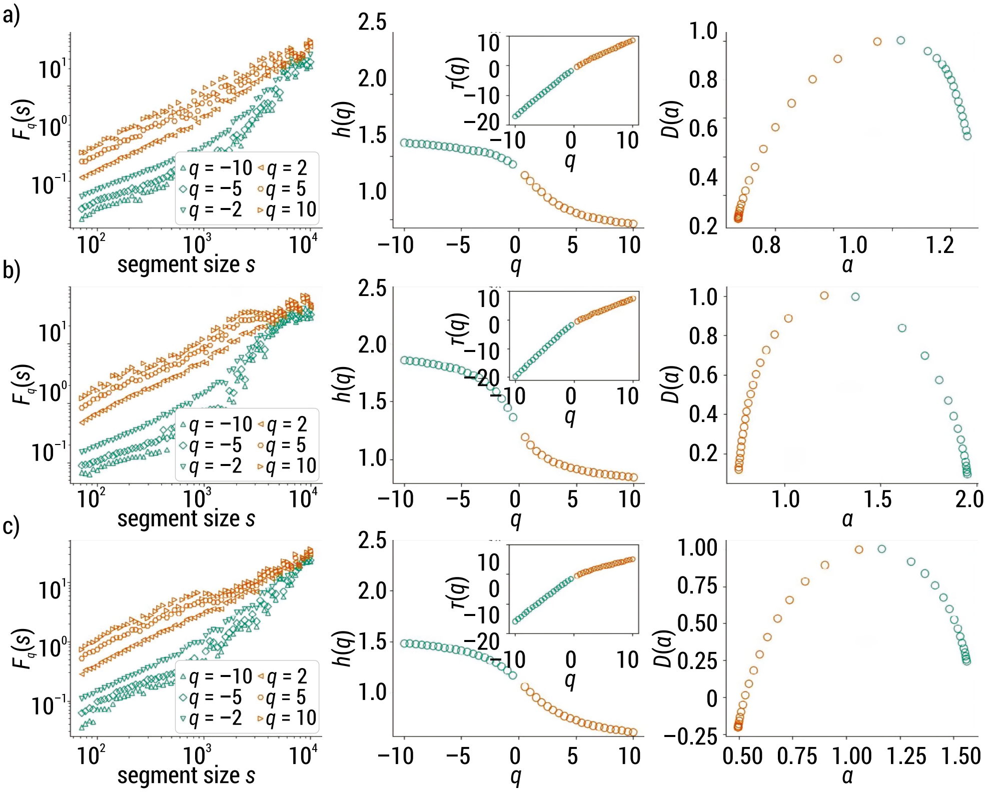


FIGURE 9.4. Graphs obtained during the MFDFA for a) group A, b) group B, c) group K at **night**. From the left there are log-log $F_q(s)$ plots as a function of the segment size s for 6 exemplary q values, the middle plot of the scaling exponent $h(q)$ as a function of q, and the inset displays the multifractal scaling exponent $\tau(q)$, on the right the spectrum the singularity $D(\alpha)$ over the singularity strength α

Figures 9.3 and 9.4 show in the first column the fluctuation function $F_q(s)$ for $q = -10, -5, -2, 2, 5, 10$ and for $s \in [70, 10000]$ in the analyzed groups measured during the day and night, respectively. The curves are not parallel, i.e. the time series of RR intervals are not monofractal. The second column shows the generalized Hurst exponent $h(q)$ as a function of q, and similarly the multifractal scaling exponent $\tau(q)$. The generalized Hurst exponent $h(q)$ as a function of q is not constant while the multifractal scaling exponent $\tau(q)$ is not linear, which clearly shows the multifractality of the time series. The third column shows the spectrum of singularity $D(\alpha)$ at singularity strength α. The strength of the singularity α extends over a wide range of values, indicating the different scales of the phenomenon under study.

There were no statistically significant differences between the groups in signals measured during the day ($p > 0.05$) (Tab. 9.1). Statistically significant differences were shown between groups for h values for $q = 1$ ($h(1)$) ($p = 0.006$) and for $q = 2$ (h(2)) ($p = 0.011$) (Tab. 9.2). An interesting observation is that all the measures that most effectively discriminated between the analyzed groups were on the positive side of the spectrum. An analysis of Table 9.2 shows that not a single measure using a negative q value was able to achieve $p < 0.05$. The best result was obtained for $q = 1$.

TABLE 9.1. Multifractal analysis parameters between groups – day

	group A			group B			group K			
parameter	***M***	***Me***	***SD***	***M***	***Me***	***SD***	***M***	***Me***	***SD***	***p***
$h(1)$	1.03	1.03	0.06	1.07	1.06	0.07	1.06	1.06	0.08	0.209
$h(-10)$	1.24	1.23	0.10	1.27	1.27	0.17	1.26	1.29	0.17	0.810
$h(-5)$	1.18	1.17	0.08	1.21	1.22	0.15	1.20	1.23	0.15	0.719
$h(-2)$	1.12	1.12	0.06	1.15	1.15	0.10	1.14	1.14	0.11	0.399
$h(2)$	1.01	0.99	0.07	1.04	1.03	0.07	1.02	1.02	0.07	0.305
$h(5)$	0.95	0.93	0.09	0.97	0.96	0.08	0.96	0.95	0.07	0.723
$h(10)$	0.90	0.88	0.10	0.91	0.91	0.09	0.90	0.89	0.08	0.849
α_{min}	0.83	0.81	0.12	0.84	0.85	0.11	0.83	0.82	0.10	0.882
α_{max}	1.32	1.30	0.12	1.34	1.33	0.19	1.33	1.34	0.19	0.881
$\Delta\alpha$	0.49	0.50	0.19	0.50	0.46	0.21	0.50	0.52	0.22	0.957
$D(\alpha_{min})$	0.27	0.29	0.18	0.27	0.27	0.23	0.29	0.31	0.27	0.919
$D(\alpha_{max})$	0.28	0.35	0.24	0.28	0.33	0.26	0.28	0.27	0.25	0.996
ΔD	0.86	0.87	0.17	0.90	0.87	0.20	0.91	0.85	0.25	0.702

M – mean; ***Me*** – median; ***SD*** – standard deviation; ***p*** – p-value

TABLE 9.2. Multifractal analysis parameters between groups – night

	group A			group B			group K			
parameter	***M***	***Me***	***SD***	***M***	***Me***	***SD***	***M***	***Me***	***SD***	***p***
$h(1)$	1.16	1.16	0.12	1.11	1.10	0.11	1.08	1.08	0.10	0.006**
$h(-10)$	1.56	1.58	0.17	1.52	1.55	0.16	1.57	1.59	0.19	0.424
$h(-5)$	1.50	1.52	0.15	1.46	1.48	0.16	1.51	1.50	0.18	0.397
$h(-2)$	1.41	1.42	0.13	1.36	1.36	0.14	1.40	1.37	0.16	0.323
$h(2)$	1.09	1.10	0.12	1.04	1.04	0.12	1.01	1.00	0.10	0.011*
$h(5)$	0.96	0.97	0.13	0.92	0.95	0.15	0.89	0.89	0.12	0.128
$h(10)$	0.88	0.90	0.14	0.85	0.86	0.17	0.82	0.84	0.14	0.273
α_{min}	0.80	0.80	0.16	0.76	0.77	0.18	0.74	0.76	0.16	0.412
α_{max}	1.63	1.65	0.18	1.59	1.60	0.17	1.64	1.67	0.20	0.514
$\Delta\alpha$	0.83	0.80	0.22	0.83	0.81	0.26	0.89	0.93	0.24	0.359
$D(\alpha_{min})$	0.13	0.14	0.22	0.18	0.13	0.24	0.20	0.19	0.24	0.488
$D(\alpha_{max})$	0.32	0.32	0.19	0.30	0.28	0.18	0.29	0.30	0.20	0.819
ΔD	0.32	0.31	0.19	0.28	0.25	0.21	0.29	0.26	0.20	0.651

*$p < 0.05$; **$p < 0.01$; ***M*** – mean; ***Me*** – median; ***SD*** – standard deviation; ***p*** – p-value

For the parameter *h*(1), significant differences were observed between group A and the control group measured at night. The highest result was obtained in group A ($M = 1.16$; $SD = 0.12$), while the lowest result was obtained in the control group ($M = 1.08$; $SD = 0.10$) (Fig. 9.5).

The parameter *h*(2) differed significantly between Group A ($M = 1.09$; $SD = 0.12$) and the control group ($M = 1.01$; $SD = 0.10$) measured at night (Fig. 9.6).

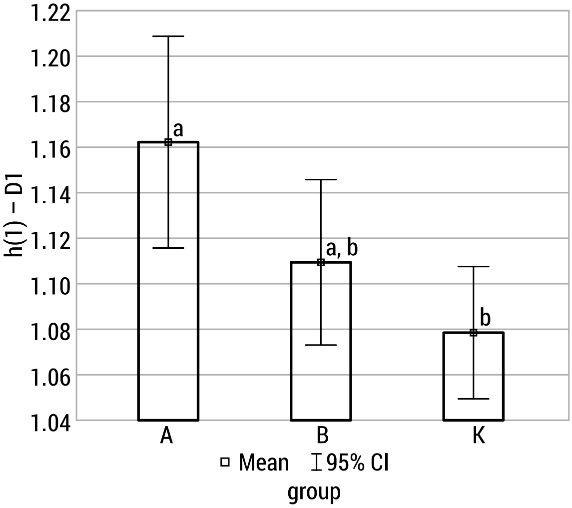


FIGURE 9.5. The *h*(1) parameter results between groups measured at night (letters a, b indicate homogeneous groups)

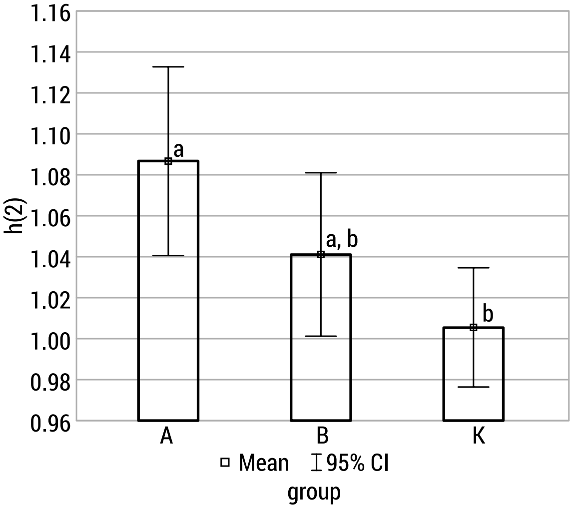


FIGURE 9.6. The *h*(2) parameter results between groups measured at night (letters a, b indicate homogeneous groups)

The time series recorded during the day and at night in each group were also compared among themselves. The graphs of the spectra of singularity $D(\alpha)$ over the singularity strength α for individual groups are presented in Figure 9.7, with a distinction for day and night. The graphs show the multifractal nature of all groups.

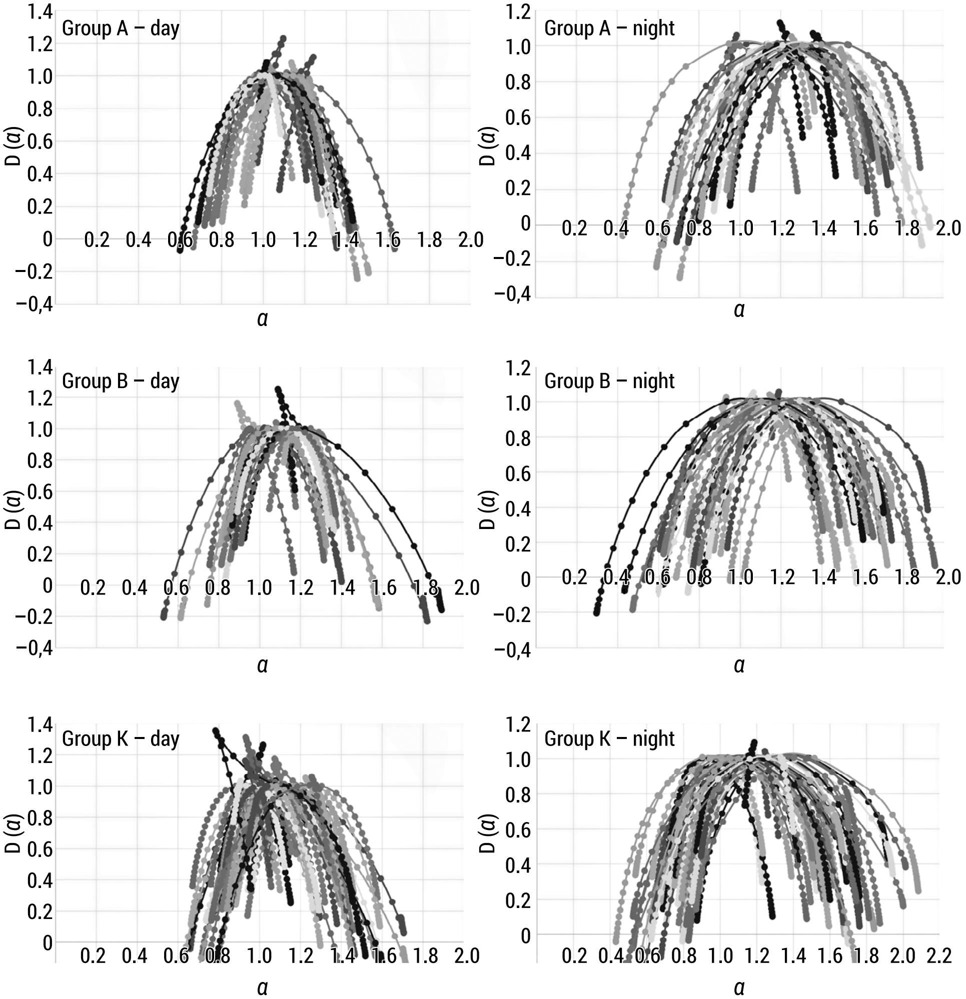


FIGURE 9.7. Summary of singularity spectra $D(\alpha)$ over the singularity strength α for individual groups with a distinction for day and night

There were statistically significant differences in Group A between day-night recordings in the $h(1) - D1$ ($p < 0.001$), $h(-10)$ ($p < 0.001$), $h(-5)$ ($p < 0.001$), $h(-2)$ ($p < 0.001$), $h(2)$ ($p = 0.003$), α_{max} ($p < 0.001$), $\Delta\alpha$ ($p < 0.001$). Parameter values were significantly higher in measurements taken during the night. Statistically significant differences were also shown in the values of D(amin) ($p = 0.006$). The values of the parameter were significantly higher in measurements taken during the day (Tab. 9.3).

TABLE 9.3. **Analysis of day-night signal statistically significant changes – group A**

	day			night			
parameter	***M***	***Me***	***SD***	***M***	***Me***	***SD***	***p***
$h(1)$	1.03	1.03	0.06	1.16	1.16	0.12	<0.001***
$h(-10)$	1.24	1.23	0.10	1.56	1.58	0.17	<0.001***
$h(-5)$	1.18	1.17	0.08	1.50	1.52	0.15	<0.001***
$h(-2)$	1.12	1.12	0.06	1.41	1.42	0.13	<0.001***
$h(2)$	1.01	0.99	0.07	1.09	1.10	0.12	0.003**
α_{max}	1.32	1.30	0.12	1.63	1.65	0.18	<0.001***
$\Delta\alpha$	0.49	0.50	0.19	0.83	0.80	0.22	<0.001***
$D(\alpha_{min})$	0.27	0.29	0.18	0.13	0.14	0.22	0.006**

***p* < 0.01; ****p* < 0.001; *M* – mean; *Me* – median; *SD* – standard deviation; *p* – p-value

Statistically significant differences were shown in Group B between day-night recordings in the $h(-10)$ ($p < 0.001$), $h(-5)$ ($p < 0.001$), $h(-2)$ ($p < 0.001$), α_{max} ($p < 0.001$) and $\Delta\alpha$ ($p < 0.001$). Parameter values were significantly higher in measurements taken during the night. Statistically significant differences were also shown in the values of the αmin ($p = 0.031$). The values of the parameter were significantly higher in measurements taken during the day (Tab. 9.4).

TABLE 9.4. **Analysis of day-night signal statistically significant changes – group B**

	day			night			
parameter	***M***	***Me***	***SD***	***M***	***Me***	***SD***	***p***
$h(1)$	1.07	1.06	0.07	1.11	1.10	0.11	0.136
$h(-10)$	1.27	1.27	0.17	1.52	1.55	0.16	<0.001***
$h(-5)$	1.21	1.22	0.15	1.46	1.48	0.16	<0.001***
$h(-2)$	1.15	1.15	0.10	1.36	1.36	0.14	<0.001***
α_{min}	0.84	0.85	0.11	0.76	0.77	0.18	0.031*
α_{max}	1.34	1.33	0.19	1.59	1.60	0.17	<0.001***
$\Delta\alpha$	0.50	0.46	0.21	0.83	0.81	0.26	<0.001***

***p* < 0.01; ****p* < 0.001; *M* – mean; *Me* – median; *SD* – standard deviation; *p* – p-value

In the K group, statistically significant differences were observed for the $h(-10)$ ($p < 0.001$), $h(-5)$ ($p < 0.001$), $h(-2)$ ($p < 0.001$), α_{max} ($p < 0.001$), $\Delta\alpha$ ($p < 0.001$). Parameter values were significantly higher in measurements taken during the night. Statistically significant differences were also shown for the values of $h(5)$ ($p = 0{,}006$), $h(10)$ ($p = 0{,}007$) α_{min} ($p = 0.008$). The values of the parameter were significantly higher in measurements taken during the day (Tab. 9.5).

TABLE 9.5. Analysis of day-night signal statistically significant changes – group K

	day			night			
parameter	***M***	***Me***	***SD***	***M***	***Me***	***SD***	***p***
$h(1)$	1.06	1.06	0.08	1.08	1.08	0.10	0.166
$h(-10)$	1.26	1.29	0.17	1.57	1.59	0.19	<0.001***
$h(-5)$	1.20	1.23	0.15	1.51	1.50	0.18	<0.001***
$h(-2)$	1.14	1.14	0.11	1.40	1.37	0.16	<0.001***
$h(2)$	1.02	1.02	0.07	1.01	1.00	0.10	0.218
$h(5)$	0.96	0.95	0.07	0.89	0.89	0.12	0.006**
$h(10)$	0.90	0.89	0.08	0.82	0.84	0.14	0.007**
α_{min}	0.83	0.82	0.10	0.74	0.76	0.16	0.008**
α_{max}	1.33	1.34	0.19	1.64	1.67	0.20	<0.001***
$\Delta\alpha$	0.50	0.52	0.22	0.89	0.93	0.24	<0.001***

$^{**}p < 0.01$; $^{***}p < 0.001$; *M* – mean; *Me* – median; *SD* – standard deviation; *p* – p-value

9.4. Conclusions

As a result of the analysis, it was found that all the three studied groups showed a multifractal character. Thanks to this, it was possible to examine how the parameters behave in each group, and compare them to each other, indicating the differences. Statistical differences between the groups were observed in the resting (night) part of the signals. The differences were noticeable for the generalized Hurst exponent $h(q)$ for positive q values.

The parameters $h(1)$ and $h(2)$ showed significant statistical differences between the group of children with type 1 diabetes without subsequent vascular complications (group A) and the group of healthy children (group K). The mean value of the $h(1)$ parameter for group A was higher than for the K group, as was the mean value of the $h(2)$ parameter.

The nature of the $D(\alpha)$ singularity spectrum (width α) for each of the considered cases indicates a strong multifractality of the data. It is noticeable, however, that in group B (children with type 1 diabetes with vascular complications) the strength of the multifractal data is greater than in the other two groups. Moreover, in all studied groups, a stronger multifractality is noticeable at night than during the day.

The results presented in the article give rise to further research on an attempt to distinguish between the presented groups. This could affect the predictability of vascular complications in children with type 1 diabetes. Further fractal analysis may improve the results.

Acknowledgement: The study was performed as part of the projects WI/WM-IIB/2/2021, WZ/WM-IIB/3/2021, and was partially financed with funds from the Polish Ministry of Science and Higher Education.

References

1. Peng C.K. et al.: "Mosaic organization of DNA nucleotides", Physical Review E, 49(2), 1685, 1994.
2. Peng CK. et al.: "Quantification of scaling exponents and crossover phenomena in nonstationary heartbeat time series", Chaos: An Interdisciplinary Journal of Nonlinear Science, 5(1), 82–87, 1995.
3. Kantelhardt J.W. et al.: "Multifractal detrended fluctuation analysis of nonstationary time series", Physica A: Statistical Mechanics and Its Applications, 316(1–4), 87–114, 2002.
4. Pavlov A.N. et al.: "Extended detrended fluctuation analysis of electroencephalograms signals during sleep and the opening of the blood-brain barrier", Chaos: An Interdisciplinary Journal of Nonlinear Science, 30(7), 073138, 2020.
5. Ivanov, P.C. et al.: "Multifractality in human heartbeat dynamics", Nature, 399(6735), 461–465, 1999.
6. Pavlov A.N. et al.: "Detrended fluctuation analysis of cerebrovascular responses to abrupt changes in peripheral arterial pressure in rats", Communications in Nonlinear Science and Numerical Simulation, 85, 105232, 2020.
7. Madanchi A. et al.: "Scaling behavior in measured keystroke time series from patients with Parkinson's disease", The European Physical Journal B, 93(7), 1–8, 2020.
8. Shadkhoo S. et al.: Scaling behavior of earthquakes' inter-events time series, Open Physics, 7(3), 620–623, 2009.
9. Meyer PG., Kantz H.: "A simple decomposition of European temperature variability capturing the variance from days to a decade", Climate Dynamics, 53(11), 6909–6917, 2019.
10. Wang F. et al.: "Multifractal detrended fluctuation analysis for clustering structures of electricity price periods", Physica A: Statistical Mechanics and its Applications, 392(22), 5723–5734, 2013.
11. Leung Y., Ge E., Yu Z.: "Temporal scaling behavior of avian influenza A (H5N1): the multifractal detrended fluctuation analysis", Annals of the Association of American Geographers, 101(6), 1221–1240, 2011.
12. Lee M. et al.: "Asymmetric market efficiency using the index-based asymmetric-MFDFA", Physica A: Statistical Mechanics and Its Applications, 512, 1278–1294, 2018.
13. Makowiec D. et al.: "Multifractal estimates of monofractality in RR-heart series in power spectrum ranges", Physica A: Statistical Mechanics and Its Applications", 388(17), 3486–3502, 2009.
14. Molina AA., Jiménez, RR., Diosdado, AM.: "Multifractal analysis of ECG time series of stress tests in healthy subjects", [in:] AIP Conference Proceedings (Vol. 2090, No. 1, p. 050001). AIP Publishing LLC, 2019.
15. Gorshkov O., Ombao, H.: "Evaluation of monofractal and multifractal properties of inter-beat (R-R) intervals in cardiac signals for differentiation between the normal and pathology classes", IET Signal Processing, 13(9), 798–805, 2019.
16. Kokosińska D. et al.: "Asymmetric multiscale multifractal analysis (AMMA) of heart rate variability", Physiological Measurement, 42(8), 085003, 2021.

17. Gorjão LR. et al.: “MFDFA: Efficient multifractal detrended fluctuation analysis in python”, Computer Physics Communications, 273, 108254, 2022.

DOI: 10.24427/978-83-67185-49-3_10

Chapter 10
Association rules in medical data classification

Anna Kasperczuk, Aleksandra W. Jasińska*,*
Małgorzata Zdrodowska, Agnieszka Dardzińska-Głębocka***

**Bialystok University of Technology, Institute of Biomedical Engineering,*
e-mail: a.kasperczuk@pb.edu.pl, ajasinska1996@gmail.com, m.zdrodowska@pb.edu.pl
***Bialystok University of Technology, Institute of Mechanical Engineering,*
e-mail: a.dardzinska@pb.edu.pl

Abstract: This chapter is focused on the application of association rules to classify medical data. There is no data mining technique that can produce consistent results for all types of medical data. The effectiveness of a data mining technique depends, among other things, on the type of data collected for study. The purpose of the following work is to see if combining two data mining methods will give better results than using them individually. For this purpose, data on breast cancer in women was used. The results showed that using a hybrid combination of association rules to produce a classifier can improve classification performance.

Keywords: association rules, classification, classification based on association (CBA)

10.1. Introduction

To make a correct diagnosis, doctors need information about the patient's health. The human body is very complex and sometimes a small deviation from the norm may indicate a disease. Therefore, it is important to analyze the collected data and extract as much useful information as possible. Due to the development of technology during the examination a huge amount of data is generated, which cannot be interpreted without the use of appropriate analytical systems. Knowledge mining methods can be useful here. The knowledge gained by means of knowledge mining methods is used in many areas of medicine. Data mining methods not only enable to make a diagnosis based on symptoms, but also estimate the prognosis of treatment progress or its costs. Relationships and patterns obtained through data mining techniques are also used to create decision support systems.

One method of data mining is classification, which involves finding a way to map data into a set of predefined classes. A model is built based on the contents of the database, which is used to classify new objects, for example, in a medical database rules can be found to classify individual diseases, and then, using the rules found, further patients can be automatically diagnosed. Another method is association discovery

– a method of finding relationships between groups of elements appearing in data sets. It belongs to one of the most popular data mining techniques. This method makes it possible to find interesting correlations occurring between data in high-volume datasets that would be hard to spot with the naked eye. This paper investigates into whether algorithms combining association discovery and classification produce measurable results in the process of extracting classification rules. The goal was to see if combining two data mining methods would yield better results than using them individually. Thus, it was tested whether using a hybrid combination of association rules to produce a classifier can improve classification performance.

The purpose of this chapter is to discuss selected association rule mining algorithms and present their application in medical data classification.

10.2. Association rules

10.2.1. Association rule construction

Association rules can be compared to "*if – then*" constructs, which have found their way into programming languages. The most popular application of association rules is the basket analysis. It involves discovering patterns of customer behavior, that is, finding groups of products that are purchased together. Although association rules have many applications, it is on the basis of the above mentioned analysis that a formal description of association rules was created [1, 2].

Let $I = \{i_1, i_2, \ldots, i_n\}$ be the set of all available goods. Any set $X \subseteq I$ will be called a set of goods. The set of goods with power k will be called the k-commodity set. The transaction base over the set I will be termed the set of transactions $D = \{T_1, T_2, \ldots, T_k\}$, where each transaction is represented by a pair $T = (id, X)$ consisting of the transaction identifier id and the set of goods $X \subseteq I$ occurring in the transaction. For simplicity, it is assumed that the transaction identifiers come from a finite set of indices $Tid = \{id_1, id_2, ..., id_m\}$, hence each transaction base D is a sort of binary relation between the sets Tid and I, and it can be expressed as $D \subseteq Tid \cdot \mathrm{P}(I)$. It can be said that a transaction $T = (id, X)$ covers a set of goods A if $A \subseteq X$. The coverage of the set of goods A in the base D is the set $\mathrm{cover}(A, D) = \{T \in D : T \text{ covers } A\}$. An association rule is any implication $A \Rightarrow B$, where A and B are the sets of goods $A \subseteq I$, $B \subseteq I$ and $A \subseteq B = C$ [1, 2].

10.2.2. Rule evaluation measures

Many pairs can be generalized in the form $A \Rightarrow B$. To select the useful ones, several measures of decision rule evaluation are introduced [2]:

- support – this is a parameter that determines how often an object appears in the dataset, or more specifically, what percentage of all association rules a particular rule represents. The formula represents the ratio of the number of transactions (containing entirely goods from set A and B) in the data set (transaction database) to the number of all transactions. Typically, support is expressed as a percentage:

$$\text{support}(A \Rightarrow B) = \frac{|P_{A\cup B}|}{|P|}; \tag{10.1}$$

- confidence – this is a parameter that describes how many percent of association rules starting with a particular set of conditional values a rule represents. In other words, it is the ratio of the number of occurrences of transactions that contain all the goods from set A and B to the number of transactions that contain only goods from set A. As with support, confidence is usually expressed as a percentage:

$$\text{confidence}(A \Rightarrow B) = \frac{|P_{A\cup B}|}{|P_A|}; \tag{10.2}$$

- conviction – this parameter was developed as an alternative to certainty, which is not as accurate in terms of the direction of association. This measure compares the probability that A would occur in a transaction without B if they were dependent on the actual frequency of occurrence of A without B. In this respect, this parameter is similar to the increment measure, but unlike it, it is a directed measure due to the fact that it also uses information about the absence of a successor. In the case, where A and B are independent of each other, the belief value equals 1. Rules with high importance reach values away from 1:

$$\text{conviction}(A \Rightarrow B) = \frac{1-\text{support}(B)}{1-\text{confidence}(A \Rightarrow B)}; \tag{10.3}$$

- lift – the increment measure is a parameter that describes the correlation between the predecessor and successor of an association rule. If the events that are represented by the predecessor and successor of the rule are independent, then the value of the increment parameter is 1. If this value is less than 1, it means that the mentioned events are negatively correlated. In such a case, the rule opposite to the one found should be considered. When the value of the parameter is greater than 1, then the events correlate positively:

$$\text{lift}(A \Rightarrow B) = \frac{\text{confidence}(A \Rightarrow B)}{\text{support}(B)}; \tag{10.4}$$

- leverage – this parameter measures the difference of A and B that occur together in the data set and what would be expected if A and B were statistically dependent.

The goal is to determine how many more units (*A* and *B* together) were in transactions than would be expected from independent transactions. The range of this parameter is between –1 and 1, with 0 indicating independence.

$$\text{leverage}(A \Rightarrow B) = \text{support}(A \Rightarrow B) - \text{support}(A) \cdot \text{support}(B). \tag{10.5}$$

A good rule is one that has high support (applies to a high number of transactions) and confidence (right side occurs frequently). In practice, it is common to assume minimum values on both of these ratios and look for all rules with support/confidence values higher than the assumed minimum [2, 3].

10.2.3. Association rules extraction

To begin searching for association rules, we first need to determine the minimum level of support ($p_{\min}$) and confidence ($s_{\min}$) that they must satisfy. This threshold should not be too low. Setting the support threshold too low leads to a large number of rules in which it is very difficult to find really useful relationships. Even if it seems that a regularity has been found, it may turn out to be an accident and within the whole database the support of the rule turns out to be too low. A rule is said to be valid and interesting when the confidence and support values exceed a pre-set threshold. Such a threshold is usually defined by a domain expert or the user himself. The set of elements that meets the support threshold $p_{\min}$ is called “attended” and is denoted as L5. An association rule can be called strong when it satisfies both the support threshold $p_{\min}$ and the confidence threshold $s_{\min}$ [2,3].

An important element when finding correlations between elements in a set is choosing the right algorithm. Determining which algorithm will be the most effective is not easy. This is because there are many algorithms and types of datasets. This makes it difficult to clearly determine which one should be used. There are two ways to check which algorithm will perform best on a given dataset. The first is to test each algorithm in turn. Then, based on the results, you can determine which one is best. The second way is to rely on the opinion of experts in the field and their experience.

The two most popular algorithms used for association discovery are the Apriori Algorithm and the FP-Growth Algorithm:

Algorithm Apriori is considered as the most popular algorithm for extracting binary association rules. It searches a dataset and in each successive iteration creates families of increasingly numerous frequent sets. These represent simple patterns that tell which elements are logically frequent [3–5].

The initial step before using the algorithm is to make sure that the transaction elements from the set *D* are ordered lexicographically. The algorithm then extracts the encountered sets of one-elements occurring in transactions from the database in the first step. Next, the algorithm checks whether they have the required support level ($p_{\min}$). Based on the found one-element frequent sets, the algorithm creates two-element candidate sets. These constitute the potential frequent sets. For each of the newly created sets, the support in the database D is calculated. If a candidate set reaches the required support level ($p_{\min}$), it is added to the list of frequent sets. In the next step, it will be used to generate three-element candidate sets. Then, the three-element frequent sets will be used to generate four-element candidate sets. Subsequently, the four-element frequent sets will be used to generate five-element candidate sets, and so on. This process is based on the frequent sets that were found in the previous step. The algorithm in each successive step creates candidate sets a size larger than the previous one. To check whether they have the required level of support, at each step of the algorithm, a read of the database D is performed. The algorithm runs until it is impossible to generate further frequent sets. The final result of the algorithm is the sum of k-element frequent sets [3–5].

During the development of the **FP-Growth Algorithm**, a different approach was used to create frequentist sets. This process is done in two steps. They are [3, 6, 7]:

- compression of the database D into an FP-tree,
- exploration of the FP-tree to find the frequent sets.

Similar to the Apriori algorithm, the FP-Growth algorithm starts by extracting all one-element frequent sets from the database. Subsequently, all elements that are not one-element frequent sets are removed from the transactions belonging to the database D. This means that the output database is transformed into a smaller database containing only one-element frequent sets. Therefore, this step is called database compression. Next, the transaction elements are sorted in terms of their support values. The list is created from the largest value to the smallest value [3, 6, 7].

The list thus sorted will be used in the next step of the algorithm to form an FP-tree. It is a rooted acyclic graph, labeled at the vertices. At the very beginning the root of the FP-tree is created. It is labeled “null”. Then, for the first transaction t1, a path is created that begins at the root of the FP-tree. The order of vertices in the path that represents the transaction corresponds to the order of the elements in the list created earlier. The value of the transaction counter for each vertex that is part of the path is initially 1. The final result of the transformation is an FP-tree [3, 6, 7].

To speed up the FP-tree search, an array of element headers is created that acts as a directory. The array informs about the location of the element in the FP-tree. If an element occurs in the FP-tree more than once, then pointers to specific vertices form a list of pointers [3, 6, 7].

In the second step of the algorithm, the FP-tree is explored to locate frequent sets. To do so, one needs to find all paths, for each 1-element alpha frequent set, that have as their end vertex a vertex representing the alpha set. A single path with an alpha final vertex is called an alpha pattern prefix path. Each such path has a path frequency counter, with its value corresponding to the value of the transaction counter at the end vertex of the path that represents the alpha set. A conditional pattern base is formed by the set of all pattern prefix paths. This base is used to construct the so-called conditional FP-tree of the pattern alpha (Tree alpha). The conditional FP-tree is then recursively explored. The goal is to find all frequent sets that contain the set alpha. When the FP-tree has only a single path p, then for each combination of beta vertices of path *p*, a set beta + alpha is generated, whose support is equal to the minimum support of the elements belonging to the set beta. The procedure looks different when the FP-tree has more than one path. Then, for each element alpha_i that belongs to the Tree element header array, a set beta = alpha_i + alpha is generated, whose support equals the support of the elements alpha_i. Then a conditional beta pattern base is created and a conditional FP-tree of the beta pattern is created, which is denoted as Tree-beta. The FP-Growth procedure with the Tree-beta and beta parameters is run until Tree-beta is empty [3, 6, 7].

10.3. Classification

10.3.1. Characteristics and steps of classification

Classification is a very important data mining method in the medical data area. In medical systems, it is used primarily to find rules for classifying individual diseases. Then, thanks to the found rules, it is possible to automatically diagnose the next patients. Classification consists in finding mappings of data into a set of predefined classes. Based on the information from the database, a model is built. The database is divided into two sets: training set and test set. The training set is used to build the model while the test set is used to test the model. The model is designed to classify new objects in the database. Classification is a data mining method with a teacher (supervision). The classification process can be divided into several steps [3, 8, 9]:

- the first step is to build a model, or classifier, that describes a predefined set of data classes. The input data for performing classification is a training set of tuples, which is a list of attribute values (descriptors) and a selected decision attribute. The classifier operates based on the values of the other attributes;
- the second step is to test the quality of the constructed classifier. The testing phase consists of checking the accuracy of the classifier based on a test set. This set should be separated before starting to build the model so that the classification result is not falsified. The most commonly used coefficient to describe the quality of classification is the overall classification accuracy. This coefficient is the ratio of cases classified correctly to all cases taken for classification;
- the last step is to apply the classification using the previously built model to the new cases based on the values assigned to their attributes.

A great number of classification methods have been developed so far. Only the classifiers used in this paper will be discussed below.

10.3.2. Naive Bayes Classifier

The Naive Bayes classifier has gained popularity among classifiers because of its simplicity and its efficiency, which is close to the results of much more complex classifiers. It belongs to the linear classifiers. Its performance is based on the conditional probability of Bayes' theorem [8–10]:

$$P\left(\frac{A}{B}\right)=\frac{P\left(\frac{B}{A}\right)P(A)}{P(B)}. \tag{10.6}$$

The classification process works by comparing the chance of a set of specifically defined object attribute values for each class. This classifier assumes that the values assigned to the attributes in the classes are independent. Because of this, the word "naïve" appears in its name. One then speaks of the so-called conditional independence of the class [8–10].

Suppose that each object is represented by a vector of attribute values $X = (F_1, F_2, ..., F_n)$. For each class C_i, the probability of an object belonging to that class is calculated. The necessary condition is that the object has the given attribute value, i.e. $P(C_i/X)$. A given object is assigned to the class that has the highest probability of such an event. To calculate this probability Bayes' theorem of conditional probability is used (1.6). When substituted into the formula the following is obtained [8–10]:

$$P(C_i / X)=\frac{P(X / C_i)P(C_i)}{P(X)}=\frac{P(f_1 / C_i)P(f_2 / C_i)\ldots P(f_n / C_i)P(C_i)}{P(X)}, \tag{10.7}$$

where: $P(X)$ – the probability of an instance of X (with a vector of attribute values $(f_1, f_2, ..., f_n)$, $P(C_i)$ – the probability of occurrence of the class C_i, $P(X/C_i)$ – the probability of occurrence of the given configuration of attribute values in class C_i.

All probabilities are taken from the probability tables, which were built using the data from the training set.

10.3.3. J48 tree

J48 tree is generated using algorithm C4.52, which works by splitting the original data set in terms of each variable. In this way there are as many types of splitting as there are explanatory variables in the set. The C4.52 algorithm is an extended version of the ID3 algorithm, whose main problem was the unnecessary growth of the tree, as well as the lack of appropriate mechanisms that would counteract the overfitting phenomenon. This led to excessively high error rates on the actual data. This phenomenon was eliminated by the so-called pruning, which aims to increase the generalization of the evaluation. The C4.52 algorithm visits each decision node recursively and selects a possible split as long as further splits are possible. For each such split, the value of the "information gain" metric is calculated. Information gain is defined as the entropy gain for each subset. The variable that has the highest information gain coefficient is assigned as the first node of the tree. This operation is repeated for all subsets until all instances are exhausted [8, 11].

10.3.4. Classification based on association

Associative classification is a branch of data mining that uses association rule discovery methods in classification problems. The term associative classification is a model used to label new records using association rules [12]. Most associative classifiers read a list of rules one at a time so that the first matching rule is used to label new records. Association classifier rules inherit some association rule indicators, such as trust or support. The indicators are used to filter or order the rules in the model and also to evaluate their quality.

The best known association-rule-based classification model is Classification Based on Association (CBA) [12–14]. The CBA algorithm uses a special type of association rules for classification. These are called class association rules (CAR). These rules have the constraint that the successor of a rule is a single item that is associated with a class label. The CBA algorithm can be divided into 3 steps [15]:

- Association Rule Extraction (CAR),
- pruning and organizing the extracted rules,
- classification of new objects.

The first step consists in modifying the Apriori algorithm. To reduce the size of the considered CAR sets, pruning based on the pessimistic rule error rate can be applied. The precision of the classifier does not depend on pruning, so implementations of the regular Apriori algorithm can also be used [12]. Simply extracted rules are not used for classification because the number of rules can be very large, even in small datasets. The generation of conflicting rules, as well as the lack of a default rule, can also be an obstacle. The CBA algorithm dealt with this problem by using data range pruning, which reduced the number of extracted rules.

There are two versions of the CBA algorithm that differ in the running time. The M1 version is slower, while the M2 version runs faster due to data access limitation. Rules in pruning are ranked in order of their strength:

- a rule is ranked higher if its confidence is greater than the confidence of the next rule;
- if a tie occurs when ranking against the trust value, then the support value is taken into account, then the rule with the higher support value is ranked higher;
- if there is a tie in both cases, then the ranking is determined by the time of rule extraction, the rule extracted earlier takes the higher place in the ranking.

Rules are processed in a ranking order. When each rule is processed, all matching transactions are removed. In case a rule does not cover at least one of the instances, it is also pruned. In the CBA algorithm, data pruning is combined with the cleaning of default rules, i.e. rules added at the end of the rule set with the majority class in the uncovered transactions on the left side and in the empty left side. This allows for a classification even when the query does not match any of the rules. If the default rule reduces the total number of errors in the learning set, then the algorithm clears all the rules below it [2].

10.4. Materials and Methods

The analysis involved breast cancer data from the Wisconsin dataset. The data was downloaded from a machine learning medical data repository, the UCI Machine Learning Repository [16]. The database contains a total of 699 rows and 10 attributes described numerically. All attributes relate to breast mammography results. The first 9 attributes take values from 1 to 10, while the last argument takes values of 2 or 4, depending on whether the tumor is benign or malignant. In the discretization run, the numerical values for the first 9 arguments were divided into 3 groups as shown in Table 10.1. The values for the last argument were also assigned to their respective group.

TABLE 10.1. Attributes and their values

Attributes	Values
Nodule thickness	1–3: {small} 4–7: {medium} 8–10: {large}
Uniformity of cell size	1–3: {small} 4–7: {medium} 8–10: {high}
Regularity of cell shape	1–3: {small} 4–7: {medium} 8–10: {high}
Marginal adhesion	1–3: {small} 4–7: {medium} 8–10: {large}
Size of individual epithelial cells	1–3: {small} 4–7: {medium} 8–10: {large}
Naked cell nuclei	1–3: {small} 4–7: {medium} 8–10: {large}
Condensed chromatin	1–3: {small} 4–7: {medium} 8–10: {large}
Nuclei with normal structure	1–3: {small} 4–7: {medium} 8–10: {large}
Mitosis	1–3: {small} 4–7: {medium} 8–10: {high}
Malignancy of tumor	2 (benign) or 4 (malignant)

The data used the analysis contain valuable information from the mammography examination that can help to make an accurate diagnosis. When cells are cancerous, they are often grouped in multiple layers, unlike benign lesions that are grouped in monolayers. An important aspect in determining whether cells are cancerous is the uniformity of their size and shape. This is because cancer cells tend to be heterogeneous. When it comes to the ability of cells to adhere marginally, or the so-called marginal adhesion, normal cells tend to clump together. Cancer cells lose this ability, so loss of marginal adhesion may be indicative of malignancy. The size of individual epithelial cells is related to the homogeneity mentioned earlier. Enlargement of epithelial cells may indicate malignancy. Naked cell nuclei are nuclei that are not surrounded by cytoplasm. They are usually found in benign lesions. Condensed chromatin describes the uniform structure of the nucleus, which may indicate that it is not a malignant lesion because the condensed chromatin in this type of cell is thicker. Normally structured nuclei are small structures that are found in the cell nucleus. In noncancerous cells, the nucleus is usually very small, sometimes even invisible. In transformed cells, the nuclei are larger, and sometimes the number of nuclei is also increased. Another important aspect is mitosis, which is the division of a cell into two identical daughter cells. Skilled pathologists are able to determine the stage of cancer by counting the number of mitoses [17].

To perform data mining, in terms of association rule mining, it is necessary to define two significant parameters: minimum confidence level and minimum support level. The values of these parameters have a great influence on the final result of the extracted association rules. This is because they depend on the number of rules that will be found.

An appropriately chosen support level allows to regulate the number of frequent sets, which will be used to create association rules. The confidence level is used to regulate the reliability of rules. This means that it is determined with what the smallest probability particular rules occur in the examined data set. Increasing the value of the support level decreases the number of selected frequent sets. Conversely, increasing the confidence level reduces the number of rules found. Decreasing these values works the other way around and increases the number of frequent sets and association rules.

The Apriori algorithm was used to perform the analysis. The minimum confidence and support values used in association mining were initially: 0.9 and 0.7. In the next trial, the minimum support value was changed to 0.65.

10.5. Results

Association rules were extracted using the program Weka. Table 10.2 shows the results of the analysis for a minimum support of 70% and a minimum confidence value of 90%. The Apriori algorithm found 489 candidate sets from which 3 strong association rules were created. The rules found during this analysis are as follows:

- at least 70% of cases in which the size of single epithelial cells was small also showed a low propensity of cells to divide with a probability of 98%;
- at least 70% of cases in which the presence of nuclei with normal structure was small also showed a low propensity of cells to divide with a probability of 98%;

- at least 70% of cases in which marginal adhesion was small also showed a low propensity of cells to divide with a probability of 97%.

The finding rules have confidence at a very high level. The value of this index for the first two rules is 0.98, while for the third rule the value dropped to 0.97. All of the found rules have a lift greater than 1, which means that they are positively correlated. The value of the leverage index is 0.04 for all rules. For the conviction index, the first rule obtained a value of 4.17. The value of this index for the second rule is 2.78, while for the third rule it is 2.59.

TABLE 10.2. Association rules for minimum support of 70% and minimum confidence of 90%

Association rules		**conf**	**lift**	**lev**	**conv**
Size of individual epithelial cells = small	Mitosis = small	0.98	1.06	0.04	4.17
Nuclei with normal structure = small	Mitosis = small	0.98	1.05	0.04	2.78
Marginal adhesion = low	Mitosis = low	0.97	1.05	0.04	2.59

TABLE 10.3. Association rules for minimum support of 65% and minimum confidence of 90%

Association rules		**conf**	**lift**	**lev**	**conv**
Malignancy = benign	Mitosis = low	0.99	1.07	0.04	8.52
Marginal adhesion = small, Nuclei with normal structure = small	Mitosis = low	0.99	1.07	0.04	4.98
Uniformity of cell size = low	Mitosis = low	0.99	1.06	0.04	4.47
Size of individual epithelial cells = small, Nuclei with normal structure = small	Mitosis = small	0.99	1.06	0.04	4.35
Marginal adhesion = small, Nuclei with normal structure = small, Single epithelial cell size = small	Mitosis = low	0.98	1.06	0.04	4.3
Epithelial single cell size = small	Mitosis = small	0.98	1.06	0.04	4.17
Condensed chromatin = small	Mitosis = low	0.98	1.06	0.04	3.27
Uniformity of cell shape = low	Mitosis = low	0.98	1.06	0.04	3.17
Nuclei with normal structure = small	Mitosis = small	0.98	1.05	0.04	2.78
Marginal adhesion = low	Mitosis = low	0.97	1.05	0.04	2.59
Cell size uniformity = low	Marginal adhesion = low	0.95	1.27	0.14	4.84

Association rules		conf	lift	lev	conv
Uniformity of cell size = low	Size of individual epithelial cells = small	0.94	1.31	0.15	4.77
Size of individual epithelial cells = small, Mitosis = small	Nuclei with normal structure = small	0.93	1.24	0.13	3.38
Size of single epithelial cells = small	Nuclei with normal structure = small	0.93	1.24	0.13	3.35
Size of individual epithelial cells = small, Mitosis = small	Marginal adhesion = small	0.92	1.22	0.12	2.91
Size of single epithelial cells = small	Marginal adhesion = small	0.91	1.22	0.12	2.89
Size of single epithelial cells = small, Nuclei with normal structure = small,	Mitosis = small	0.91	1.25	0.13	3.03
Marginal adhesion = small, Mitosis = low	Nuclei with normal structure = small	0.91	1.22	0.12	2.73
Nuclei with normal structure = small, Mitosis = small	Marginal adhesion = small	0.91	1.21	0.12	2.68
Nuclei with normal structure = small, Mitosis = small	Size of single epithelial cells = small	0.9	1.25	0.13	2.83
Size of single epithelial cells = small, Marginal adhesion = small,	Mitosis = small	0.9	1.24	0.12	2.69

Table 10.3 shows the results of the analysis for a minimum support of 65% and a minimum confidence value of 90%. The Apriori algorithm found 454 candidate sets from which 21 strong association rules were created. The rules found during this analysis read as follows:

- at least 65% of the cases in which cancer appeared benign also showed a low propensity of the cells to divide with a probability of 99%;
- at least 65% of cases in which marginal adhesion was low and the occurrence of normal nuclei was low also showed a low propensity of cells to divide with a probability of 99%;
- at least 65% of cases in which the uniformity of cell size was small also showed a low propensity of cells to divide with a probability of 99%.

The finding rules also have confidence at a very high level. The smallest value of this index is 0.90. All of the rules found have a lift greater than 1, meaning that they are positively correlated. The value of the leverage index for the argument with the lowest value is 0.04, while the highest value is 0.12. In the case of the conviction index, the first rule obtained a value of 8.52. The value of this index for the other rules remains below 5.

To complete the classification, 2 classifiers were used which will serve as a comparative element against the classification using association rules. These classifications were performed in the Weka program. J48 and Naive Bayes classifier were used to perform them. The results obtained during these classifications are included in the appendix.

The next stage of the work was to perform classification based on association rules using the CBA model. This stage of the work was performed using the R programming language. The algorithm was upgraded with association rule indices. The algorithm, in addition to calculating the measure of support and confidence, was enriched by calculating the measure of increment and conviction, which allows for better evaluation of the generated association rules, which in turn would be impossible to obtain using Weka software.

To create a classifier using association rules, rules containing information about the malignancy of the tumor are needed. From the results shown in Table 10.2, it can be seen that the minimum support level needs to be lowered because with a support of 0.65, not enough rules containing this information were found. In order to search for a sufficient number of rules containing the needed information, the minimum support level was lowered to 0.3. This resulted in 504 rules.

Next, it was verified how many rules among those found are relevant in terms of the performed analysis. The obtained results are shown in Figure 10.1.

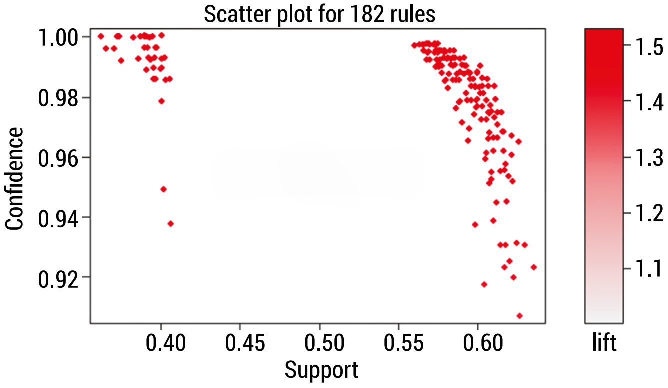


FIGURE 10.1. Rules significant in their support, confidence and lift values

After removing irrelevant rules, 182 rules were obtained, which are presented in a graph containing information about support, confidence and growth. It can be seen that the rules are found in two clusters gathered with support values of 0.4 and 0.6. Of the rules obtained, a significant number have a very high confidence index. All of the obtained significant rules have a high level of increment.

In the next step, using the created CBA model with the Apriori algorithm, a classifier was created. Before its creation, it was necessary to partition the dataset. That is, the dataset was partitioned into a learning set and a test set. The caret module was used for this purpose.

The created classifier was constrained by the values of support and confidence parameters. Since the number of rules affected the precision of the classifier, these values were chosen so that the most valuable association rules were selected to create the classifier. The minimum support and minimum confidence had values equal to, respectively: 0.1 and 0.3.

The results obtained for all 4 classifiers performed are presented below in Table 10.4.

TABLE 10.4. Comparison of classification results

	Classifier		
	J48	**NaiveBayes**	**CBA**
Operating time (sec)	0.12	0.11	1.75
Cases classified correctly	669	673	686
Cases misclassified	30	26	13
Precision	95.70%	96.30%	98.10%

Comparing the performance of the classifiers, it can be concluded that the best precision was obtained by the CBA classifier created in the R programming language. The classifier assigned 686 cases correctly, which is 98.1% of all diagnoses. Only 13 cases were classified incorrectly. The Naive Bayes classifier achieved a precision of 96.3%, meaning that 673 diagnoses were assigned correctly, while 26 were assigned incorrectly. The J48 classifier assigned 669 cases correctly, accounting for 95.7% of all diagnoses. The cases classified incorrectly include 30 diagnoses.

As can be seen, the time to build the model and test it differs significantly for the classifications performed in Weka and the classification performed in the R programming language. For the J48 classifier, the running time was 0.12 sec, while the Naive Bayes classifier did the job in 0.11 sec. The difference in the running time between the two classifiers is small at only 0.01 sec. For the CBA model built in R programming language, the time to build and test the model was 1.75 sec. The difference in the running time between the classifiers performed in Weka and the classifier performed in R programming language is about 1.63 sec.

By implementing the CBA algorithm in R programming language, 182 class association rules (CARs) were found. Table 10.5 shows the analysis results for the 10 most significant class association rules.

TABLE 10.5. Class association rules (CAR)

Association rules		conf	lift	lev	conv
Nodule thickness = small, Regularity of cell shape = small, Naked cell nuclei = small	Tumor malignancy = benign	0.371	0.99	1.53	62.9
Nodule thickness = small, Uniformity of cell size = small, Naked cell nuclei = small	Tumor malignancy = benign	0.375	0.99	1.53	62.5
Nodule thickness = small, Marginal adhesion = small, Naked cell nuclei = small	Tumor malignancy = benign	0.373	0.99	1.53	62.7
Nodule thickness = small, Uniformity of cell size = small, Regularity of cell shape = low	Malignancy of tumor = benign	0.401	0.99	1.53	59.9
Nodule thickness = small, Regularity of cell shape = small, Condensed chromatin = small	Tumor malignancy = benign	0.388	0.99	1.53	61.2
Nodule thickness = small, Regularity of cell shape = small, Size of individual epithelial cells = small	Tumor malignancy = benign	0.392	0.99	1.53	60.8
Nodule thickness = small, Marginal adhesion = small, Condensed chromatin = low	Tumor malignancy = benign	0.389	0.99	1.53	61.1
Nodule thickness = small, Size of individual epithelial cells = small, Naked cell nuclei = small, Nuclei with normal structure = small	Tumor malignancy = benign	0.363	0.99	1.53	63.7
Nodule thickness = small, Regularity of cell shape = small, Marginal adhesion = small, Nuclei with normal structure = small	Tumor malignancy = benign	0.395	0.99	1.53	60.5
Nodule thickness = small, Regularity of cell shape = small, Nucleus with normal structure = small, Mitosis = low	Tumor malignancy = benign	0.392	0.99	1.53	60.8

The support values for the found rules range from 0.363 to 0.401. The rules have confidence at a very high level. All of them have a value close to 1. All of the retrieved rules have an lift equal to 1.53, which means that they are positively correlated. For the conviction index, the smallest value is 59.9 and the largest – 63.7.

10.6. Discussion

This chapter attempts to combine two data mining methods, which are association rule mining and classification. It was examined how this combination would affect the classification results of medical data relative to using these methods separately.

Breast cancer accounts for 22.9% of all cancer cases in women worldwide. Data from 2008 show that as many as 458,503 diagnosed patients died from the disease. This represents 13.7% of all cancer deaths in women [18, 19]. Research on breast cancer shows that statistically 1 in 9 women will develop the said cancer in their lifetime. The risk of developing the disease increases with the age of the woman. For women under 25 the risk is 1 in 15,000, for women between 25 and 30 the risk increases to 1 in 1900, and for women between 30 and 40 the risk is 1 in 200 [18, 20, 21]. The course of the disease and the chance of cure is determined by the moment of detection. Therefore, preventive screening is a very important aspect. Self-examination of the breast allows for the detection of the disease at an early stage of development, which significantly increases the chance of recovery. However, the most effective prevention of breast cancer is mammography examination. Digital mammography is a radiological imaging method that is able to detect cancer before it becomes clinically apparent. However, this method is one of the most difficult to read among medical imaging methods. This is due to the low contrast and differences in tissue types. Some of the first signs of tumor formation on mammographic imaging are initial signs of masses and clusters of calcifications. However, making an accurate diagnosis is not easy. Tumors that form have a variety of shapes, with some showing features of normal tissue. This makes it difficult for even qualified specialists to make a diagnosis in the early stages of breast cancer [22]. A false positive diagnosis can result in an unnecessary biopsy. It is estimated that as many as 70 to 80 percent of all breast biopsy results do not show cancerous lesions. In the opposite case, when the diagnosis is false negative, the cancer remains undetected. This can lead not only to increased medical costs, but also life-threatening [23].

A correct diagnosis at an early stage requires a reliable and accurate diagnostic procedure that allows doctors to distinguish between malignant and benign breast cancers. Therefore, artificial intelligence is being used increasingly, which is able to classify a given case, based on the previously entered data. This avoids errors when making a diagnosis, as well as shortens the waiting time for results [24].

Based on the obtained results, it can be concluded that the combination of two data mining methods gives better results when classifying medical data. The proposed hybrid method scored 2.4% more accurate than classification using the J48 classifier and 1.8% better than classification using the naive Bayes classifier. Of all 699 patients, 686 cases could be positively classified.

A more accurate classification allows for:

- reducing the cost of further diagnosis,
- saving valuable time of doctors and patients,
- shortening the time from noticing symptoms to making a diagnosis,

- reducing stress associated with suspected cancer,
- reducing the number of unnecessary invasive procedures in the event of a false suspicion of cancer.

However, by analyzing the time in which the classification was performed, it can be concluded that the hybrid method is slower than the classification performed using the J48 classifier and Naïve Bayes. Although the time difference between the hybrid classifier and the traditional classifiers was only 1.63 sec, in the case of a huge database this difference can be much larger, which can affect the functionality of the method.

Comparing the results obtained for association rule mining, it can be seen that decreasing the value of the minimum support parameter increased the number of obtained association rules. This took place because decreasing the value of the minimum support parameter leads to an increase in the number of generated frequent sets. It can be observed that the rules extracted at a higher level of support are also in the group of rules extracted for support with a lower value.

Citing the expertise of cancer experts from the King Hussein Cancer Center [25], the attributes classified in terms of their importance in diagnosis and prediction of breast cancer are presented below in Table 10.6. The attributes were described by attribute weight and three measures of importance: low, medium, and high.

TABLE 10.6. Classification of breast cancer diagnostic attributes in terms of importance in diagnosis

Attributes	Validity	Weight
Nodule thickness	High	9
Uniformity of cell size	Medium	7
Regularity of cell shape	Average	5
Marginal adhesion	High	8
Single epithelial cell size	Medium	6
Naked cell nuclei	High	10
Condensed chromatin	High	9
Normal-structure nuclei	Low	3
Mitosis	Medium	6

Thus, considering the attribute importance weights and association rates, it can be concluded that some of the most valuable rules that were extracted are:

- at least 65% of the cases in which the cancer appeared benign also showed a low propensity of the cells to divide with a probability of 99%;
- at least 65% of cases in which the condensed chromatin structure in the nucleus was uniform also showed a low propensity of cells to divide with a probability of 98%.

Although mitosis has a medium weight in terms of importance in the diagnostic process the first rule can be considered valuable. It is the only one of the retrieved rules that shows a relationship between the diagnostic attribute and the diagnosis. It also has very good indicators of association rules, which were discussed in the theoretical part.

The second presented rule was found to be important because condensed chromatin has a very high weight in terms of importance in the diagnostic process. Its weight is 9 on a scale of 1–10. It is the attribute with the highest weight found because neither naked cell nuclei nor nodule thickness were found in the extracted rules.

For class association rule mining (CAR) in Weka, the only measure calculated is confidence. Therefore, the R programming language was used to perform the analysis, where additional metrics were calculated for the rules. This makes it possible to determine whether the created classifier is based on strong or weak rules. Analyzing the obtained results, it can be concluded that the rules that took part in the construction of the CBA classifier are valuable rules. This is evidenced by the high value of the belief parameter, which ranges from 59.9 to 63.7.

The obtained results confirm that the application of the hybrid CBA method gives measurable benefits in the process of knowledge extraction. This shows that it is worthwhile to keep looking for new solutions that may prove to be more effective than those better known and commonly used.

10.7. Conclusions

The analysis of medical data is particularly difficult due to its specificity. Information and data are obtained from medical records. If this documentation is not carefully drawn up, properly completed and collected, the information extracted from it will not have any reflection in reality. In Poland alone, the flow of information in the health care system is estimated as follows: more than 2 billion health services, 7 billion data on medical services, 1 billion data on medical events, and 38 million data on medical device [26, 27]. Due to inadequate health care funding, hospitals lack assistants to complete the database during the patient visit. This results in the doctor having to not only take a history, examine, interpret results, and inform the patient in a short period of time, but also enter this information into the database. In a hurry, it is not easy to make a mistake, and any such mistake while entering data into the system may result in falsifying the results of further examinations.

An additional problem associated with medical data is that it is particularly protected by law because of its sensitivity. This means that only authorized persons have access to them. In order to gain access to the data, special consent is needed from the patient to release it for further research. It is also important that the data is anonymized before it is passed on so that it cannot be recognized who it belongs to. This makes it difficult to search for important information about a disease in a wider group of patients. There is no such data mining technique that works for all types of medical data. Classification seeks to discover a class of attributes, but this does not take into account the relationships of those attributes. The relationships between attributes are dealt with by the association rule mining technique. In the CBA method implemented in this paper, many association rules are considered. Their number and validity can be adjusted by setting the rule indices appropriately, but this does not ensure that the discovered rule will be relevant. Only an expert in a given domain is able to determine the usefulness of the found rule. This makes that in the hybrid method for creating a classifier rules that do not contribute to anything can also be used.

Acknowledgments: This work is supported by the Ministry of Science and Higher Education of Poland under research projects No. WZ/WM-IIB/3/2021 and WZ/WM-IIM/3/2020.

References

1. Agrawal R., Imielinski T., Swami A.: "Mining Association in Large Databases", Proc 1993 ACM SIGMOD Int Conf Manag Data – SIGMOD '93, 1993:207-16.
2. Hahsler M.: A Probabilistic Comparison of Commonly Used Interest Measures for Association Rules, 2015.
3. Morzy T., Morzy M., Leśniewska A.: „Algorytmy odkrywania binarnych reguł asocjacyjnych", n.a. https://wazniak.mimuw.edu.pl/images/3/3f/ED-4.2-m03-1.0-kolor.pdf.
4. Du J., Zhang X., Zhang H., Chen L.: "Research and improvement of Apriori algorithm", 6th Int Conf Inf Sci Technol ICIST 2016, 2016:117–21. https://doi.org/10.1109/ICIST.2016.7483396.
5. Yang J., Li Z., Xiang W., Xiao L.: "An improved apriori algorithm based on features", Proc – 9th Int Conf Comput Intell Secur CIS 2013, 2013:125–8. https://doi.org/10.1109/CIS.2013.33.
6. Chen J.B., Hu W.Y., Ying K.H., Li G.N.: "A log analysis technology based on FP-growth improved algorithm", Proc – 2021 Int Conf Artif Intell Big Data Algorithms, CAIBDA 2021, 2021:219–23. https://doi.org/10.1109/CAIBDA53561.2021.00053.
7. Chen M., Gao X.D., Li H.F.: "An efficient parallel FP-Growth algorithm", CyberC 2009 – Int Conf Cyber-Enabled Distrib Comput Knowl Discov, 2009:283–6. https://doi.org/10.1109/CYBERC.2009.5342148.
8. Aggarwal C.C.: Data classification: Algorithms and applications. Boca Raton: CRC Press, Taylor & Francis Group; 2014. https://doi.org/10.1201/b17320.
9. Karpiuk M.: Klasyfikacja dokumentów tekstowych z wykorzystaniem uogólnionych wzorców sewkencyjnych. Politechnika Warszawska, 2013.
10. Yang F.J.: "An implementation of naive bayes classifier", Proc – 2018 Int Conf Comput Sci Comput Intell CSCI 2018, 2018:301–6.
11. https://doi.org/10.1109/CSCI46756.2018.00065.
12. Saravanan N., Gayathri V.: "Performance and classification evaluation of J48 algorithm and Kendall's based J48 Algorithm (KNJ48)", Int J Comput Trends Technol 2018; 59:73–80. https://doi.org/10.14445/22312803/ijctt-v59p112.
13. Liu B., Hsu W., Ma Y.: "Integrating Classification and Association Rule Mining", KDD-98 Proc 1998;98:80–6.
14. Jiří F., Kliegr T.: "Classification based on associations (CBA) – A performance analysis", CEUR Workshop Proc 2018; 2204:1–9.

15. Zheng T., Hanhu W., Mei C., Xiaoping Z.: "Improved CBA classification algorithm based on rough set", 2009 1st Int Conf Networked Digit Technol NDT 2009, 2009:43–6. https://doi.org/10.1109/NDT.2009.5272128.
16. Vanhoof K., Depaire B.: "Structure of association rule classifiers: A review", Proc 2010 IEEE Int Conf Intell Syst Knowl Eng ISKE 2010, 2010:9–12. https://doi.org/10.1109/ISKE.2010.5680784.
17. UCI Machine Learning Repository, n.a., https://archive.ics.uci.edu/ml/datasets/
18. Breast+Cancer+Wisconsin+%28Original%29.
19. Salama G.I., Abdelhalim M.B., Zeid M.A.E.: "Experimental comparison of classifiers for breast cancer diagnosis", Proc – ICCES 2012 2012 Int Conf Comput Eng Syst 2012:180–5. https://doi.org/10.1109/ICCES.2012.6408508.
20. Azar A.T., El-Said S.A.: "Performance analysis of support vector machines classifiers in breast cancer mammography recognition", Neural Comput Appl 2014; 24:1163–77. https://doi.org/10.1007/s00521-012-1324-4.
21. Boyle P., Levin B.: "World CanCer Report 2008", Cancer Control 2008; 199:512. https://doi.org/10.1016/j.cma.2010.02.010.
22. Chaurasia V., Pal S.: "A novel approach for breast cancer detection using data mining techniques", Int J Innov Res Comput Commun Eng 2014; 2:2456–65.
23. McAree B., O'Donnell M.E., Spence A., Lioe T.F., McManus D.T., Spence R.A.J.: "Breast cancer in women under 40 years of age: A series of 57 cases from Northern Ireland", Breast 2010; 19:97–104. https://doi.org/10.1016/j.breast.2009.12.002.
24. Jotwani A.C., Gralow J.R.: "Early detection of breast cancer: New biomarker tests on the horizon?" Mol Diagnosis Ther 2009; 13:349–57. https://doi.org/10.1007/BF03256340.
25. Lévy D., Jain A.: Breast Mass Classification from Mammograms using Deep Convolutional Neural Networks, n.a.
26. Mitchell T.M.: Machine Learning. New York: McGraw-Hill; 1997.
27. Alwidian J., Hammo B.H., Obeid N.: "WCBA: Weighted classification based on association rules algorithm for breast cancer disease", Appl Soft Comput 2018; 62:536–49. https://doi.org/10.1016/J.ASOC.2017.11.013.
28. Krok E.: „Pozyskiwanie i gromadzenie danych medycznych – analiza jakości." Stud i Mater Pol Stowarzyszenia Zarządzania Wiedzą 2011; 54:288–99.
29. Szymczyk D.M., Horoch A.: „Implementacja elektronicznej dokumentacji medycznej. Część 2 – korzyści dla uczestników systemu ochrony zdrowia", Med Ogólna i Nauk o Zdrowiu 2013:324–30.

DOI: 10.24427/978-83-67185-49-3_11

Chapter 11

The concept of applying gamification in biomedical engineering

Jolanta Pauk[*], *Bartosz Śliwecki*[**], *Bogdan Hościło*[*****],
Andrzej Sobolewski[****], *Begoña Garcia-Zapirain Soto*[***],
Eliza Romańczuk-Ruszuk[*], *Zofia Dzięcioł-Anikiej*[******]

[*]*Bialystok University of Technology, Institute of Biomedical Engineering*
[**]*Bialystok University of Technology, Faculty of Architecture*
[***]*University of Deusto, Faculty of Engineering*
[****]*ChM Implants&Instruments*
[*****]*Bialystok University of Technology, Institute of Mechanical Engineering*
[******]*Medical University of Bialystok, Department of Rehabilitation*

Abstract: Traditional academic education is not sufficient in adult education and not adequate for teaching either technical skills or medical skills. In recent years, especially in times of the Covid-19 pandemic, the research and teaching staff have experienced problems related to the lack of interest and decreased motivation, concentration disorders, and the deterioration of the quality of communication and interpersonal relationships in students. The aim of the study is to develop a new concept on the use of gamification in biomedical engineering aimed at improving the quality of academic teaching. The main slogan of the proposed game is Orthopedic engineering and pathology, which is part of an educational module called Introduction to biomedical engineering. The presented approach is supported by an online educational platform including gaming resources available to create an innovative biomedical engineering itinerary. The result of these works leads to the creation of a multi-layered system of interactive means of teaching, which include the use of 3D games designed in the form of escape rooms, and augmented and virtual reality tools for the display and interaction with surgical instruments. In addition, a foreseeable map of interconnected topics will be made available to students as a direct and deliberate practice of increasing awareness of the possible directions of development, as well as to let students plan out their own route in order to achieve similar and expected results. Further development of the teaching tools is planned for the upcoming year, as the implementation of the mentioned tools will be done with selected students from each university.

Keywords: biomedical engineering, gamification, remote education

11.1. Introduction

Traditional academic education, mainly teacher-centered and based on reading, listening, or watching, this time seems to be insufficient in adult education and inadequate for teaching either technical skills or medical skills. In recent years, especially in times of the Covid-19 pandemic, the research and teaching staff experienced problems related to the lack of interest and decreased motivation, concentration disorders, and the deterioration of the quality of communication and interpersonal relationships in students [1, 2]. It may indicate that the education process must be changed by moving from traditional educational methods and forms to those experience-oriented and collaborative and taking into account the features of the digital generation. In the digital generation, the educational process in medical and technical science needs to evolve beyond traditional approaches and infuse new tools and media into the curriculum. The new approach in education may be gamification, which seems to be desirable by a digital generation who are looking for new sensations [3, 4–8] and spend a lot of time playing and participating in virtual communities [9, 10]. Gamification in the education process should use different mechanisms including remote education, classroom technologies, avatar-based self-representation, narrative storylines, and tutorials [1, 2].

Within the last years, technology-enhanced active learning and multimedia applications have been developed. Moreover, many organizations supported training systems by elaborating gamification solutions to improve skills that include digital literacy, creativity, and complex thinking. The literature, offers many examples of the use of gamification applied in the academic environment. They are usually concerned with individual courses and educational platforms [11, 12], and focus on the analysis of players' behavior during the game and the impact of the use of game mechanics on the results of teaching or measuring student involvement in the teaching process [8, 13–16].

One of the significant challenges in the field of biomedical engineering is the development of effective tools for the development of research and teaching staff by increasing their digital competencies in the higher education system. Biomedical engineering is considered an interdisciplinary science focused on advances in technology and medicine development to improving human health. It applies the principles and problem-solving techniques of engineering to biology and medicine. Technical and medical universities are currently facing a shift in their teaching paradigm in both fields of medicine and biomedical engineering. The increasing amount of information and research makes it difficult for biomedical engineering education to stay current in its curriculum. Moreover, many professionals need practical experience without endangering themselves and others. Innovative approaches and modalities for education in biomedical engineering are constantly sought after to improve teaching and learning. Taking into account that biomedical engineering students are young and keen on technologies, games can be used in biomedical engineering education [3, 4].

The aim of the study is to develop a new concept on the use of gamification in biomedical engineering aimed at improving the quality of academic teaching.

11.2. Methods and results

The main slogan of the proposed game is *Orthopedic engineering and pathology*, which is part of an educational module entitled *Introduction to biomedical engineering*. The presented approach is supported by an online educational platform including gaming resources available to create an innovative biomedical engineering itinerary.

The game is a conglomerate of various types of tasks using the university's infrastructure and people associated with it (students, medical doctors, teachers) including traditional tasks, identifying places, identifying people, finding an artefact, doing a specific job, etc. Solution sub-tasks are arranged into successive tasks (e.g. gradual shading of the image, which is another puzzle). The game comes to be more and more difficult and complicated levels. The game adapts the classical elements of game mechanics: points, levels, and badges. Players will receive bonuses not only for completing tasks but also for additional, even simple activities, e.g. posting a comment on social media, or answering a question about studies or university.

Gamification has been widely used in the medical field of practice and higher education, as has – for the most part – resulted in modern technical advancements for the many branches of associated with it. Simulations made it possible to create more sophisticated human computer interaction solutions for modern medical procedures, as was the case for the intuitive surgical robot "Da Vinci" by Intuitive Surgical, Inc. With the adaptation of game driven simulators and advanced interactive game pads, the first prototype of the world-renowned surgical platform was made possible [17].

The scope of the project is prioritized in three important aspects; accurate representation of traditional educational programs, ease of use by new users – i.e. students, and scalability. The challenge associated with the first is the problem of concentrating a seemingly limitless library of case studies, content and variations, as the specificity of the medical field is as vast as the number of patients with a given issue [18]. In order for games to be able to be classified as playable, a predictable and replicable experience should be designed in such a manner that each and every player can expect to encounter similar problems and try to propose further similar solutions, which in the case of the highly individualized medical cases in patients may seem impractical. As a result of this, a detailed set of cases was configured so that both the needs of the vast library of possibilities could be covered on the most part, as well as the entry level experience of some users (students) could be within reach of the primary stages of play. Along with covering as many details as it would be practical, the game would have to remain interesting enough so that players would recognize the platform as both a learning tool and a means of personal development.

One of the main methods used to organize the ongoing work of translating a traditional course syllabus into an understandable and predictable database was the use of technology trees. Through active adaptation and the use of cause and effect relationships between certain topics of a given course, a visually understandable and relatable map is created in order to both plan and foresee the different paths a player can take in order to reach the same final destination as the other players [19]. In the case of the traditional approach, students carry out learning sequences in a linear fashion, often with multiple subjects and fields at the same time, with possible relationships being addressed at the further stages of a given course. In the case of the gamified approach, players can see the nonlinear and divergent vectors of approach which they can freely define and modify in the course of the semester as well are able to foresee possible alternatives to their strategy (Fig.11.1) [20].

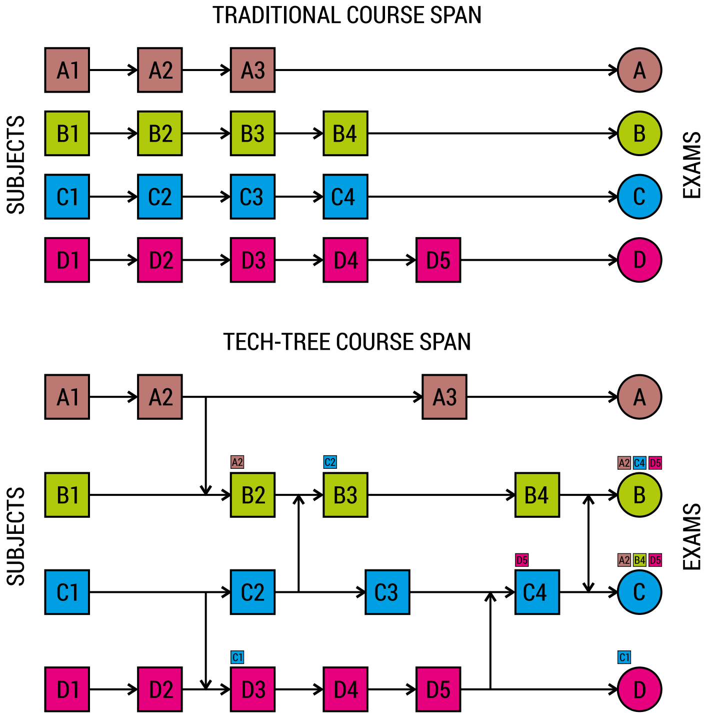


FIGURE 11.1. Comparison of a traditional approach and technology tree approach of a simulated semester of a college course

Having chosen such methods of content organization, it became clear that the crucial stage of the development would revolve around the logical allocation of game-time to element ratio throughout the span of the technology tree. A well designed and developed game has consequential ratios at different stages of the evolution of the technology tree, most of which reflect the game design style and theme. Fast-paced games would generally favor a slow movement through the stages of the tech tree as a buffering measure to ensure that the game is not finished too quickly, while slow-paced strategy games tend to favor a somewhat fast-paced evolution rate throughout game play to ensure that the game stays appealing and involving to the player base.

In the case of this project, a relatively fast-paced advancement was put in place in addition to other measures to allow the game to keep the student player base involved in the game and to increase random non-crucial character development practices to be ready to be put in place. Such actions would include side tasks that do not necessarily make substantial contributions to the knowledge about the topic, but instead involve the player in the game itself. Examples of such would be character customization, fun trivia quests and quizzes, peer-to-peer trading and communication via the platform, in game “easter egg” [21] hunts, and time events regarding seasons/times of the year. More importantly than side tasks, the importance to keep the player evenly involved in the game is just as important as the quality of the content made available in the database itself [22].

While somewhat easy at the beginning of the course, the ability to upkeep the statistics throughout the span is one of the harder challenges for game designers. The initial rush of new content featured in the game loses its glamour after a while and if not dosed correctly, could lead to monotonous labor like struggle to finalize the game [23]. In order to combat this effect, the mentioned game will host several elementary features that will become apparent only after a set amount of time has been spent playing. Some of these will include surprise mechanics, time-based events, such as challenges and bonus content, and monthly winners based on scoreboards. Although at this point in time these functions are still under development, their usage is widespread in online gaming communities, as certain games are almost solely based on constantly evolving themes and the content of a given game.

Teaching content in the game will revolve around the commonly accessible materials made available to university domains. Most of these materials are in the form of texts, graphic representations, 3D models, pictures and audio-visual files, and so their implementation into the game will focus primarily on the use of such. While the use of fully animated and interactive elements in games may be a standard, it is imperative to point out that the sheer amount of content that will be placed within this teaching game would render a seamless online play undoable due to performance issues. Along with technical problems, the necessary manpower and labor needed to create such a game is beyond the scope of the research project, and so certain simplifications will need to take place in order for the game to be delivered on time. For example, the use of 3D models may be limited to the ability to host observable and roughly interactive models (move, rotate, zoom, take snapshot, comment). In other cases, video materials would be converted to specific cuts to reduce the file size, but with links to the full video made readily available via a third party video streaming service. Large text fields would also become reedited to better suit the game, with full text alternatives being made available if necessary in a separate window or module. The purpose of such drastic simplifications of highly complex and specialized content is to minimize the loss of interest within the player base but at the same time as a modernized method of conveying condensed information with the ability to reach for more information if needed.

Teaching materials would be layered so that the topmost firstly visible information would be understandable by most beginner-leveled users, but with the implementation of more and more sophisticated and detailed examples, would allow more intent students to have a tailored dose of information for their specific need and requirements [24]. Players who would reach for the more in-depth “layers” would be awarded through a point system for going beyond the minimum requirements, which in a given academic term would promote the less intent students to put more effort in the long run.

The scalability of this project is also one of the main aspects, as each partnering university would have to adapt their collected content and syllabus to best fit the still to be set example. In order to better prepare for the upcoming contextual setup, partnering universities were given an example technology tree layout based on a very simple procedure – baking challah bread. By writing out a relatively simple procedure such as a baking recipe, partnering universities would be able to relate to the topic and adapt their highly specialized courses into the mentioned tech tree system, as well as to foresee any complications further down the road as a result (Fig.11.2).

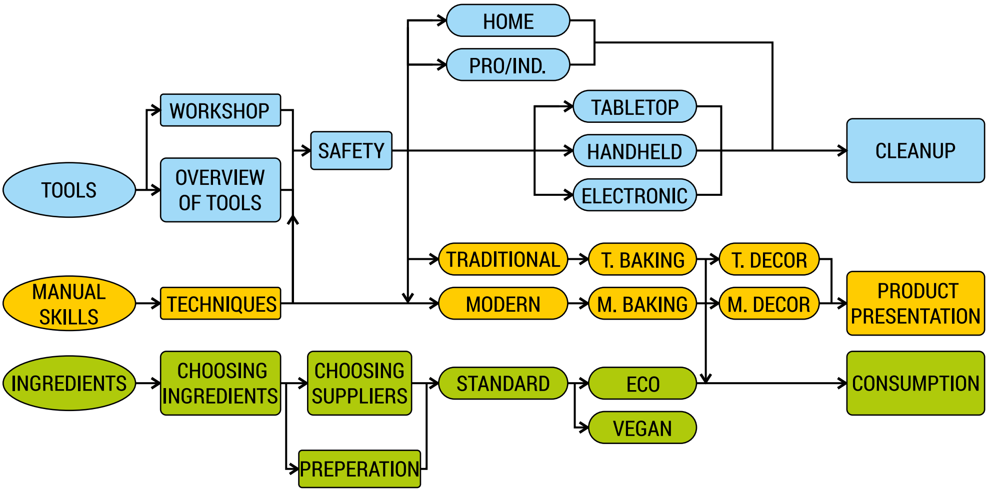


FIGURE 11.2. **Technology tree as an example based on the approach of making challah bread**

As seen above, the relatively simple process of following a recipe could be extended into a more complex learning experience with an in-depth curriculum as well as a basic insight into certain topics. This model could be adapted to any field of medicine as well as any college course in general, as it is made out of modular and easy to understand sections that can be of linear fashion or more complex dependencies between chosen sections.

In addition to the planning and coordination aspect of this educational game design for medical students, a peer-to-peer rivalry system is also introduced into the mainframe, as a motivational and integration aspect of gamification [25]. By using a point system that gratifies each player after completing a certain task, multi-level competitive interaction will take place during the span of the course. Local level competitiveness would take place within small groups in accordance to the standard university procedures, in which top students would be rewarded with such bonuses as more time on timed tasks, the ability to use notes during an exam, or anything the course moderators would deem necessary and acceptable. The next levels of competitiveness would bring forward the best of the local levels as well as reward those that excel at certain time events/tasks as an additional motivational factor. The last and most global level of competitive learning would bring out the best of the best and grant the highest levels of rewards as well as allow for special gratification by the hosting university. The promotion of excellence and positive reinforcement at several levels of the populace [25] are just some of the important features that the gamified system will host. The final product may vary slightly between consortium partners, as the system has lenience towards interpretation and modifications, but the framework should allow for seamless adaptation to any and all fields of research and teaching.

11.3. Summary

The existing studies on games and virtual patient simulations for biomedical engineering education measure domains of engagement: student satisfaction, flow (fun, enjoyment, and concentration), and variety. Well-designed games keep students engaged and facilitate progression through difficult tasks, specifically in different aspects: strategizing, collaborating, decision making, competing, evidence gathering, reviewing feedback, and reflecting. Real-World Application Games and virtual patient simulations may be designed to allow students to solve real-world problems. Contextualizing patient-case practice allows students to safely apply medical theory to a specific instance, sometimes mediated by a mentor. While students benefit from deliberate practice in risk-free environments, educators benefit from the analytics (scoring systems, statistical reports) offered by games and virtual patient simulation platforms.

The presented works are part of an ongoing process of implementation, trial and error, as well as refining predicted outcomes of higher engagement in the given curriculum. Having conceptualized and created the basic mainframe, the authors will proceed with further development and use in the upcoming semesters with bio-engineering students from partner universities. If applied appropriately, the results should favor the use of gamified means of the teaching methods and will provide a basepoint of ongoing development of the teaching tools and a practical approach to the evolving interactive means of holding classes and conveying knowledge.

With similar levels of importance, it should be noted that the above described skill tree approach is of the greatest importance in this paper, as the development of any and all further tools will be based on it. A gamified approach to the course syllabus will improve students' understanding of the upcoming sections, their relationship to other topics/sections of the course, as well as allow them to foresee a possible direction of growth for the span of a given semester. This in turn should grant students the ability to better prepare for the nearing topics and make it easier to understand the reason for such paths of progression.

All in all, the field of gamification for biomedical engineering education is innovative and dynamic. Gamified training platforms, mobile applications and virtual patient simulations can be used in biomedical curricula to real-world application, clinical decision-making, distance training, learning analytics and swift feedback. Rigorous research regarding the effectiveness of games, simulations, and mobile applications for health care learning is still in its infancy. Games and gamification elements should introduce fun and excitement in stressful environments. Well-designed games are cognitively challenging. They should keep students engaged and facilitate progression through difficult tasks. Games should support the need for adult learners to inventory and master short-term and long-term aims by breaking activities into a series of networking activities that are varied and interesting. These tasks should engage learners in different aspects of a serious play, such as strategizing, collaborating, decision-making, competing, evidence gathering, reviewing feedback, and reflecting.

Acknowledgments: The paper is supported by project „Technology for Healthcare education using smart gamification (HealthTEK)", No. 2021-1-ES01-KA220-HED-000035794 under the program Erasmus+.

References

1. Palmer D.: Lunceford S., Patton A.J., "The engagement economy: how gamification is reshaping businesses", Deloitte Rev., 11, 52–69, 2012.
2. Cugelman B.: "Gamification: what it is and why it matters to digital health behavior change developers", JMIR Serious Games, 1 (1), e3, 2013.
3. Haoran G., Bazakidi E., Zary N.: "Serious games in health professions education: Review of trends and learning efficacy", Yearb Med Inform., 28(1), 240–248, 2019.
4. Kapralos B., Fisher S., Clarkson J., et al.: "A course on serious game design and development using an online problem-based learning approach", Interactive Technol Smart Educ., 12(2), 116–136, 2015.
5. Petouhoff N.: "Social Media, Gamification And Gaming Statistics", February 8, www.drnatalienews.com/blog/social-media-gamification-and-gaming-statistics (accessed: 15.11.2015), 2012.
6. Koivisto J., Hamari J.: "Demographic differences in perceived benefits from gamification", Computers in Human Behavior, 35, 179–188, 2014.
7. Hamari J.: "Transforming homo economicus into homo ludens: A field experiment on gamification in a utilitarian peer-to-peer trading service", Electronic Commerce Research and Applications, 12, 236–245, 2013.
8. Hamari J.: "Do badges increase user activity? A field experiment on the effects of gamification", Computers in Human Behavior, 71, 469–478, 2017.
9. Katchuck M.: "Gamification in PR" [in:] Hutchins A., Tindall N. (eds.): Public Relations and Participatory Culture: Fandom, Social Media and Community Engagement, Routledge, London, 45–57, 2016.
10. Hainey T., Westera W., Thomas M., Connolly T., Boyle L., Baxter G., Beeby R., Soflano M.: "Students' attitudes toward playing games and using games in education: comparing Scotland and the Netherlands", Computers & Education, 69, 474–484, 2013.

11. Jakubowski M.: "Gamification in business and education – project of gamified course for university students", Developments in Business Simulation and Experiential Learning, 41, 339–342, 2014.
12. Kuo M.S., Chuang T.Y.: "How gamification motivates visits and engagement for online academic dissemination – an empirical study", Computers in Human Behavior, 55, 16–27, 2016.
13. Barata G., Gama S., Jorge J., Gonçalves D.: "Improving participation and learning with gamification", Gamification, Stratford, ON, Canada, 2013.
14. Burton B., Brothers L., Martin B.: "An Examination of the outcomes of gamification when used in a university level classroom", American Education Research Association (AERA), 2013.
15. Simões J., Redondo R., Vilas A.: "Using flow as a measurement of students engagement in a gamified social learning environment", Proceedings of ICERI, Seville, Spain, 2015.
16. Seixas L., Gomes A., Filho I.: "Effectiveness of gamification in the engagement of students", Computers in Human Behavior, 58, 48–63, 2016.
17. Chitwood W.R. et al.: "Robotic mitral valve repair: trapezoidal resection and prosthetic annuloplasty with the da Vinci surgical system", The Journal of thoracic and cardiovascular surgery 120.6 (2000): 1171–1172.
18. Short A.H., and Tomlinson D.R.: "How to…: Use games in medical education", Medical Teacher 2.3 (1980): 116–117.
19. Durand T.: "Dual technological trees: Assessing the intensity and strategic significance of technological change." Research policy 21.4 (1992): 361–380.
20. Kaur D.P., Vijayant S., and Jatin S.: "Surveying the role of Gamified MOOCs, Augmented and Virtual Reality for Engineering Education" (1989).
21. Mago, Z.: "Easter eggs in digital games as a form of textual transcendence (case study)", Acta Ludologica 2.2 (2019): 48–57.
22. Gajadhar BJ., De Kort Y.A.W., and IJsselsteijn W.A.: "Rules of engagement: Influence of co-player presence on player involvement in digital games", International Journal of Gaming and Computer-Mediated Simulations (IJGCMS) 1.3 (2009): 14–27.
23. Stasieńko J.: "Why are they so boring? – the educational context of computer games from a design and a research perspective" (2013).
24. Liu Z., Wei L., and Gao X.: "A study on self-regulated micro-course learning and implicitly layered flipped classroom", Theory and Practice in Language Studies 6.4 (2016): 870.
25. Barros P. et al.: "Incorporating rivalry in reinforcement learning for a competitive game", arXiv preprint arXiv:2011.01337 (2020).

DOI: 10.24427/978-83-67185-49-3_12

Chapter 12
Modification of PLA with nano-iron particles for biomedical applications

Hubert Wiński, Dorota Trusiewicz, Zbigniew Oksiuta

Bialystok University of Technology, Institute of Biomedical Engineering,
e-mail: hubertwinski99@gmail.com, dorkas99@wp.pl, z.oksiuta@pb.edu.pl

Abstract: The current problem of biopolymer composites in biomaterial engineering is their long degradation over time under environmental interactions and their potential application in biomedical engineering. The present research work aims to study the properties of a PLA composite with addition of iron and nano-iron powders with a particle size of 45 μm and 80 nm, respectively, in an isotonic solution in the Ringer fluid. The paper describes the general properties of polylactide composites and the mechanisms of degradation after modification with different additives, among them iron powders and polyethylene glycol (PEG)-based blends, where the PLA-nanoFe-PEG liquid blend has the best mechanical properties. The areas of biomedical applications of the PLA-based composite were also presented. In the research part of the PLA composite fabricated by the extrusion method, hardness, tensile strength and degradation in Ringer's solution of the obtained composites were studied and discussed.

Keywords: PLA, modifiers, iron powders, mechanical properties, composite, PEG blend

12.1. Introduction

The current use of polymers in medicine is very broad, ranging from medical devices, such as spatulas, specula, syringes through drains, surgical threads to artificial heart valves or pharmacological agents introducing drugs and blood substitutes into the body [1, 2]. All these devices are characterized by one important parameter, i.e. persistence in the human body. Invasive medical devices must meet numerous requirements, such as biocompatibility, biofunctionality and biotolerance in the environment in which they work, and especially their chemical composition, which, above all, should not harm the patient. The best solution is to create a material that meets the above mentioned aspects and has the appropriate mechanical properties, such as increased ductility and higher tensile and impact strength in comparison to the pure PLA.

The present work is devoted to the investigation of PLA-based composites with iron powders of various particle sizes (micro- and nano-particles), with the addition of PEG blends, in order to obtain a biodegradable composite with better ductility in comparison to pure PLA, with the possibility of controlling time of its degradation. The composite will be obtained by extrusion followed by the pressing process. This research work also aims to reveal the influence of additives on selected physicochemical properties of the obtained composites.

12.1.1. Characteristics of PLA

Polylactide or Poly(lactic acid) is a biodegradable and biocompatible aliphatic polymer that is produced from starch. It consists of an L- or D-twisted lactic acid, from which the polymer is obtained by using a synthetic method [3]. Starch found in corn or sugarcane potatoes is used to produce PLA. This polymer has an ability to crystallize by stretching, temperature crystallization, and its impact strength can be modified, copolymerized and processed using a variety of processing methods [4].

PLA has numerous advantages, such as biodegradability, low melting point, low polymerization shrinkage, easy doping, as well as some disadvantages which include brittleness, hydrophilicity, low flexibility and high price. Pure PLA is transparent and thermoplastic and therefore easy to process [5, 6].

The properties of PLA depend on the stoichiometric composition of the repeating units and their distribution along the chain. The melting point (Tt) of PLA is in the range 150÷180°C [4]. The atactic polylactide is made of heterochiral PLDLA units, which is an amorphous polymer with a glass transition temperature (Tg) of 59°C and no distinct melting point. A mixture of homochiral and isotactic PLLA and PDLA chains form stereocomplexes with a melting point of about 230°C [4].

The properties of PLA are similar to polystyrene, but after modification with additives its properties are like polypropylene and polyethylene. It is worth noting that good physical and rheological properties of PLA can be optimized by combining with modifiers, such as fillers or fibers. Prominent among them are polysaccharides such as starch, which significantly reduces the cost of PLA production and shortens the biological degradation time [4].

12.1.2. Mechanisms of PLA degradation

The degradation of a polymeric material is a phenomenon whereby the structure of a molecule is broken down into smaller equivalents and elements. The presence of hydroxyl groups accelerates the kinetics of the hydrolysis reaction through the absorption of water into the material. This process depends on the origin of the polymer and its physicochemical properties, such as chemical composition, degree of crystallinity, surface area, molecular weight and the nature of the polymer (e.g., hydrophilic) [7].

There are three types of degradation of biodegradable polymers: photodegradation, hydrolysis and microbial degradation. In the context of PLA application on the patient's body, only two of these degradations should be considered: hydrolysis, due to the tissue environment in which the material can be placed, and microbial degradation, due to the presence of micro- and macrophages in the body.

Hydrolytic degradation is characterized by the breaking of hydrolytic bonds of PLA with the production of water-soluble by-products. Polymer degradation in this case intensifies hydrophilicity, which is initiated from the inside towards the outer surface layer, causing a slight dimensional change in the material. On the other hand, progressive surface erosion, which causes a loss of mass from the outer surface of the polymer towards its core, does not affect the mechanical and structural integrity of the polymer. Hydrolysis, on the other hand, involves the dissociation of the bonds of the polymer backbone in the presence of water, in which molecules that have been converted to water-soluble monomeric acids by bulk or surface erosion are dissociated. During bulk erosion, water molecules diffuse very rapidly into the amorphous regions of the polymer, causing loss of strength and changes in structure. The degree of hydrolysis is influenced by the temperature and pH of the solution, morphology, degree of crystallinity and porosity of the tested material [8].

Microbial degradation can be caused by microorganisms which can change the strength and color of the polymer through compounds (enzymes) secreted as by-products during the microbial growth process. These enzymes are secreted outside the cell membrane to extract energy and nutrients from the polymer and catalyze the circulation of organic matter. During this degradation, there is a reduction in the size of the polymer, which is dissolved in water and ready to be transported across the cell wall. Then, it is mineralized by microbial metabolic pathways involving chemical transformation. During depolymerization, fragments of the biopolymer are consumed as carbon substrates forming various products such as carbon dioxide, methane, hydrogen, ammonia, sulfur dioxide, water, or salts and minerals. Degradation proceeds faster under aerobics than under anaerobic conditions [8, 12].

12.1.3. Areas of application for PLA-based composites

PLA has degradable properties which allow it to be used in medicine to create composites with the main purpose to disintegrate (biodegrade) after some time of usage. The resulting composite products based on PLA make it possible to reduce the intervention of invasive surgery. As an example, the composite can be used to produce bioresorbable surgical sutures, or "olives" for intramedullary stabilization of humerus fractures with a nail, so that the surgeon does not have to remove them again, as is the case with traditional treatments [1].

The second type of PLA-based medical devices are implants, plates and screws used as a bond for damaged bones. They allow for limiting the intervention of a surgeon to reopen the patient's body to remove the device after the medical device has completed its function [9].

In tissue engineering, porous PLA can provide a scaffold for cell culture in which well-defined cells proliferate for subsequent implantation into traumatic sites. Another widely used PLA product is pharmaceuticals, which exploit the material's biodegradability property for time-controlled drug release to prevent side effects and prolong drug duration [3].

A number of different additives can be introduced into PLA to influence the properties of the polymer. Some additives are added to reduce the cost of production of final products, others to improve or adjust the mechanical properties for specific requirements.

Literature reports also show that the introduction of PEG into PLA matrix improves ductility and impact properties, but at the same time reduces tensile strength and tensile modulus [10].

Also, these are the very PLA compositions with additives of ceramics, polymers as well as other organic substances. Few reports only deal with the addition of metallic elements such as magnesium, titanium, manganese and 316L steel [13, 14]. The addition of metallic compounds, such as iron, provides higher stiffness of the composite, an increase in Young's modulus with increasing metal content values, and a higher coefficient of friction is exhibited by the metal-reinforced PLA composite. Preliminary studies on metallic additives to PLA [14] indicate that among several selected elements, such as Mg, Fe or Ti, the PLA composite with Fe powder addition has better mechanical properties, especially ductility, compared to pure PLA, while tensile strength results were reduced by 20%. Therefore, in this study, a PEG compatibilizer was used to improve the mechanical properties of the PLA+Fe+PEG composite. It is also worth noting that compared to traditional biomaterials, such as stainless steel, titanium alloys and cobalt alloys, pure iron is more reactive and faster biodegradable in human body fluids and thus has a low risk of adverse complications and can be more easily absorbed by the human body. It also promotes the formation of iron-containing proteins. In tissue engineering, the PLA-iron composite shows a controlled degradation rate and a beneficial tissue regeneration function [11].

12.2. Test materials and methods

PLA 2003D in granular form from the Nature Works (USA) was used for the study. Three composite materials were prepared with the addition of iron powder (average particle size 45 m) and nano-iron (average particle size 80 nm) at 10 and 1%

(wt.%), respectively. To improve bonding between the PLA and iron powder particles, PEG in a powder or liquid form was added at 1% and 20% (in wt.%). Literature data show that the addition of this blend has a beneficial effect on improving the ductility of the obtained material compositions [10].

Subsequently, several tests were carried out on the materials in order to obtain information on their mechanical and physicochemical properties.

The nominal chemical composition of the tested composites is presented below:

- PLA + 1% nanoFe + 1% PEG (powder) – molecular weight 200 (designated 1nFp),
- PLA + 1% nanoFe + 1% PEG (liquid) – molecular weight 600 (designated 2nFc),
- PLA + 10% Fe + 20% PEG (liquid) – molecular weight 600 (determined as 3F20c).

Stability studies in an isotonic solution of 0.9% NaCl were carried out for 10×10×2 mm samples in a greenhouse at a constant temperature of 37±1°C. Three samples from each batch were prepared for testing.

Prior to testing, the materials were measured and weighed on a laboratory scales (Radwag 1000). Iron nanopowder was added to the PLA in a glovebox under argon shielding. The ingredients were then mixed by hand in a mortar for 15 minutes. Each sample (of 100 g each) was extruded and additionally pressed using a PXW 100 hydraulic press, applying a unit pressure of 20 MPa.

The extrusion process parameters are presented below:

- temperature in the charging zone – 50°C,
- feeding zone – 170°C,
- compression zone – 180°C,
- dosing zone – 180°C,
- extruder head – 175°C,
- extrusion speed 5% of maximum machine speed 10 rpm.

Samples of equal size were placed in 0.9% NaCl solution to determine their changes in weight, pH and electrical conductivity of the saline, as a function of time. The tests were conducted for 90 days at 7-day intervals to gather more data. In addition, hardness tests were carried out using a Shore scale D apparatus, tensile strength according to EN-PN-PN 868:2005 using an Instron 2620-601 extensometer and observations of the microstructure and fractures of the samples were made using confocal microscopy and scanning electron microscopy (SEM) to examine as accurately as possible the structure of the material before and after the degradation test.

12.3. Research findings

12.3.1. Results of hardness testing

The results of Shore's hardness tests of the composites before and after 90 days storage in 0.9% NaCl solution at 37±1°C are shown in Table 12.1.

TABLE 12.1. Comparison of initial and final hardness values of PLA composite samples

	As received			After 90 days		
Sample	1nFp	2nFc	3F20c	1nFp	2nFc	3F20c
Average	58.2	65.6	67.6	52.1	61.1	62.5
SD	3.9	1.5	4.1	5.2	1.6	2.5
CV	0.07	0.02	0.06	0.1	0.03	0.04

The hardest sample before testing in saline (as received) turned out to be the PLA+1% nanoFe+1% PEG (liquid) and the lowest hardness has the PLA+1%nanoFe+1% PEG (powder) composition. After 90 days of testing, the hardness of all the samples decreased, while the trend remained unchanged. It is noteworthy that PLA+1%nanoFe+1% PEG (powder) sample were partially degraded and it was difficult to obtain reliable test results, as indicated by the highest value of the coefficient of variation (see Table 12.1).

It is noteworthy that the lowest hardness obtained for the PLA+1% nanoFe+1% PEG (powder), with addition of iron nanopowders and in which the PEG blend was in a powder form. For comparison, pure PLA obtained under the same process conditions has a Shore hardness of 68±2°ShD. The decrease in hardness of the tested samples was influenced by the addition of iron particles and iron nanopowders, as well as blends, especially in the form of powder, weakening the internal structure of PLA and increasing surface porosity.

12.3.2. pH tests in a physiological solution

The results of the pH solution as a function of the storage time of the samples are shown in Figure 12.1.

From the results presented in Fig. 12.1 it can be seen that the 1nFp samples, in the initial period up to 35 days, the pH of the solution decreases to a value of 3.0 mol/cm^3 and then remains unchanged until the end of the tests. It should be recalled that in this material the PEG additive was in the powder form. A similar course of changes in solution pH is observed in the case of the sample with micro-iron powder 3F20c, while an opposite trend is presented by the graph obtained for the 1nFc sample, where initially, up to 21 days, a slight increase in pH from 5.82 mol/cm^3 to the value of 6.27 mol/cm^3 is observed, followed by a decrease to the value of about 5.60 mol/cm^3 after 35 days of testing, and a steady increase of this parameter again.

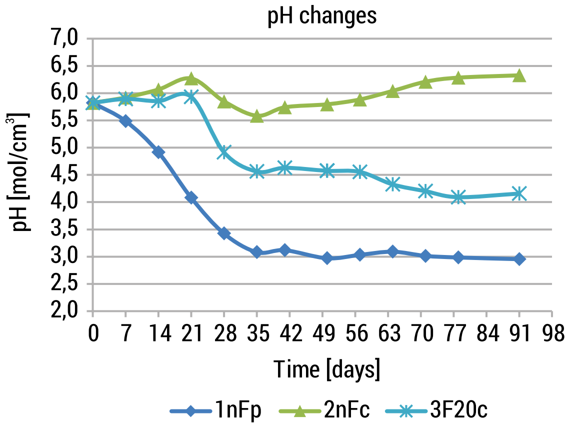


FIGURE 12.1. pH changes of the composites as a function of time

The highest pH value was obtained for PLA+1%nanoFe+1%PEG (liquid) samples (acid reaction), and its value was 6.33 mol/cm^3. The PLA+1%nanoFe+1%PEG (powder) sample had the lowest pH value, and its value was 2.89 mol/cm^3. It is noteworthy that none of the samples obtained a pH of alkaline. The observed pH trends are to some extent correlated with the mass loss (see section 12.3.3) due to hydrolytic degradation and the consequent release of H+ ions into solution, which are acidic in nature.

12.3.3. Studies on the mass changes of the samples after conditioning

The results of the change in mass of the samples after conditioning in the saline solution are presented in Fig. 12.2.

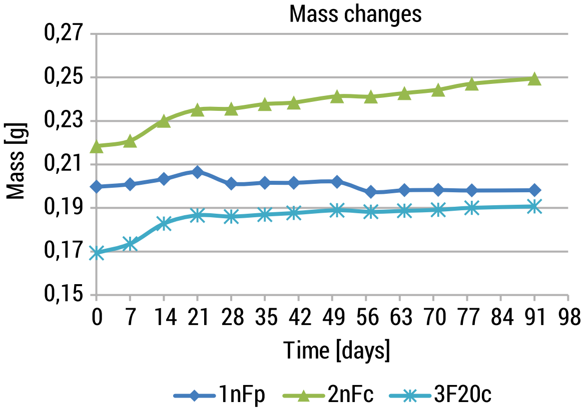


FIGURE 12.2. Plots of the specimens mass variation with time of conditioning in the 0.9% NaCl solution

The presented results indicate a continuous increase of the mass of two tested samples, with the exception of the 1nFp, where the mass of which slightly fluctuating during the tests. The highest increase in the mass can be observed in the case of the sample 2nFc. The continuous increase in mass content can be explained by absorption of the solution into the samples.

At the end of the test, the mass of sample 1nFp increased by 1.0% of its initial mass, whereas the sample 2nFc by 18.9% of its initial mass and the sample 3F20c by 13.5% of its initial mass. The obtained results demonstrate the effect of PEG on the mechanism of absorption of 0.9% NaCl solution by PLA, and it seems that less important in this case is the % of blend, but rather its molecular weight, of the powder and liquid state. It should be recalled that of powder PEG blend is 3 times lower in comparison to the liquid form, 200 and 600, respectively. Hence, it can be concluded that the addition of PEG to PLA in the form of powder effectively inhibits the sample weight changes of such a composition during conditioning in physiological fluid.

12.3.4. Electrical conductivity tests of samples in brine

The results of the electrical conductivity of the 0.9% NaCl solution are shown in Fig. 12.3.

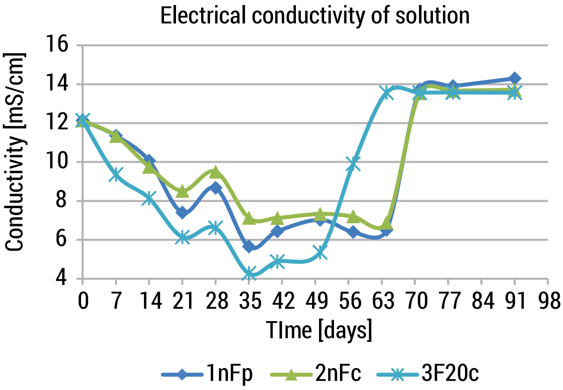


FIGURE 12.3. Plots of electrical conductivity vs time of the tested composite samples

On the basis of the presented results, it can be concluded that all the samples in the initial period of testing, up to about 35 days, reduce the electrical conductivity parameter. The value of this parameter is variable and ranges from 4.08 to 7.29 [mS/cm]. After another 14 days (49 days of the testing) an increase in conductivity was observed for the sample 3F20c, and after 63 days of conditioning, there was a sudden increase of this parameter to the value of 14.00 [mS/cm].

A more detailed analysis of the obtained results shows that between the 21st and 28th day, all the samples showed an increase in electrical conductivity; however, further measurements 1 week later showed that the value of electrical conductivity decreased.

The PLA+1%nanoFe+PEG (powder) sample has the highest electrical conductivity of 14.41 mS/cm. As the 1nFp and 2nFc composites differ from the 3F20c20 sample by a smaller amount of the PEG content and a smaller amount of nano-iron powder, in relation to the higher content of iron particles (10 wt.%) in this composite. It should be concluded that the decrease in electrical conductivity in the initial period of the research up to 35 days is due to the dissolution of PLA in saline. Thus, the higher the proportion of polylactide particles in the liquid volume, the lower the electrical conductivity. After a stagnation period of about 4 weeks, the increase in electrical conductivity should be attributed to the appearance of iron ions from the exposed metal particles contacted with 0.9% NaCl water solution, after the dissolution of the top layer of the polymer.

12.3.5. Tensile testing of PLA composites

Tensile plots of the tested composite specimens after manufacturing are shown in Fig. 12.4. Also, the results of the tensile testing are summarized in Table 12.2. For comparison, Table 12.2 shows the tensile results obtained for 100% PLA, produced under the same conditions as the rest of composite specimens.

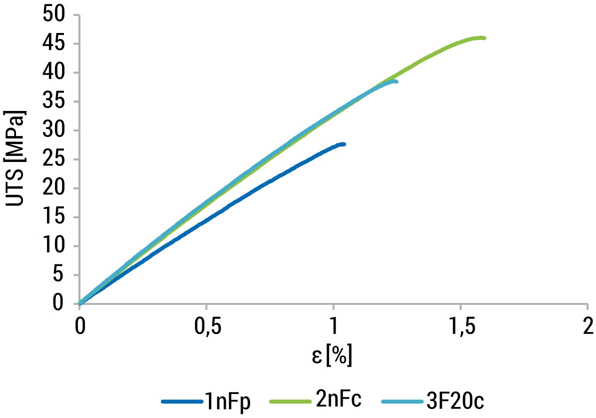


FIGURE 12.4. Typical tensile strength curves of three tested composite specimens

TABLE 12.2. Summarized averaged tensile properties of the PLA based specimens

Sample	ε [%]	UTS [MPa]	E [GPa]
1nFp	0.99±0.02	27.7±3.2	2.78
2nFc	1.51±0.05	45.8±3.5	3.03
$3F_{20}c$	1.22±0.03	39.6±2.9	3.25
Pure PLA	1.40±0.03	47.5±2.5	3.48

The worst tensile properties has the 1nFp sample, whereas the 2nFc material has the highest tensile strength and elongation. The highest Young's modulus is observed in the case of 3F20c material and it was 3.25 [GPa], and the lowest E parameter is measured for the 1nFp sample, and it was 2.78 [GPa]. Comparing the obtained data with the pure PLA, in the case of PLA+1%nanoFe+1%PEG (liquid) composite, the strain value is about 0.1% higher than that of pure PLA, while the stress value decreased about 1.7 MPa (about 3.6%). In the PLA+1%nanoFe+1% PEG (powder) composite, the strain was close to 1% and the stress was 27.7 MPa, which is certainly lower (about 42%) than that of the pure PLA. The differences in strain may be due to the density of the composite as well as the type and valence fraction of additives.

The UTS value correlates well with the hardness data measured for the as received composites (see Table 12.1). In addition, the PLA composite with the PEG additive in the powder form (sample 1nFp), probably due to insufficient homogenization during the extrusion process, has the weakest mechanical properties. In contrast, the material with PEG in the form of a liquid (PLA+1%nanoFe+1%PEG liquid), has comparable UTS properties to pure PLA, while the plasticity of this composite is slightly better (about 8%) compared to the pure polymer. Therefore, further studies should be carried out to optimize the chemical composition of the composite with the addition of iron nanopowder using PEG liquid blends.

12.4. Discussion

The obtained results allowed to determine the characteristic features of the produced composites based on PLA as a degradable polymer matrix and modifying additives in the form of iron powders and nano-iron as well as the PEG with different molecular weight with the possibility of their application in medicine.

Analyzing the results obtained for mechanical properties of the samples achieved immediately after their manufacturing (extrusion and axial pressing), it should be tated that the most favorable properties, both hardness, UTS and the E modulus were obtained for the PLA+1%nanoFe+1%PEG (liquid) composite, while much weaker mechanical properties were found for the PLA+1%nanoFe+1%PEG (powder) sample, where the blend additive was in the form of powder. A correlation was found between the hardness and tensile strength results. The hardness and mechanical properties of the composites are affected by the addition of PEG, which in the form of a powder was difficult to homogenize with PLA, which in turn probably caused delamination and weakening of the properties. Thus, it can be concluded that PEG in different forms (liquid vs. powder) added to PLA can affect the mechanical properties of the composites.

However, in the case of the E modulus, the composite with 10% of iron powder PLA+10%Fe+20%PEG (liquid) obtained the highest value. This fact can be explained by the highest volume fraction of additives, both iron and blend: 10% Fe and 20% PEG. In the case of the material composed of PLA+1%nanoFe+1%PEG (powder), the value of the E modulus is the smallest.

It was also found that the pH of the solution with the samples of the composites used for the study, carried out for a period of 90 days, is in correlation with the hardness and tensile strength tests. Despite the lowest tensile properties obtained for the PLA+1%nanoFe+1%PEG material (powder) sample, the lowest pH value was measured that indicates its stability in saline.

In contrast, the studies on electrical conductivity showed that the electrical conductivity curves of all the samples was similar and comparable, which suggests that the mechanism of iron and PLA ions release into solution is similar.

12.5. Conclusions

Based on the results presented here, PLA+1%nanoFe+1%PEG with the addition of a liquid blend proved to be the best material, as it has one of the highest hardness levels and the best tensile properties. This composite material has also a pH closest to the neutral, so no acidification or inflammation should occur in the tissues. The material is also characterized by stable electrical conductivity, and negligible mass changes during the testing in the saline solution conducted up to 90 days.

The results of the electrical conductivity show that regardless of whether the PEG blends are added, in a powder or liquid form, the composites have similar electrical conductivity, while, the PEG form related to their molecular weight strongly affects mechanical properties.

The PLA+1%nanoFe+1%PEG (powder) composite has the highest mass gain than the other composites, which indicates an increased hydrophilicity of this material due to porosity caused by the inferior homogeneity of the composite sample.

Also noteworthy is the fact that the composite with the addition of the PEG blend in the form of a liquid, practically does not change the mass of sample during the whole process of incubation in 0.9% NaCl water solution. This feature can be used when designing the resorbable properties of the PLA-based implant as a function of time.

This research will be continued to optimize the chemical composition of PLA-based composites. It seems that polylactide-based composites with the addition of metal powders are promising materials in terms of their potential for medical applications with controlled degradation rates.

Reference

1. Janicki S. et al.: Farmacja Stosowana, Wydawnictwo Lekarskie PZWL, IV, Warszawa, 2002.
2. Ruśkowski P., Gadomska-Gajadhur A.: “Polilaktyd w zastosowaniach medycznych”, Tworzywa sztuczne w przemyśle, 2, 32–35, 2017.
3. Chandika P. et al.: “Recent advances in biological macro molecule based tissue-engineered composite scaffolds for cardiac tissue regeneration applications”, International Journal of Biological Macromolecules, 164, 1, 2329–2357, 2020.
4. https://www.tworzywa.pl/wiedzopedia/baza-tworzyw/121,polilaktyd-pla,polimer.html [accessed: 20.05.2022].
5. Jun C.L.: “Reactive blending of biodegradable polymers: PLA and starch”, Journal of Polymers and the Environment, 8(1), 33–37, 2000.
6. Liang H. et al.: “Thermal, rheological, and mechanical properties of polylactide/poly (diethylene glycoladipate)”, Polymer Bulletin, 70(12), 3487–3500, 2013.
7. Kabir E. et al.: “Prospects of biopolymer technology as alternative option for non-degradable plastics and sustainable management of plastic wastes”, Journal of Cleaner Production, 258, 120536, 2020.
8. Nowak B., Pająk J.: “Biodegradacja polilaktydu (PLA)”, Archiwum Gospodarki Odpadami i Ochrony Środowiska, 2, 1–10, 2010.
9. Burczyk B.: Biomasa: surowiec do syntez chemicznych i produkcji paliw, Oficyna Wydawnicza Politechniki Wrocławskiej, 2011.
10. Mohapatra, AK., Mohanty S., Nayak S.K.: “Properties and characterization of biodegradable poly (lactic acid)(PLA)/poly (ethylene glycol)(PEG) and PLA/PEG/organoclay: A study of crystallization kinetics, rheology, and compostability”, Journal of Thermoplastic Composite Materials, 29(4), 443–463, 2016.
11. Jiang D., Ning F., Wang Y.: “Additive manufacturing of biodegradable iron-based particle reinforced polylactic acid composite scaffolds for tissue engineering”, Journal of Materials Processing Technology, 289, 116952, 2021.
12. Oliveira M. et al.: “The role of shear and stabilizer on PLA degradation.” Polymer Testing, 51, 109–116, 2016.
13. Kuentz L. et al.: “Additive manufacturing and characterization of polylactic acid (PLA) composites containing metal reinforcements”, [in:] International Conference and Expo on Advanced Ceramics and Composites (No. GRC-E-DAA-TN29146), 2016.
14. Oksiuta Z. et al.: “Mechanical and thermal properties of polylactide (PLA) composites modified with Mg, Fe, and polyethylene (PE) additives”, Polymers, 12(12), 2939, 2020.

DOI: 10.24427/978-83-67185-49-3_13

Chapter 13
Modified titanium dioxide for 3D printing applications

Bogna Sztorch, Julia Głowacka**, Anna Hanuszewska**,*
*Eliza Romańczuk-Ruszuk***, Roksana Konieczna** Robert Przekop**

**Adam Mickiewicz University in Poznań, Centre for Advanced Technologies,*
e-mail: bogna.sztorch@amu.edu.pl, rprzekop@amu.edu.pl;
***Adam Mickiewicz University in Poznań, Faculty of Chemistry*
****Bialystok University of Technology, Institute of Biomedical Engineering*

Abstract: This chapter describes the methodology of obtaining polylactide (PLA) biocomposites with the TiO2 modified by organosilicon compounds: disiloxane and functionalized octaspherosilicate designed for use in additive technologies. The paper presents general information about polylactide, titanium dioxide and the used organosilicon modifiers (silsesqioxanes and polysiloxanes). In the paper the authors compare the effect of two types of TiO2 from different manufacturers on the properties of polymer biocomposites, which can be used in biomedical applications. The properties of the obtained composites with the use of additive FDM technology were characterized mechanically (tensile strength, flexural strength, impact strength), and the particle size of the TiO2 pigment nanofillers used was measured. All the composites were pretreated using a plastic processing technique with two different mass ratios of TiO2 filler content (1.5%; 0.5%). The properties of PLA/TiO2 composites depend on the filler concentration. The applied fillers had a positive effect on the improvement of the mechanical properties of the produced polymer composites. This effect was additionally intensified by the introduction of organosilicon modifiers.

Keywords: biocomposites, PLA, titanium dioxide, pigments, organosilicon modifiers, functional nanocomposites

13.1. Introduction

Additive techniques commonly known as 3D printing are a rapidly developing field of new technologies. The history of printing begins in the 1986, when Charles Hull patented the stereolithographic apparatus (SLA). Another patent appeared three years later. It concerned the selective laser sintering (SLS) process and belonged to Carl-Deckard. In the same year, i.e. 1989, a patent for Fused Deposition Modeling (FDM) was filed, for the publication of which Scott Crump waited three years (*History of 3D printing:*

The Free Beginner's Guide, 2014) [1]. Additive technologies are based on the production of three-dimensional objects by selectively adding successive layers of material. The process is computer-controlled on the basis of a given three-dimensional digital model, making it possible to manipulate the geometry of the object [2].

They have found their application in almost every industry branch. They are used, e.g. in accordance with the original assumption for rapid prototyping, limiting the use of specialized equipment [3]. The automotive, architectural, military, construction, fashion and aviation sectors, and even medical and prosthetics, also use 3D printing [2]. 3D printing, despite the many applications mentioned, has drawbacks that need to be eliminated. One of the weaknesses of 3D printing is still the lack of suitable materials to create more functional prototypes with a longer useful life. In this work, efforts were made to modify the properties of polylactide to make it an even better material for 3D printing.

13.1.1. FDM (Fused Deposition Modeling)

The most popular additive techniques include SLA (Stereolithography), SLS (Selective Laser Sintering) as well as the most profitable for the 3D printing market and the most commonly used FDM technique. The FDM method consists in extruding a linear thermoplastic material melted in the head and depositing it on the platform (Fig. 13.1).

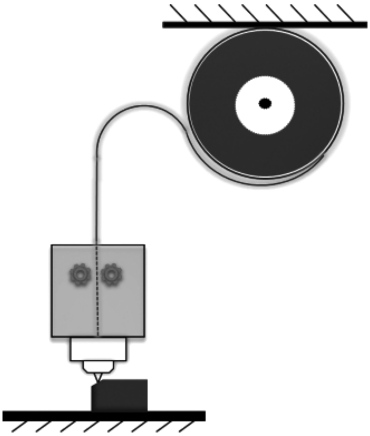


FIGURE 13.1. Scheme of FDM technique

The nozzle moves in the X-Y plane, allowing for various shapes. The growth of the model in the Z direction is possible by depositing successive layers of plasticized material to the layers located below them. The presence of the second nozzle is necessary when support layers are used, i.e. slides that allow model objects to be given complex shapes, the supports are removed after the printing process. The main advantages of this method are: ease of printing with thermoplastics, lower cost and less waste than in the case of other techniques, and its simplicity. The disadvantages of the FDM method are the relatively long time to print larger parts and often the need for post-processing [3].

In the FDM technique, thermoplastics, such as ABS (acrylonitrile-butadiene-styrene copolymer), HIPS (high-impact polystyrene), PET-G (glycol-modified poly (ethylene terephthalate), TPU (thermoplastic polyurethane), nylon (polyamide) are the most popular due to the ease of processing and the environmental impact of PLA (polylactide) [4].

13.2. Polylactide – synthesis, properties, modification methods

13.2.1. Chemical structure of PLA

Polylactic acid (PLA) is an aliphatic linear polyester of natural origin, composed of lactic acid monomers and obtained from raw materials of natural resources, such as corn or sugar beet. Lactic acid (2-hydroxypropanoic acid, $HO(CH_3)CHCOOH$, LAc) is a weak, water-soluble organic acid. It is the simplest α-hydroxy acid with an asymmetric carbon atom in the molecule. It can exist as two enantiomers, *D*(–)-LAc or *L*(+)-LAc. Due to the fact that there are two stereogenic forms of lactic acid *D* and *L*, cyclic diesters can exist in the form of three different stereogenic forms: *D,D*-lactide, *L,L*-lactide and *D,L*-lactide [5]. There are two reactive functional groups, hydroxyl and carboxyl, in the structure of the lactic acid molecule. Both can undergo intermolecular and intramolecular esterification reactions. Intermolecular esterification is the condensation of carboxylic acids with nucleophilic hydroxyl forms and occurs between molecules. On the other hand, the intramolecular esterification reaction takes place in the structures from dimeric upwards. It occurs for large molecules and involves the protonation of the carbonyl, followed by a nucleophilic attack within the molecule and the transfer of the proton to the hydroxyl group. The water molecule is removed and the deprotonation step takes place. The stereochemical structure of PLA due to the presence of two stereoisomeric forms – *D* and *L*, can be easily modified thanks to the controlled polymerization reaction of isomers [6].

13.2.2. Synthesis of PLA

Polylactide can be obtained directly from lactic acid or from its lactide. The synthesis of PLA from lactic acid is the most common as a result of direct or azeotropic condensation [7]. The production of PLA from lactide is carried out by ring-opening polymerization (ROP, Figure 13.2). In this way, a polymer with the desired molecular weight can be obtained in a continuous process. On the other hand, the direct polycondensation method is used to obtain a polymer with a higher molecular weight [8].

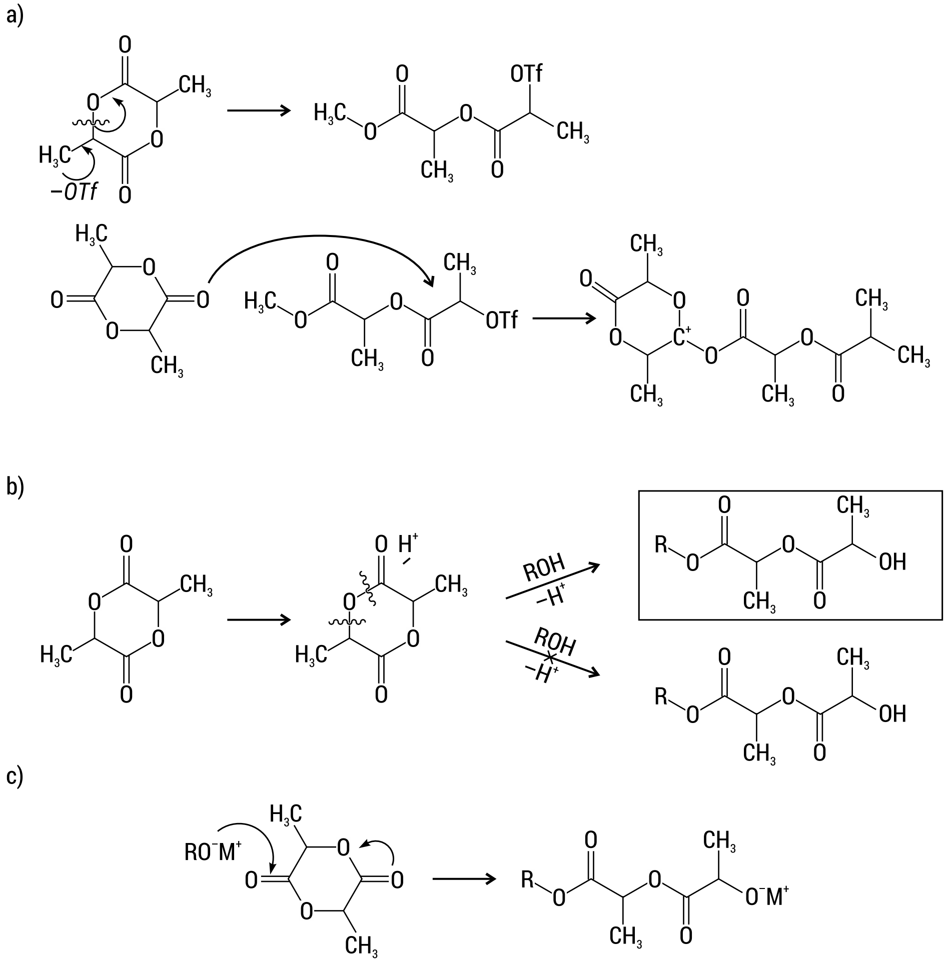


FIGURE 13.2. ROP schemes: a – cationic polymerization, b – cationic polymerization with activated monomer, c – anionic polymerization

13.2.3. Properties and application of PLA

Polylactide has many advantages, such as ease of processing, transparency, relatively good durability and tensile strength. PLA is degraded to the form of lactide [9]. Its biodegradability is a great advantage in times of increasing amount of residual nondegradable polymer waste. Contrary to most bioplastics, PLA shows good mechanical properties, easy processability in technological processes, as well as high transparency similar to PET and PS, which is why it is very attractive for many industries as an alternative material. On the other hand, its natural properties are not satisfactory, in particular in the so-called specialized applications. First of all, low impact strength (high brittleness) limits the application abilities of PLA [10]. The disadvantages of PLA (limitations) pose serious scientific challenges. For this reason, there is a need to increase the versatility of PLA in order to compete with petrochemicals. The impact of the negative properties of PLA on its use in many industries can be eliminated by the addition of appropriate modifiers, e.g. plasticizers and inorganic and fibrous nucleates or fillers, such as glass fibers or carbon fibers [11]. Until now, PLA has mainly found its application in the packaging, 3D printing and textile industry. There are also known cases of the use of polylactide-based copolymers in medicine for the production of bone scaffolds placed in the body during bone regeneration with the use of 3D bio-printing methods [12].

13.3. Titanium dioxide as a pigment filler

Fillers and pigments are added to polymers to modify their properties, e.g. color, strength, dimensional stability. The most frequently used substances, due to their low price, can also be an effective method of reducing the production costs of polymer composites, if their addition in large amounts does not cause a drastic decrease in thermal or mechanical properties.

Due to their origin, fillers can be classified as: synthetic (e.g. glass, carbon, graphite fibers), inorganic (e.g. talc, chalk, silica, TiO_2) belonging to the category of microfillers and organic (e.g. wood flour, cellulose fibers) [13]. Of the many described in the filler literature, titanium dioxide plays a special role in the industry.

Titanium dioxide (TiO_2) occurs in three polymorphs (rutile, brukite, anatase), which undergo the following phase changes under the influence of temperature: brukite-rutile (500–600°C) and anatase-rutile (850°C). TiO_2 is a typical n-type, band-gap semiconductor with the value of 3.0 eV for rutile and 3.2 eV for anatase, which translates into wavelengths of 386 nm and 413 nm, respectively, therefore it exhibits photocatalytic properties [14]. On an industrial scale, titanium dioxide is produced by two methods: sulphate and chloride. It is a non-toxic, biocompatible white powder with high thermal stability and good coverage; therefore it is a commonly used pigmentation additive used for bleaching plastics, e.g. in packages, foils, window frames. TiO_2 is also used in the paint and varnish industry, acting as a thickening agent, improving their strength, light reflectance and hardness. It is also used as an antibacterial additive to materials used in the medical industry for orthopedic and dental implants. In addition to its applications in the polymer industry, TiO_2 is widely used in cosmetics and in photocatalysis. Due to the high refractive index, it is also used in the form of thin foils, as anti-reflective coatings in optics and photovoltaics [15].

13.4. Organosilicon modifiers

13.4.1. Silsesquioxanes – structure and methods of obtaining

Polyhedral oligomeric silsesquioxanes are organosilicon compounds of the general formula $(RSiO_{3/2})_n$ (where R may be, e.g. an alkyl, aryl, hydrogen substituent). Silsesquioxanes are broad class modifiers of various structures: cage, incompletely closed cage, ladder, or with a disordered network structure [16]. The most popular and most frequently used in the processes of further functionalisation are cubic silsesquioxanes and spherosilicates. Their structure is shown in Fig. 13.3.

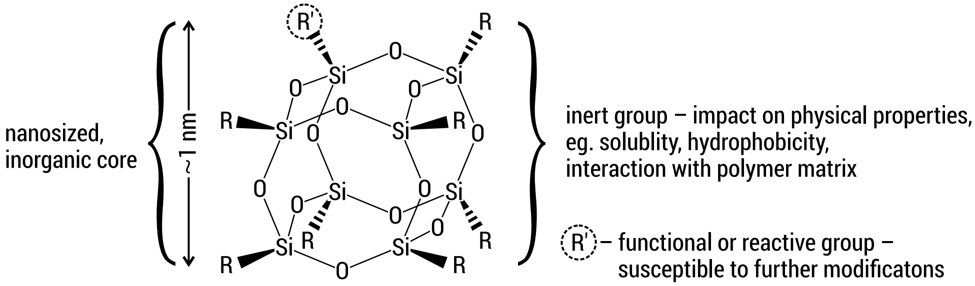


FIGURE 13.3. Scheme of silsesquioxane

Due to the possibility of modifying and controlling the mechanical and thermal properties, the use of these compounds in various industries has increased. Their advantages include the ability to modify the properties of polymers and the possibility of giving them new functional properties [17]. Silsesquioxanes can be used as functional additives influencing both the rheological and mechanical properties of polymers. By using various organic substituents and by selecting the appropriate processing method, the strength and rheological properties of the polymer can be improved. According to literature reports, the addition of silsesquioxanes also allows the viscosity of the polymer melt to be changed. In this case, the use of an appropriate concentration of the additive will be of key importance, as the rheology will change with the increase of the modifier content [18, 19]. Silsesquioxanes can occur as modifiers in the production of thermosetting polymers [20], thermoplastic polymers [21], including filaments used in 3D printing [22] and composite materials (improving both thermal and mechanical properties). As drug carriers they are used in pharmacy [23], but the medical [24] and dental sector [25] also use silsesquioxanes.

Octo-substituted cubic silsesquioxanes having a cage structure belong to the most widely studied group of silsesquioxanes. Most often they are obtained by the hydrolytic condensation reaction of trialkoxy or trichlorosilanes. Silicon atoms in their structure can be connected with almost any organic substituent, in the case of hydrogen-substituted silsesquioxanes it can be done by catalytic hydrolyzation reaction with appropriate olefins [16].

13.4.2. Polysiloxanes – preparation and application

Polysiloxanes, unlike cage spherosilicates, are organosilicon polymers with a long, flexible chain consisting of alternately arranged –Si–O–Si–O–Si–O– atoms and side groups attached to Si atoms [26]. The substituents on the silicon atom can be organic groups, such as: methyl, alkyl, cyclic structures (phenyl) or unsaturated groups (vinyl). The most widely studied polymer from this family is polydimethylsiloxane (PDMS), which contains methyl groups in its structure (they are substituents and are attached to the silicon atom).

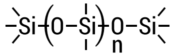


FIGURE 13.4. Chemical structure of PMDS

Polysiloxanes are characterized by a low Tg (–123°C), very high hydrophobicity, high thermal stability, low surface tension, low viscosity variation with temperature, high gas permeability and non-toxicity. These properties have made PDMS and related compounds with the most important inorganic polymers attractive from a commercial point of view [27].

Polysiloxanes are used in implantology, e.g. breast implants and facial modeling implants (e.g. tissue expanders, artificial finger joints), in drug delivery systems, contact lenses, thermal and electrical insulation, lubricants, adhesives, sealants and hydraulic fluids, motor oils, silicone rubbers, children's toys as well as in emulsions to make fabrics, papers and other materials waterproof [28].

13.5. Materials and method

PLA Ingeo2003D (NatureWorks, Minnetonka, MN, USA) has been modified with the use of pigment fillers in the form of two grades of TiO_2 from different manufacturers: TYTANPOL® R-001 (Grupa Azoty, Zakłady Chemiczne "Police" S.A.) and TIOXIDE® TR28 (Huntsman P&A Asia Sdn Bhd). TiO_2 was added to PLA in concentrations: 0.5% and 1.5%. Then, these systems were modified using organosilicon derivatives: octaspherosilicate (SS-5OD-3VTMOS) and polysiloxane (TMDS-2OD).

In order to prepare filled polymer systems, masterbatches with a concentration of 5% TiO_2 were prepared on a laboratory rolling mill (ZAMAK Mercator WG 150/280), then diluted to lower concentrations (1.5%; 0.5%). The TiO_2 filler and 1% organosilicon modifier relative to the TiO_2 mass were introduced into the polymer melt at 215°C: bis(octadecyl) tetramethyldisiloxane1,3,5,7,9,11,13,15-penta((octadecyl)dimethylsiloxy) tri ((trimethoxysilyl)dimethylsiloxy)pentacyclo[9.5.1.13,9.15,15.17,13]octasiloxane (SS-5OD-3TMOS) and stirred for 10 min.

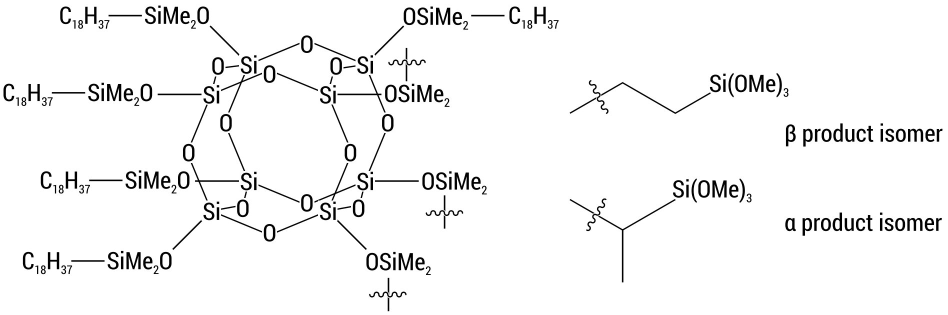


FIGURE 13.5. Chemical formula of (SS-5OD-3TMOS) – M1

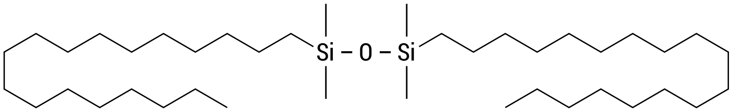


FIGURE 13.6. Chemical formula of TMDS-2OD – M2

In order to prepare the filaments for the 3D printing process, the dried masterbach (T = 60°C, 24h) was milled and diluted to concentrations of 0.5 and 1.5% on a Thermo Scientific HAAKE ™ extruder equipped with a single-screw module with six-zone temperature control. The drive system was controlled with the PolyLab ™ OS program. Extrusion was carried out under the following temperature conditions: TS1: 200°C; TS2: 205°C; TS3: 210°C; TSD1: 200°C.

Standardized test pieces for mechanical testing were printed on a Creatlity Ender 3 printer using the parameters summarized in Table 13.1.

TABLE 13.1. 3D printing parameters

Parameter	Value
Nozzle diameter	0.4 mm
Layer height	0.2 mm
Filling density	100%
Printing speed	60 mm/s
Printing temperature	210°C
Table temperature	60°C
Filament diameter	1.75 mm

The particle size (Dynamic Light Scattering – DLS) distribution of the fillers used to prepare the composites was measured using the MALVERN MASTERSIZER 300 device. The study involved dynamic light scattering and measurement of Brownian motion of macromolecules in solution. Measurements were made for samples in water suspension (Hydro EV attachment). Using DLS, the particle size distribution of both TiO_2: TR28 Huntsman and R001 Tytanpol was examined. The test samples were prepared by dissolving TiO_2 nanofillers in the surfactant. Before starting the measurement, the refractive angle index (T = 25°C) was measured using an INSMARK refractometer. After all the measurements were made, the test was performed again, this time with the use of ultrasound, in order to refine the results. The particle size distribution is determined by the Stokes-Einstein equation:

$$D_t = \frac{K_b T}{3\pi\eta\, d_H}. \tag{13.1}$$

where: d_H – hydrodynamic diameter [m], K_b – Boltzman's constant [J/K], T – temperature [K], η – solvent viscosity [kg/ m·s], D_t – diffusion coefficient [m^2/s].

To determine the mechanical properties of prepared composites, material bending and tensile tests were carried out in accordance to PN-EN ISO 527:2020 and PN-EN ISO 178:2019. Tests of the obtained specimens were performed on a universal testing machine INSTRON 5969 with a maximum load force of 50 kN. The traverse speed for tensile strength measurements was set at 2 mm/min, and for flexural strength was also set at 2 mm/min.

Charpy impact test (with no notch) was used to investigate the resistance of the materials to dynamic loads. Standardized 3D printed samples were tested on the Instron Ceast 9050 impact-machine according to PN-EN ISO 179:2020.

For all mechanical tests (bending, tensile, impact), the average value was determined for each series of specimen measurements. The results were compiled based on the corresponding norm references. The measurement error was calculated as the standard deviation from the mean. The measurement results are presented in the form of bar charts.

13.6. Results and discussion

13.6.1. The particle size (Dynamic Light Scattering – DLS) distribution of TiO_2 fillers used to prepare the composites

Prior to the preparation of biocomposite materials, DLS was investigated to assess the suitability of commercial TiO_2 grades for plastics processes. Table 13.2 presents DLS results in the form of Dv (10), Dv (50) and Dv (90) percentages, which are defined as:

- Dv (10) (µm), where 10% of the particle distribution is below this value,
- Dv (50) (µm) is the median, where 50% of the particle distribution is above this value and 50% below,
- Dv (90) (µm), where 90% of the particle distribution is below this value.

Analyzing these values, a significant difference in the distribution of Dv particles (90) can be noticed. Huntsman's TiO_2 TR28 tends to form agglomerates. It can disperse less, which makes it difficult to carry out processing. The difference between the particle size of TiO_2 R001 Tytanpol and Huntsman's TiO_2 TR28 is as high as 39.656 µm.

TABLE 13.2. DLS analysis

Type of TiO_2	Specific surface area	Dv (10)	Dv (50)	Dv (90)
TYTANPOL® R-001	27890 m²/kg	0.114 µm	0.268 µm	0.544 µm
TIOXIDE® TR28	21280 m²/kg	0.130 µm	0.357 µm	40.2 µm

13.6.2. Bending test results

In Figure 13.7 the stiffness moduli and bending strength were summarized for the tested systems in the three-point bending test.

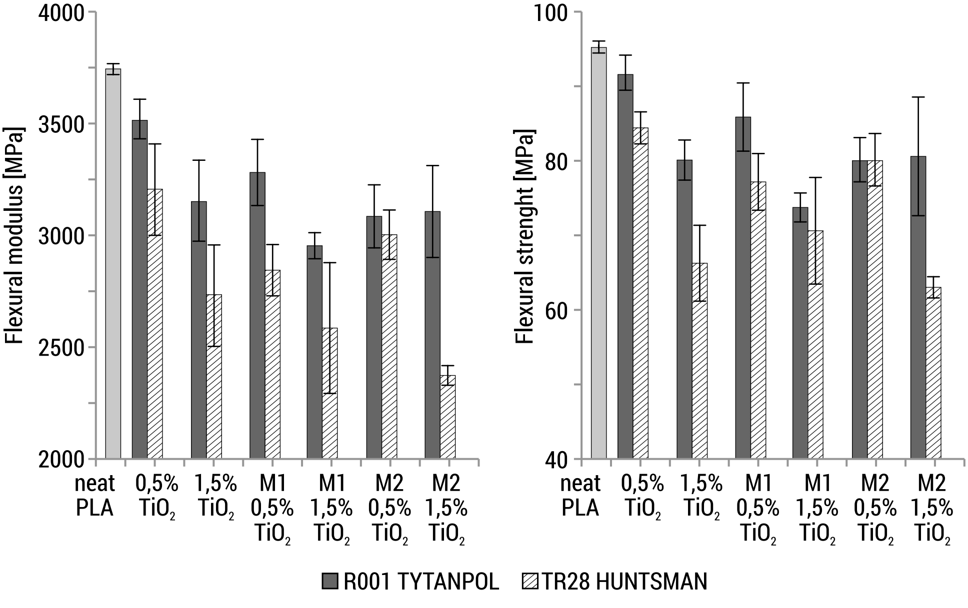


FIGURE 13.7. Results from the three-point bending test

The incorporation of TiO_2 to the PLA matrix resulted in a decrease in the stiffness modulus and strength in relation to the results obtained for the unmodified polymer. The decrease of parameters value depends on the percentage of the filler. For samples containing 1.5% TiO_2, the obtained results are lower than for samples with 0.5% TiO_2 concentration.

Comparing the effect of the type of TiO_2 used, it can be said that the samples containing ROO1 Tytanpol TiO_2 are characterized by better bending strength properties than Huntsman TR28. These differences may be due to agglomeration of the Huntsman TR28 filler particles as shown by DLS studies.

The introduction of organosilicon modifiers resulted in a reduction in bending stiffness for all tested systems, which is probably due to their plasticizing effect. The bending strength of the tested composites is also dependent on the proportion of the modifier. The tendency of the decrease in bending strength is maintained in relation to pure polylactide and composites containing only TiO_2 fillers.

13.6.2. Tensile test results

Figure 13.8 presents the results of measuring the tensile strength and Young's modulus for identical systems. The addition of TiO_2 had a positive effect on the tensile strength of the tested systems. The tensile strength increased with the introduction of TiO_2 into PLA. In the case of systems containing ROO1 Tytanpol, the effect of the filler concentration on the strength is visible, which decreases with the increase of its percentage, while for the TR28 Huntsman white, there is practically no effect of the concentration on this strength parameter. The introduction of organosilicon modifiers to the tested systems resulted in changes in the tensile strength of the tested composites. The modifiers M1 and M2 reduced the strength of the composites with Huntsman TR28, regardless of its concentration in PLA; however, they improved the tensile strength of the system containing 1.5% TiO2 ROO1 Tytanpol.

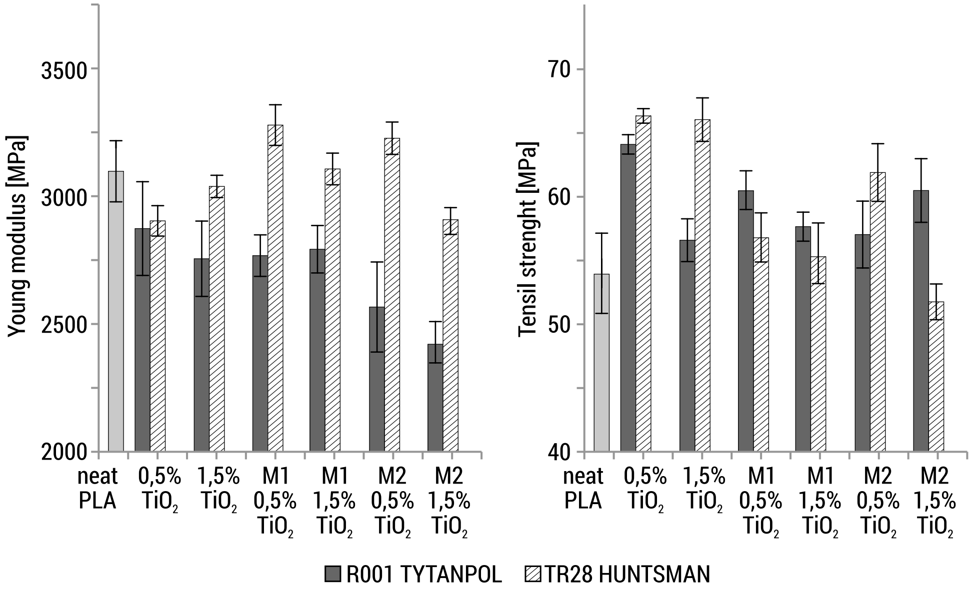


FIGURE 13.8. **Results of uniaxial static stretching**

Figure 13.9 (left) shows the percentage elongation of the specimens with the tensile strength. When analyzing all the samples, we can see that the deformation of the samples is positively influenced by the addition of TiO_2 and a modifier to the systems. For samples containing ROO1 Tytanpol, an increase in filler concentration also improves plastic properties. Huntsman TiO_2 TR28 composites do not show deformation dependence on the filler concentration.

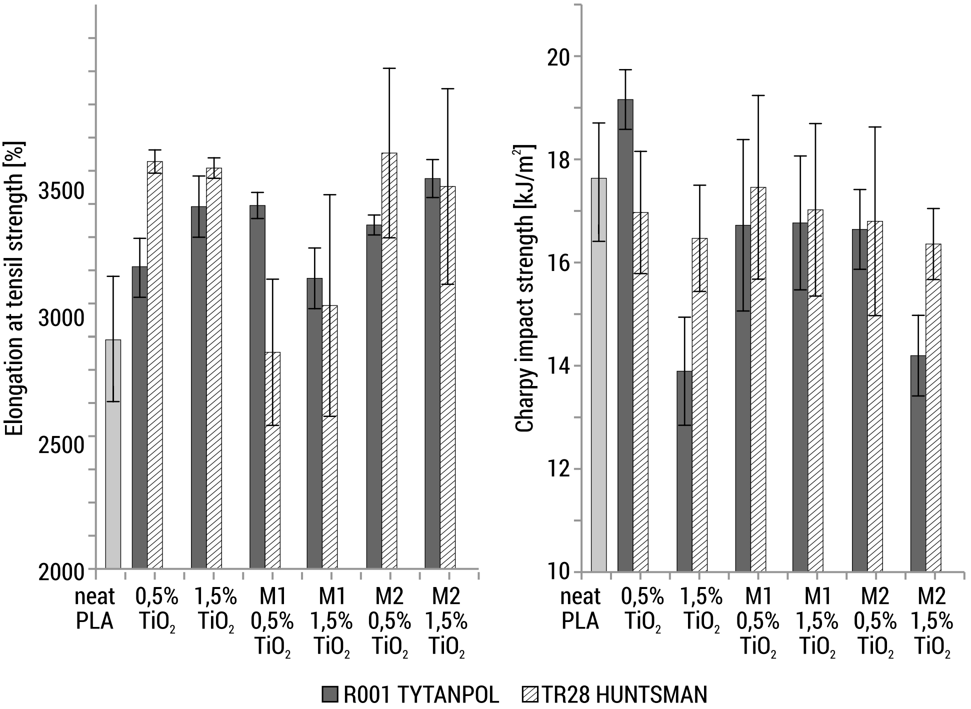


FIGURE 13.9. **Elongation at tensile strength – left, Charpy Impact – right.**

The Young's modulus of the obtained filled composites is lower than that of pure PLA and for systems containing ROO1 Tytanpol whiteness it decreases with increasing TiO_2 concentration, while for systems containing TR28 Huntsman the opposite relationship occurs.

The addition of organosilicon modifiers caused a significant increase in Young's modulus for composites containing TR28 Huntsman titanium white. This effect is most visible for systems with a lower concentration of TiO_2. The stiffness modulus for systems filled with the white ROO1 Tytanpol decreases after adding both modifiers and is lower than for PLA.

The modifiers M1 and M2 introduced into the composites improve their elastic properties, however, for some tested systems, the negative influence of the modifier M1 on the elongation at Rm was also observed. Systems containing TR28 Huntsman show deterioration of properties after adding the M1 compound to the composite, while for the system containing 0.5% TiO_2 concentration ROO1 Titanpol, an improvement in flexibility is visible. The introduction of the M2 modifier causes an increase in deformation regardless of the concentration of TiO_2 and the type of filler used.

13.6.3. Charpy impact test results

Figure 13.9 (right) shows impact strength results of the polymer composites tested. The impact strength of the produced composites is close to that of pure PLA and lies within the measurement error limit. The impact strength decreases with increasing TiO_2 concentration regardless of the TiO_2 grade used. The highest impact strength value was determined for the composite with 0.5% TiO_2 concentration ROO1 Titanpol. The use of organosilicon compounds causes a substantial increase in the impact strength of composites with 1.5% TiO_2. For composites with 0.5% TiO_2, the impact strength increases for TR 28 Huntsman TiO_2 and decreases for TiO_2 ROO1 Tytanpol

13.7. Conclusion

In these studies, a new class of polymer composites for FDM 3D printing applications was successfully obtained and it was shown that various TiO2 species and organosilicon compounds can be successfully used as modifiers of PLA biopolymer properties. The paper compares the effect of two types of TiO_2 from different manufacturers on the properties of polymer biocomposites for applications in the FDM incremental technology. The particle size distribution of both used fillers was determined. On the basis of the DLS analysis, the tendency of TIOXIDE® TR28 powder to agglomerate particles was observed (D_v (90) = 40.2 μm), which may directly translate into the strength characteristics of the composites filled with it. The properties of TiO_2 composites depend on the filler concentration. Flexural strength and modulus at bending reach lower values for 1.5% concentration of TiO_2. However, the tensile strength and elongation are higher for this percentage of TiO_2 in PLA matrix. The applied fillers had a positive effect on the improvement of the elastic properties of the produced polymer composites and the tensile strength. This effect was additionally intensified by the introduction of organosilicon modifiers, which also improved the impact toughness of TiO_2 filled composites in relation to the reference systems. The introduction of fillers and modifiers to the PLA matrix resulted in the plasticizing effect, therefore the stiffness and bending strength of the tested composites decreased. On the basis of the results obtained in mechanical tests, it can be concluded that the used organosilicon modifiers may therefore interact differently at the polymer/TiO_2 interface, depending on the grade of titanium white used.

Acknowledgments: This research was funded by the National Centre for Research and Development, Poland, grant number LIDER/01/0001/L-10/18/NCBR/2019.

References

1. Prince, J.D.: "3D printing: an industrial revolution", Journal of Electronic Resources in Medical Libraries, 11 (1), 39–45, 2014.
2. Shahrubudin N., Lee T.C., Ramlan, R.: "An overview on 3D printing technology: Technological, materials, and applications", Procedia Manufacturing, 35, 1286–1296, 2019.
3. Upcraft S., Fletcher R.: "The rapid prototyping technologies", Assembly Automation, 23 (4), 318–330, 2003.
4. Bates-Green K., Howie T.: "Materials for 3D printing by fused deposition", Technical Education in Additive Manufacturing and Materials, 1–21, 2017.
5. Chen Y., Geever L.M., Killion J.A., Lyons J.G., Higginbotham, C.L., Devine D.M.: "Review of multifarious applications of poly (lactic acid)", Polymer-Plastics Technology and Engineering, 55 (10), 1057–1075, 2016.
6. Maharana T., Mohanty B., Negi Y.S.: "Melt – solid polycondensation of lactic acid and its biodegradability", Progress in Polymer Science (Oxford) Elsevier Ltd, 34 (1) 99–124, 2009.
7. Cadar O., Cadar S., Senia M., Majdik C., Roman C.: "Synthesis of poly (l-lactic acid) by direct polycondensation", Stud Univ Babes-Bolyai Chem, 56 (2), 57–64, 2011.
8. Cheng Y., Deng S., Chen P., Ruan R.: "Polylactic acid (PLA) synthesis and modifications: A review", Frontiers of Chemistry in China, 4(3), 259–264, 2009.
9. Banerjee R., Ray S.S.: "An overview of the recent advances in polylactide-based sustainable nanocomposites", Polymer Engineering & Science, 61(3), 617–649, 2021.
10. Gálvez J., Correa Aguirre J.P., Hidalgo Salazar M.A., Vera Mondragón B., Wagner E., Caicedo C.: "Effect of extrusion screw speed and plasticizer proportions on the rheological, thermal, mechanical, morphological and superficial properties of PLA", Polymers, 12(9), 2111, 2020.
11. Cheng G.Z., Estepar R. S.J., Folch E., Onieva J., Gangadharan S., Majid A.: "Three-dimensional printing and 3D slicer: powerful tools in understanding and treating structural lung disease", Chest, 149(5), 1136–1142, 2016.
12. Gao Q., Feng M., Li, E., Liu C., Shen C., & Liu X.: "Mechanical, thermal, and rheological properties of Ti3C2Tx MXene/thermoplastic polyurethane nanocomposites", Macromolecular Materials and Engineering, 305(10), 2000343, 2020.
13. Marquis D.M., Guillaume E., Chivas-Joly C.L.: "Properties of nanollers in polymer", Nanocomposites and Polymers with Analytical Methods, 2011.
14. Wunderlich W., Oekermann T., Miao L., Hue N.T., Tanemura S., Tanemura M.: "Electronic properties of nano-porous TiO 2- and ZnO thin films – comparison of simulations and experiments", Journal of Ceramic Processing & Research, 5(4), 343–354, 2004.
15. Luo Y., Wang L., Zou Y., Sheng X., Chang L., Yang D.: "Electrochemically deposited Cu_2O on TiO_2 nanorod arrays for photovoltaic application", Electrochemical and Solid-State Letters, 15(2), H34, 2011.
16. Marciniec, B. (ed.): Hydrosilylation: A Comprehensive Review on Recent advances, p. 215, 2009.
17. Scott D.W.: "Thermal rearrangement of branched-chain methylpolysiloxanes1", Journal of the American Chemical Society, 68(3), 356–358, 1946.

18. Barczewski M., Czarnecka-Komorowska D., Andrzejewski J., Sterzyński T., Dutkiewicz M., Dudziec B.: “Właściwości przetwórcze termoplastycznych tworzyw polimerowych modyfikowanych silseskwioksanami (POSS)”, Polimery, 58(10), 805–815, 2013.
19. Perrin F.X., Panaitescu D.M., Frone A.N., Radovici C., Nicolae C.: “The influence of alkyl substituents of POSS in polyethylene nanocomposites”, Polymer, 54(9), 2347–2354, 2013.
20. Efrat T., Dodiuk H., Kenig S., McCarthy, S.: “Nanotailoring of polyurethane adhesive by polyhedral oligomeric silsesquioxane (POSS)”, Journal of Adhesion Science and Technology, 20(12), 1413–1430, 2006.
21. Chen J. H., & Chiou, Y. D.: “Crystallization behavior and morphological development of isotactic polypropylene blended with nanostructured polyhedral oligomeric silsesquioxane molecules”, Journal of Polymer Science Part B: Polymer Physics, 44(15), 2122–2134, 2006.
22. Brząkalski D., Sztorch B., Frydrych M., Pakuła D., Dydek K., Kozera R., Przekop R. E.: “Limonene derivative of spherosilicate as a polylactide modifier for applications in 3D printing technology”, Molecules, 25(24), 5882, 2020.
23. McCusker C., Carroll J.B., Rotello V.M.: “Cationic polyhedral oligomeric silsesquioxane (POSS) units as carriers for drug delivery processes”, Chemical Communications, (8), 996–998, 2005.
24. Kannan R.Y., Salacinski H.J., Ghanavi J.E., Narula A., Odlyha M., Peirovi H., Seifalian A.M.: “Silsesquioxane nanocomposites as tissue implants”, Plastic and Reconstructive Surgery, 119(6), 1653–1662, 2007.
25. Kim S.K., Heo S.J., Koak J.Y., Lee J.H., Lee Y.M., Chung D.J., Hong S.D.: “A biocompatibility study of a reinforced acrylic-based hybrid denture composite resin with polyhedraloligosilsesquioxane”, Journal of oral rehabilitation, 34(5), 389–395, 2007.
26. Rajeshkumar G., Seshadri S.A., Devnani G.L., Sanjay M.R., Siengchin S., Maran, J.P., Anuf A.R.: “Environment friendly, renewable and sustainable poly lactic acid (PLA) based natural fiber reinforced composites – a comprehensive review”, Journal of Cleaner Production, 310, 127483, 2021.
27. Pricop L., Durdureanu-Angheluta A.: “Polysiloxane Networks Synthesis and thermo-mechanical characterization”, Materiale Plastice, 48(4), 308–314, 2011.
28. Mark J. E., Schaefer D.W., Lin G.: The polysiloxanes, Oxford University Press, 2015.

DOI: 10.24427/978-83-67185-49-3_14

Chapter 14
Effect of porosity on the properties of nickel-free stainless steel obtained by powder metallurgy route

Eliza Romańczuk-Ruszuk, Paweł Lipowicz*, Bogna Sztorch**, Robert E. Przekop***

*Bialystok University of Technology, Institute of Biomedical Engineering,
e-mail: e.romanczuk@pb.edu.pl, p.lipowicz@pb.edu.pl
** Adam Mickiewicz University in Poznań, Centre for Advanced Technologies,
e-mail: bogna.sztorch@amu.edu.pl, rprzekop@amu.edu.pl*

Abstract: This paper presents the results of research on nickel-free austenitic stainless steel with the composition Fe-18% Cr-12% Mn, obtained by powder metallurgy, with the use of mechanical alloying of elementary powders, manganese nitride and various process atmospheres. The influence of different protective atmosphere of the process (nitrogen and argon) and the addition of manganese nitride on porosity, microstructure, hardness and corrosion resistance of the sinters was described. Steel samples after pressing and sintering at 1200°C (1473 K) in nitrogen were characterized by a relative density above 80% and a porosity among 12–19%. In tested materials the pores on the surface are visible in each of the tested materials, which is confirmed by SEM images. The material obtained from powders with manganese nitride after mechanical alloying in nitrogen is characterized by the highest hardness and corrosion resistance. The addition of manganese nitride can be another way of introducing nitrogen into nickel-free steel.

Keywords: nickel-free austenitic stainless steel, powder metallurgy, porosity, mechanical alloying, corrosion resistance

14.1. Introduction

The constantly developing implantology requires the use of modern biomaterials that should meet specific biofunctionality requirements. The material used for implantation should be characterized by the following features: biocompatibility, high strength and fatigue resistance, corrosion resistance in the environment of body fluids, modulus of elasticity similar to the bone elasticity modulus and low price [1, 2].

In recent years, there has been increased demand for implants, plates, screws and bone wires, therefore the subject of producing medical materials is very important.

The 316LV austenitic steel has been most commonly used stainless steel. In the composition of this steel it can be distinguished an adverse element for the human body – nickel. It is an alloy addition strongly stabilizing austenite, but at the same time, too high a concentration of it in the human body can cause allergy. Since the 90s of the 20th century, the amount of allergic reactions to this element among the human population is still growing. Every tenth human tested is allergic to nickel, mainly women. Therefore, materials without nickel used in medicine are an important scientific and social issue [3–6].

The promising material that may replace commonly used 316LV steel in the future, is the nickel-free austenitic stainless steel. The literature about the production of nickel-free austenitic stainless steels indicates that nickel can be replaced by another austenitizing element, like nitrogen and manganese. At the same time, there are some difficulties with implementing nitrogen into steel during melting, because it is a gas. Fabrication methods of the nickel-free austenitic stainless steels include high-pressure melting technologies, such as pressurized induction melting (PIM), pressurized electroslag remelting (PESR) and pressurized plasma arc melting (PARP) [7, 8]. High-pressure melting technologies need special equipment, which is expensive and difficult to operate [9, 10]. Contrary to this method, nickel-free austenitic stainless steel can be synthesized in a solid state by mechanical alloying (MA) of elemental or prealloyed powders [11, 12]. The main advantage is that the MA process generates a number of structural defects that can cause an increase in solubility of nitrogen in the steel matrix [13–16]. Also the amorphous phase obtained during prolonged ball milling can be another factor influencing the final properties of steel because of a strong impact on nitrogen solubility and nano- (sub)grain microstructure creation [1]. Therefore, the powder metallurgy (PM) route including mechanical alloying can be an alternative method for the production of nickel-free austenitic stainless steels.

In this study, the influence of porosity on the properties of nickel-free austenitic steel obtained by powder metallurgy was examined.

14.2. Materials and methods

Elemental powders of iron, chromium, manganese (Fe, Cr, Mn) with an average particle size of ~45 μm and a purity of 99.95% (basic metal) and, optionally, an addition of manganese nitride (Mn_4N) with a particle size of ~15 μm and a purity of 99.95% supplied by Alfa Aesar were used in the work, which were mechanically alloyed in a planetary ball mill Pulverisette 6 (Fritsch). The ball-to-powder ratio was 8:1, the rotational speed was 250 rpm. The mechanical alloying (MA) process was carried out in an atmosphere of pure nitrogen or argon (99.999% purity). The addition of manganese nitride was chosen so that the percentage of Mn in the alloy was 12%. The powders were mechanically alloyed for 90 hours. One MA cycle lasted 20 minutes, including a 10-minute break.

The mechanically alloyed powder was cold compacted in a cylindrical die of 20 mm in diameter and 10 mm high, using an uniaxial hydraulic press with a pressure of 600 MPa and sintered at 1200 °C (1473 K) for 2h, in nitrogen with cooling rate of 600°C/h (873 K/h). No lubricant or process controlled agent was applied.

Changes in the density of sintered materials were tested in accordance with the PN-EN ISO 2738:2001 standard. For metallographic investigation, the sintered samples were polished, etched and observed using the Scanning Electron Microscope (SEM, Hitachi 3000N). The grain size of the tested specimens was examined according to ISO 643:2003. The chemical analysis of the consolidated specimens was performed by the LECO TCH600 analyzer and Scanning Electron Microscope equipped with the Energy Dispersive Spectrometer (SEM-EDS) analysis (Hitachi 3000N, Tokyo). The porosity of the samples was examined using the X-ray computed microtomography (μCT) method. The study was carried out on the Skyscan 1172 microtomograph (Bruker company). The device is equipped with a 10 W X-ray tube and an 11 mpx CCD matrix. The matrix resolution was set to 2000x1332px, and the parameters of the lamp were 88 kV/112 μA. An aluminum filter with a thickness of 0.5 mm was used. The exposure time was 1500 ms. The resolution used for the test is 3.40 μm and the rotation angle is 0.40°. In order to fully overexpose the object, a 360° rotation was used. The images were made in a 16-bit detail representation. The tests were carried out on cuboidal samples with dimensions of 3 mm x 3 mm x 7 mm, cut from the center of the samples after their consolidation. Vickers hardness (HV_{10}) of the samples was measured by using the Anova Hardness Tester under a load of 98 N (10 kg) and an indentation time of 10 s. The microhardness (μHV0.2) tests were performed using the PMT-3 tester under a load of 1.96 N (0.2 kg) for 10 s. Phase analysis was performed on an X-ray diffractometer (XRD) Bruker Eco D8 Advance, using Cu-$K\alpha_2$ radiation of 40 kV and 25 mA. For all samples, the angular range (2θ) of 20° to 100° with a step width of 0.01 and a step time of 5 s was used. The Al_2O_3 sample was used as the XRD standard to correct an instrumental line broadening. A background correction the $K\alpha_2$ stripping were performed. Corrosion resistance of the sintered specimen was tested according ISO 10993-15. The PGP201 VoltaLab galvanostat/potentiostat (Radiometer Analytical, France) equipped with the VoltaMaster 4 software was used. As a corrosion medium, 70 mL of the Hank's solution was applied. The opening potential (E_{OCP}) was determined under no-current conditions for 2 hours. Then the anodic polarization curves were recorded from $E_{start} = E_{OCP} - 100$ mV. A three-electrode electrochemical measuring circuit was adapted in each corrosion test. The reference electrode was a saturated calomel electrode. An auxiliary electrode was a platinum plate. On the basis of the obtained results, the corrosion potential of Ecor, the polarization resistance Rp and the corrosion current density (Icor) were calculated. The corrosion potential and the polarization resistance were determined by the Stern linear polarization method. Corrosion current density was determined using Tafel extrapolation and the Stern-Geary equation [17].

The following designations were used in the study: 12Mn – manganese percentage; E – the use of elementary powders Fe, Cr, Mn; Mn_4N – the use of manganese nitride and iron and chromium powders; Ar/N_2 – MS process atmosphere: argon/nitrogen

14.3. Results

Table 14.1 shows the total, closed and open porosity, as well as the relative and absolute densities of the sintered samples. In all tested samples, the relative density was above 80% and open porosity has a higher value compared to the closed porosity. The total porosity of the sample with manganese nitride is 12±1%, while the sample with manganese elemental powder with nitrogen and argon atmosphere is 15±1% and 19±1%, respectively. Closed porosity of materials obtained from elementary powders mechanically alloyed in nitrogen and argon atmosphere is 2±1% and 4±1%, respectively, while for material with manganese nitride it is 3±1%.

TABLE 14.1. The density and porosity of the sinters

Samples	**Total porosity [%]**	**Closed porosity [%]**	**Open porosity [%]**	**Green sample density [kg/m^3]**	**Sinter density [kg/m^3]**	**Relative density [%]**	**Compaction parameter Ψ [%]**
12Mn-E-N$_2$	15±1	2±1	13±1	5232±56	6786±59	85±1	0.609±0.021
12Mn-E-Ar	19±1	4±1	15±1	5145±56	6330±59	81±1	0.449±0.086
12Mn-Mn$_4$N-N$_2$	12±1	3±1	9±2	5432±52	6849±93	88±1	0.602±0.011

Figure 14.1 compares the phase structure of the samples with 12% Mn obtained from elemental powders, after MS in nitrogen and powders with the addition of manganese nitride alloyed in nitrogen and argon. The presented diffractograms show that only the sinter obtained from powders with manganese nitride after mechanical alloying in nitrogen has an austenitic structure. In other materials tested, there was an austenitic and ferritic structure.

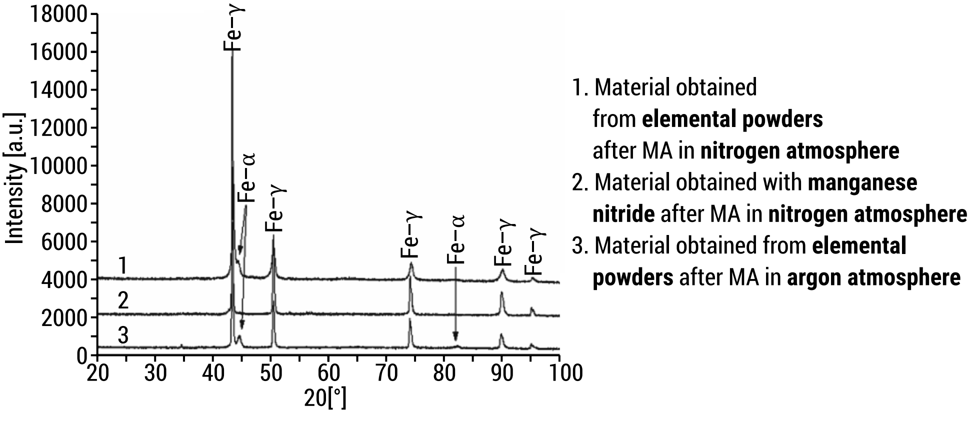


FIGURE 14.1. Phase structure of sinters with the composition Fe-18% Cr-12% Mn obtained from elemental powders or manganese nitride MS in various protective atmosphere, pressed and sintered at *T* = 1200°C, atm. Nitrogen

In the sinter obtained from elementary powders after MA in nitrogen, austenite grains with pores and irregular precipitates are visible (Fig. 14.2a). In the sample obtained from powders with Mn_4N, the grains are more aligned, although average grain sizes in both samples are comparable (Tab. 14.2).

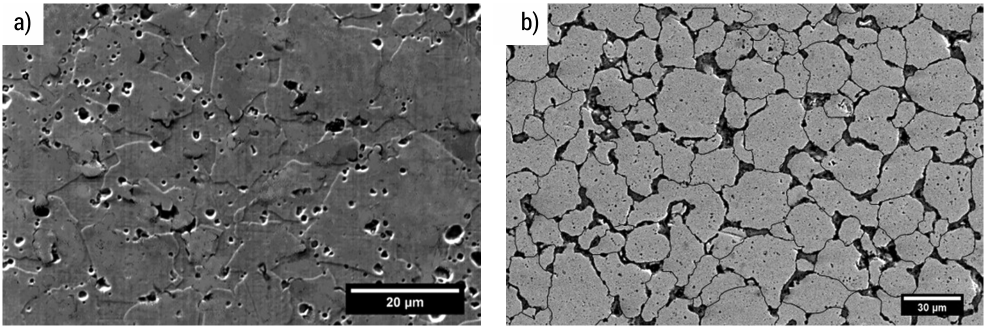


FIGURE 14.2. Microstructure of: a) 12Mn-E-N_2 sample, b) 12Mn-Mn_4N-N_2

TABLE 14.2. Grain size of selected materials

Samples	Grain size [µm]
12Mn-E-N_2	16 ± 6
12Mn-E-Ar	20 ± 7
12Mn-Mn_4N-N_2	19 ± 4

The chemical composition analysis (Tab. 14.3) confirmed the basic proportion of the main alloying elements (18% of Cr, 12% of Mn). Nitrogen was present in each analyzed material. The highest nitrogen content was in the 12Mn-Mn_4N-N_2 sinter, as expected.

TABLE 14.3. Chemical composition of the samples obtained. The results were obtained from a LECO spectrometer and SEM-EDS analysis

Samples	Elements					
	Fe	Cr	Mn	N	O	C
12Mn-E-N_2	68.93±1.01	17.96±0.49	12.13 ±0.39	0.526 ±0.026	0.390 ±0.041	0.0579 ±0.0011
12Mn-E-Ar	68.51±0.98	18.06±0.59	11.95 ±0.42	0.218 ±0.073	0.389 ±0.023	0.0662 ±0.0018
12Mn-Mn_4N- N_2	68.77±1.16	18.01±1.01	12.03±0.35	0.813 ±0.052	0.311 ±0.016	0.0612 ±0.0021

Figure 14.3 shows the dependence of hardness HV_{10} and relative density. Higher sinter relative density results in higher hardness of the material. The HV_{10} hardness of the material obtained from elementary powders mechanically alloyed in argon and nitrogen atmosphere is 289±12 and 261±20, respectively. The hardness of the material obtained from manganese nitride mechanically alloyed in nitrogen atmosphere is 295±17. At the same time, the material with manganese nitride (88%) is characterized by the highest density, and 12Mn-E-Ar has the lowest density.

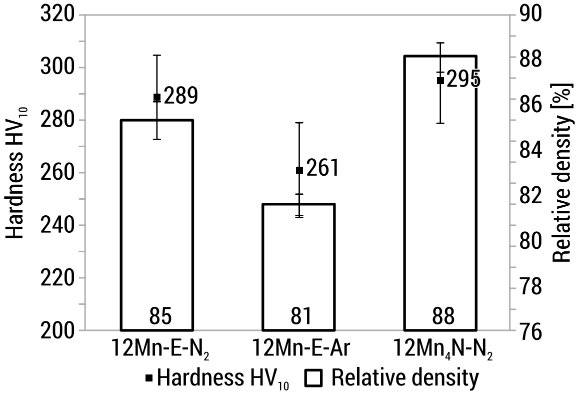


FIGURE 14.3. Dependence of relative density on hardness HV_{10} of sintered materials

Figure 14.4 shows the dependence of the microhardness of materials as a function of grain size. In each material there is a constant dependence of the microhardness on the grain size, described by the Hall-Petch formula. The smaller the grain size, the higher the microhardness. In the 12Mn-E-Ar sample, lower microhardness may be due to the larger grains observed in this material and the lowest nitrogen content, compared to the other samples. The 12Mn-Mn_4N-N_2 sinter was characterized by higher microhardness, which can be explained by its fine-crystalline structure, higher density (approx. 88%) and a relatively high nitrogen content in the steel.

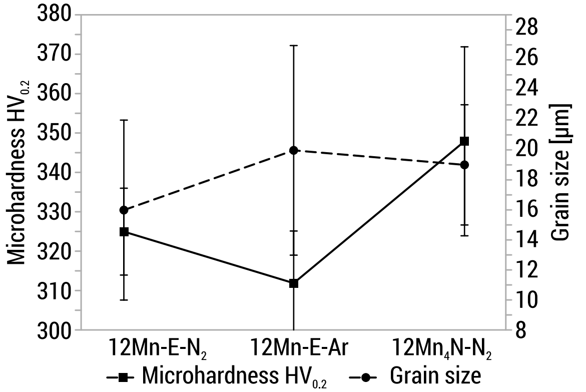


FIGURE 14.4. Dependence of grain size and microhardness of sintered materials

Changes in the open circuit potential of the tested materials are shown in Figure 14.5a. The analysis of the results from Figure 14.5a shows that in the tested materials, there is a decrease in the open circuit potential during the first 5 minutes of the test for the 12Mn-Mn_4N-N_2 sample and 40 minutes for the 12Mn-E-N_2 sample. In each sinter samples, the value of the open circuit potential after 2 hours is lower than the initial value.

The anodic polarization curves (potentiodynamic curves) are shown in Figure 14.5b, and in Table 14.4 the basic parameters of corrosion resistance are given: corrosion potential (Ecor), polarization resistance (Rp) and corrosion current density (Icor) of the tested materials. The active and passive areas are clearly visible in the polarization curves. The EOCP in the 12Mn-Mn4N-N2 sample takes more constant values, which indicates a higher corrosion resistance of the material. In the material obtained from powders with the addition of manganese nitride after MS in nitrogen, there is a narrow hysteresis loop. Moreover, only in this material the potentiodynamic curve crosses the graph again in the passive area after the potential change.

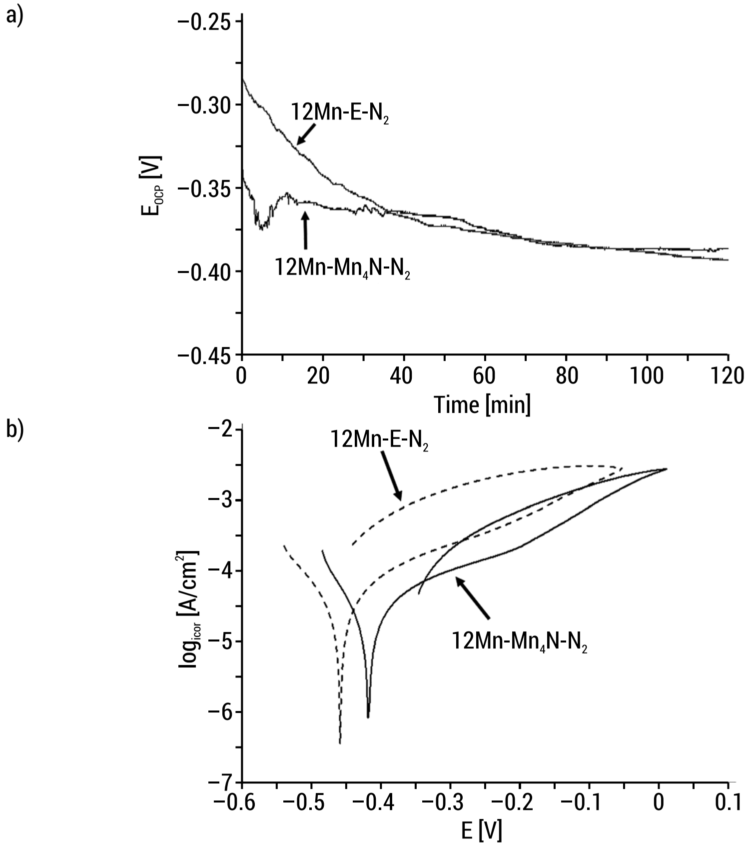


FIGURE 14.5. a) E_{OCP} corrosion potential changes over time, recorded with an open circuit without external polarization, b) potentiodynamic curves of the 12Mn-E-N_2 and 12Mn-Mn_4N-N_2 samples

TABLE 14.4. Basic parameters of the corrosion process: Ecor – corrosion potential, Rp – polarization resistance, Icor – corrosion current density

Samples	Ecor [V]	Rp [ohm]	Icor [uA/cm^2]
12Mn-E-N$_2$	−0.462±0.029	10430±210	3.687±0.453
12Mn-Mn$_4$N-N$_2$	−0.413±0.022	697350±398	0.044±0.004

14.4. Discussion

The analysis of the results of the density and porosity of the tested samples showed that the samples in which manganese nitride was used for mechanical alloying had the lowest total porosity, and thus the highest density. The highest porosity and the lowest density were characteristic of samples obtained from elementary powders alloyed in an argon atmosphere. By comparing the samples obtained from MA elemental powders in different atmospheres, it can be concluded that the nitrogen atmosphere increases the relative density and reduces the porosity. The compaction parameter Ψ determines whether the sample has shrinkage or swelling in the tested sinters. When this parameter is positive, it indicates a shrinkage of the samples, otherwise the samples swell. The swelling of the sample is disadvantageous, causes de-densification, a reduction of the sintering kinetics, and opposes sintering [18]. The compaction parameter was positive in each tested sinter, which means that the materials shrink after sintering.

Phase structure studies using X-ray diffraction indicated the complex nature of phase transformations taking place in the tested nickel-free austenitic stainless steel. In order to increase the possibility of obtaining austenite in nickel-free austenitic stainless steel, in addition to nitrogen as the MA atmosphere, manganese nitride can be used.

As mentioned in chapter 14.3, the highest nitrogen content was in the 12Mn-Mn$_4$N-N$_2$ sinter. This can be explained by the fact that in this material nitrogen was introduced in two forms: as a nitride and as a milling atmosphere.

The analysis of the hardness of the materials showed that the material obtained from MS powders in nitrogen is characterized by higher hardness. In addition, the open porosity of this material is lower, so there are less pores on the surface. The material obtained from powders with manganese nitride mechanically alloyed in nitrogen was characterized by the highest hardness, and its porosity was the lowest. The higher hardness of this material may be related to the highest nitrogen content and the use of nitrogen in various forms (as the atmosphere of the MA process, or as an alloying powder).

A decrease in the open circuit potential indicates the chemical activity of the alloying elements on the surface of the material and its faster corrosion. The narrow hysteresis loop present in the results of potentiodynamic corrosion resistance tests may indicate the metastable nature of the pitting formation and the material's ability to repassivate the surface layer.

14.5. Conclusion

In the study, samples of nickel-free steel were obtained using powder metallurgy techniques and introducing nitrogen in two ways: as a protective atmosphere and as an addition of manganese nitride. On the basis of the conducted research, the following general conclusions were formulated:

a) the porosity of the samples after sintering at 1200°C in nitrogen atmosphere was among 12–19% and the relative density was in the range of 81–88%;
b) X-ray diffraction studies showed that only in the 12Mn-Mn_4N-N_2 material there was an austenitic phase;
c) the highest hardness and microhardness was characteristic of the 12Mn-Mn_4N-N_2 material, which had the smallest porosity and grain size;
d) the electrochemical tests show that the material obtained from powders with manganese nitride mechanically alloyed in nitrogen has the best corrosion properties.

Acknowledgments: This research was funded by the Institute of Biomedical Engineering, Faculty of Mechanical Engineering, Bialystok University of Technology, project No. WZ/WM-IIB/2/2020.

References

1. Yang F., Yibin R.: “Nickel-free austenitic stainless steels for medical applications”, Science and Technology Advanced Materials, 2010.
2. Di Schino A., Kenny J.: “Grain refinement strengthening of a micro-crystalline high nitrogen austenitic stainless steel”, Materials Letters, 57, (12), 1830–1834, 2003.
3. Tuliński M., Jurczyk M.: “Mechanical and corrosion properties of Ni-free austenitic stainless steels”, Archives of Metallurgy and Materials, 563, (3), 2008.
4. Wataha J., O'Dell N., Singh B., Ghazi M., Whitford G.: “Relating nickel-induced tissue inflammation to nickel release in vivo”, Journal of Biomedical Materials Research, 58, (537), 2002.
5. Eliades T., Pratsinis H., Kletsas D., Eliades G., Makou M.: “Characterization and cytotoxicity of ions released from stainless steel and nickel-titanium orthodontic alloys”, American Journal of Orthodontics and Dentofacial Orthopedics, 125, (1), 24–29, 2004.
6. Uggowitzer P.J., Magdowski R., Speidel M.O.: “Nickel free high nitrogen austenitic steels”, ISIJ international, 36, 901–908, 1996.
7. Rawers J.C., Dunning J.S., Asal G., Reed R.P.: “Characterization of stainless steels melted under high nitrogen pressure”, Metallurgical Transactions A, 23,1992.

8. Niinomi M., Nakai M., Hieda J.: "Development of new metallic alloys for biomedical applications", Acta Biomaterialia, 8, 2012.
9. Ahmed A., Ghali S., Eissa M., El Badry S.: "Influence of partial replacement of nickel by nitrogen on microstructure and mechanical properties of austenitic stainless steel", Journal of Metallurgy,1–6, 2011.
10. Cisneros M.M., Lopez H.F., Mancha H., Vazquez E.: "Development of austenitic nanostructures in high-nitrogen steel powders processed by mechanical alloying", Metallurgical and Materials Transactions A, 33A, 2139–2144, 2002.
11. Salahinejad E., Amini R., Hadianfard M. J.: "Structural evolution during mechanical alloying of stainless steels under nitrogen", Powder Technology, 215–216, 247–253, 2012.
12. Sumita M., Hanawa T., Teoh S.H.: "Development of nitrogen-containing nickel-free austenitic stainless steel for metallic biomaterials – review", Material Science and Engineering C, 24, 2004.
13. Suryanarayana C.: "Mechanical alloying and milling", Progress in Materials Science, 46, 1–184, 2001.
14. Duan C., Chen C., Zhang J., Shen Y., Feng X.: "Nitriding of Fe-18Cr-8Mn stainless steel powders by mechanical alloying method with dual nitrogen source", Powder Technology, 294, 2016.
15. Loginov P., Levashov E., Kurbatkina V., Zaitsev A., Sidorenko D.: "Evolution of the microstructure of Cu–Fe–Co–Ni powder mixtures upon mechanical alloying", Powder Technology, 276, 166–174, 2015.
16. Azabou M., Ibn Gharsallah H., Escoda L., Suñol J., Kolsi A., Khitouni M.: "Mechanochemical reactions in nanocrystalline Cu–Fe system induced by mechanical alloying in air atmosphere", Powder Technology, 224, 338–344, 2012.
17. Stern M., Geary A.: "Electrochemical polaryzation I. A theoretical analysis of the shape of polaryzation curves", Journal of the electrochemical Society, 56–63, 1957.
18. Manière C., Saccardo E., Lee G., McKittrick J., Molinari A., Olevsky E.: "Swelling negation during sintering of sterling silver: An experimental and theoretical approach", Results in Physics, 11, 79–84, 2018.

List of Tables

List of Figures

Politechnika
Białostocka