BIOTECHNO 2025

The Seventeenth International Conference on Bioinformatics, Biocomputational Systems and Biotechnologies

ISBN: 978-1-68558-246-3

March 9th –13th, 2025

Lisbon, Portugal

BIOTECHNO 2025 Editors

Claus-Peter Rückemann, Universität Münster / Leibniz Universität Hannover, Germany

BIOTECHNO 2025

Foreword

The Seventeenth International Conference on Bioinformatics, Biocomputational Systems and Biotechnologies (BIOTECHNO 2025), held between March 9 - 13, 2025, covered these three main areas: bioinformatics, biomedical technologies, and biocomputing.

Bioinformatics deals with the system-level study of complex interactions in biosystems providing a quantitative systemic approach to understand them and appropriate tool support and concepts to model them. Understanding and modeling biosystems requires simulation of biological behaviors and functions. Bioinformatics itself constitutes a vast area of research and specialization, as many classical domains such as databases, modeling, and regular expressions are used to represent, store, retrieve and process a huge volume of knowledge. There are challenging aspects concerning biocomputation technologies, bioinformatics mechanisms dealing with chemoinformatics, bioimaging, and neuroinformatics.

Biotechnology is defined as the industrial use of living organisms or biological techniques developed through basic research. Bio-oriented technologies became very popular in various research topics and industrial market segments. Current human mechanisms seem to offer significant ways for improving theories, algorithms, technologies, products and systems. The focus is driven by fundamentals in approaching and applying biotechnologies in terms of engineering methods, special electronics, and special materials and systems. Borrowing simplicity and performance from the real life, biodevices cover a large spectrum of areas, from sensors, chips, and biometry to computing. One of the chief domains is represented by the biomedical biotechnologies, from instrumentation to monitoring, from simple sensors to integrated systems, including image processing and visualization systems. As the state-of-the-art in all the domains enumerated in the conference topics evolve with high velocity, new biotechnologes and biosystems become available. Their rapid integration in the real life becomes a challenge.

Brain-computing, biocomputing, and computation biology and microbiology represent advanced methodologies and mechanisms in approaching and understanding the challenging behavior of life mechanisms. Using bio-ontologies, biosemantics and special processing concepts, progress was achieved in dealing with genomics, biopharmaceutical and molecular intelligence, in the biology and microbiology domains. The area brings a rich spectrum of informatics paradigms, such as epidemic models, pattern classification, graph theory, or stochastic models, to support special biocomputing applications in biomedical, genetics, molecular and cellular biology and microbiology. While progress is achieved with a high speed, challenges must be overcome for large-scale bio-subsystems, special genomics cases, bio-nanotechnologies, drugs, or microbial propagation and immunity.

We take here the opportunity to warmly thank all the members of the BIOTECHNO 2025 Technical Program Committee, as well as the numerous reviewers. The creation of such a high quality conference program would not have been possible without their involvement. We also kindly thank all the authors who dedicated much of their time and efforts to contribute to BIOTECHNO 2025.

Also, this event could not have been a reality without the support of many individuals, organizations, and sponsors. We are grateful to the members of the BIOTECHNO 2025 organizing committee for their help in handling the logistics and for their work to make this professional meeting a success.

We hope that BIOTECHNO 2025 was a successful international forum for the exchange of ideas and results between academia and industry and for the promotion of progress in the fields of bioinformatics, biocomputational systems and biotechnologies.

We are convinced that the participants found the event useful and communications very open. We also hope that Nice provided a pleasant environment during the conference and everyone saved some time for exploring this beautiful city

BIOTECHNO 2025 Chairs:

BIOTECHNO 2025 Steering Committee
Petre Dini, IARIA, USA/EU
Hesham H. Ali, University of Nebraska at Omaha, USA
Xiaobo Zhou, University of Texas Health Science Center – Houston, USA

BIOTECHNO 2025 Publicity Chairs
Francisco Javier Díaz Blasco, Universitat Politècnica de València, Spain
Ali Ahmad, Universitat Politècnica de València, Spain
Sandra Viciano Tudela, Universitat Politecnica de Valencia, Spain
José Miguel Jiménez, Universitat Politecnica de Valencia, Spain

BIOTECHNO 2025

Committee

BIOTECHNO 2025 Steering Committee
Petre Dini, IARIA, USA/EU
Hesham H. Ali, University of Nebraska at Omaha, USA
Xiaobo Zhou, University of Texas Health Science Center – Houston, USA

BIOTECHNO 2025 Publicity Chairs
Francisco Javier Díaz Blasco, Universitat Politècnica de València, Spain
Ali Ahmad, Universitat Politècnica de València, Spain
Sandra Viciano Tudela, Universitat Politecnica de Valencia, Spain
José Miguel Jiménez, Universitat Politecnica de Valencia, Spain

BIOTECHNO 2025 Technical Program Committee

Behrooz Abbaszadeh, University of Ottawa, Canada
Antonino Abbruzzo, University of Palermo, Italy
A M Abirami, Thiagarajar College of Engineering, Madurai, India
Don Adjeroh, West Virginia University, USA
Aftab Ahmad, City University of New York, USA
Jens Allmer, Hochschule Ruhr West - University of Applied Sciences, Germany
Joel P. Arrais, University of Coimbra, Portugal
Simone Avesani, Università di Verona, Italy
Yoseph Bar-Cohen, Electroactive Technologies / NDEAA Lab - Jet Propulsion Laboratory (JPL), USA
Kais Belwafi, King Saud University, Saudi Arabia
Boubaker Ben Ali, University of Bordeaux, France / University of Manouba, Tunisia
Razvan Bocu, Transilvania University of Brasov, Romania
Vincenzo Bonnici, Università di Parma, Italy
Paolo Cazzaniga, University of Bergamo, Italy
Peter Clote, Boston College, USA
Matteo Comin, University of Padova, Italy
Giovanni Cugliari, University of Turin, Italy
Adithi Deborah Chakravarthy, University of Nebraska at Omaha, USA
Maria Evelina Fantacci, University of Pisa, Italy
Abdolhossein Fathi, Razi University, Iran
Silvana Giuliatti, University of São Paulo, Brazil
Henry Griffith, University of Texas at San Antonio, USA
Md Shafaeat Hossain, Southern Connecticut State University, USA
Chih-Cheng Hung, Kennesaw State University, USA
Dimitrios G. Katehakis, Center for eHealth Applications and Services FORTH | Institute of Computer Science, Greece
Jan Kubicek, VSB - Technical University of Ostrava, Czech Republic
Antonio LaTorre, Universidad Politécnica de Madrid, Spain
Chen Li, Monash University, Australia

Copyright Information

For your reference, this is the text governing the copyright release for material published by IARIA.

The copyright release is a transfer of publication rights, which allows IARIA and its partners to drive the dissemination of the published material. This allows IARIA to give articles increased visibility via distribution, inclusion in libraries, and arrangements for submission to indexes.

I, the undersigned, declare that the article is original, and that I represent the authors of this article in the copyright release matters. If this work has been done as work-for-hire, I have obtained all necessary clearances to execute a copyright release. I hereby irrevocably transfer exclusive copyright for this material to IARIA. I give IARIA permission or reproduce the work in any media format such as, but not limited to, print, digital, or electronic. I give IARIA permission to distribute the materials without restriction to any institutions or individuals. I give IARIA permission to submit the work for inclusion in article repositories as IARIA sees fit.

I, the undersigned, declare that to the best of my knowledge, the article is does not contain libelous or otherwise unlawful contents or invading the right of privacy or infringing on a proprietary right.

Following the copyright release, any circulated version of the article must bear the copyright notice and any header and footer information that IARIA applies to the published article.

IARIA grants royalty-free permission to the authors to disseminate the work, under the above provisions, for any academic, commercial, or industrial use. IARIA grants royalty-free permission to any individuals or institutions to make the article available electronically, online, or in print.

IARIA acknowledges that rights to any algorithm, process, procedure, apparatus, or articles of manufacture remain with the authors and their employers.

I, the undersigned, understand that IARIA will not be liable, in contract, tort (including, without limitation, negligence), pre-contract or other representations (other than fraudulent misrepresentations) or otherwise in connection with the publication of my work.

Exception to the above is made for work-for-hire performed while employed by the government. In that case, copyright to the material remains with the said government. The rightful owners (authors and government entity) grant unlimited and unrestricted permission to IARIA, IARIA's contractors, and IARIA's partners to further distribute the work.

Table of Contents

Cardiomyopathy Mechanisms: Meta-analysis of Expression Profiles for Z-disc-associated Genes Across Multiple Microarray Datasets

Yulia Einav and Steve Solun
Faculty of Engineering and Digital Medical Technologies, Holon Institute of Technology
Holon, Israel
Email: yulia_e@hit.ac.il, solunsteve@gmail.com

***Abstract*–Cardiomyopathies (CM) encompass ischemic and non-ischemic subtypes, with evidence increasingly supporting the involvement of Z-disc-associated genes in their pathogenesis. This study conducted a meta-analysis of ten such genes across six microarray datasets, including 142 samples, to explore their expression patterns in CM. Statistical analyses revealed significantly higher expression levels in CM groups compared to controls, with ischemic CM showing greater variability and heterogeneity than non-ischemic CM. Differential expression analysis highlighted substantial downregulation of Filamin C in non-ischemic CM and significant upregulation of PDZ and LIM domain proteins in ischemic CM. These findings enhance our understanding of the molecular mechanisms underlying CM and suggest potential therapeutic targets.**

Keywords-Cardiomyopathies; Z-disc; Gene expression

I. INTRODUCTION

Cardiomyopathies (CM) are a diverse group of cardiac disorders, broadly categorized into non-ischemic forms, primarily caused by genetic factors or non-coronary conditions, and ischemic subtypes, resulting from reduced myocardial blood flow due to coronary artery disease [1]. The Z-disc, a crucial sarcomeric structure, maintains cardiac muscle integrity and function, serving as an anchor point for thin filaments and a mechanotransduction signaling hub [2]. Several Z-disc-associated genes have been implicated in various cardiomyopathy subtypes [2][3]. To date, no systematic analysis has been conducted on the expression levels of these genes across different cardiomyopathy subtypes.

II. RESULTS

This study focused on ten genes selected based on two criteria: structural or functional association with the Z-disc and published involvement in cardiomyopathy in human and/or animal models. Their expression patterns were compared across control, non-ischemic CM, and ischemic CM to elucidate their role in different CM subtypes. We employed a meta-analysis approach, integrating data from six microarray datasets (Table I). Expression data from 3,882 genes shared across 142 samples were retrieved from GEO, normalized, and merged using the ComBat function (sva package). Differential expression analysis was conducted using the limma package in R.

TABLE I. SUMMARY OF GEO SERIES AND SAMPLES IN THE STUDY

GEO Series	Control	Non-Ischemic CM	Ischemic CM	Total Samples
GDS651 / GSE1145	11	15	11	37
GDS1362 / GSE1869	6	21	10	37
GDS2205 / GSE3585	5	7	-	12
GDS2206 / GSE3586	15	13	-	28
GDS3115 / GSE9128	3	4	4	11
GDS4772 / GSE42955	5	12	-	17

Both CM subtypes showed distinct expression patterns compared to the control group, with some variations between subtypes. Statistical analyses revealed significantly higher median expression values in CM groups, highest in ischemic CM, intermediate in non-ischemic CM, and lowest in controls ($p < 2.2e\text{-}16$).

The ten target genes mirrored these overall trends. Analyzing these genes separately showed narrow interquartile ranges (IQRs) in non-ischemic CM samples, while ischemic CM exhibited wider IQRs, validating the selection of these genes as representative markers. Fold change analysis ($p < 0.05$, $|\log2Fc| > 0.5$) revealed substantial downregulation of Filamin C and slight upregulation of PDZ and LIM domain protein 3 in non-ischemic CM. In ischemic CM, Alpha-actinin-2 and PDZ and LIM domain protein 5 were moderately upregulated, while FH1/FH2 domain-containing protein 3, LIM domain-binding protein 3, and PDZ and LIM domain protein 3 showed the most profound upregulation.

Multiple mutations in the target genes of this study have been previously associated with cardiomyopathies in the literature. A notable example is the Filamin C gene, where mutations have been linked to non-ischemic cardiomyopathy [4]. Our meta-analysis revealed substantial downregulation of Filamin C in non-ischemic

 ISBN: 978-1-68558-246-3

CM samples, aligning with clinical observations. This concordance validates our meta-analysis approach and underscores the potential clinical relevance of our findings.

III. CONCLUSION

This study represents the first comprehensive meta-analysis of Z-disc-associated gene expression across multiple CM subtypes, providing a unique perspective on the role of these genes in CM pathogenesis. Overall, gene expression levels differed significantly among the three groups, with several Z-disc-associated genes exhibiting notable up- or down-regulation in either non-ischemic or ischemic CM compared to the control group. Our comprehensive description of gene expression distributions, coupled with the differential expression analysis of ten target Z-disc-associated genes across multiple microarray datasets, provides valuable insights into the pathogenesis of cardiomyopathies. These findings may contribute to the identification of novel therapeutic targets and enhance our understanding of the molecular mechanisms underlying different forms of cardiomyopathy.

REFERENCES

[1] P. M. Seferović, et al., "Genetic basis of cardiomyopathies," in The ESC Textbook of Heart Failure, The European Society of Cardiology Series, Oxford, 2023.

[2] D. Frank and N. Frey, "Cardiac Z-disc signaling network," J Biol Chem, vol. 286, no. 12, pp. 9897-9904, 2011.

[3] M. L. Bang, J. Bogomolovas, and J. Chen, "Understanding The Molecular Basis of Cardiomyopathy," Am J Physiol Heart Circ Physiol, vol. 322, no. 2, pp. H181-H233, 2021.

[4] J. A. J. Verdonschot, et al., "A mutation update for the FLNC gene in myopathies and cardiomyopathies," Hum Mutat, vol. 41, no. 6, pp. 1091-1111, 2020.

A Four-Dimensional Mathematical Model for DNA/RNA Classification

Dorota Bielińska-Wąż
Department of Radiological Informatics and Statistics
Medical University of Gdańsk
80-210 Gdańsk, Poland
email: djwaz@gumed.edu.pl

Piotr Wąż
Department of Nuclear Medicine
Medical University of Gdańsk
80-210 Gdańsk, Poland
email: phwaz@gumed.edu.pl

***Abstract*—This document reviews the 4D-Dynamic Representation of DNA/RNA Sequences, a four-dimensional bioinformatics model developed and published by the authors for classifying deoxyribonucleic acid (DNA) and ribonucleic acid (RNA) sequences.**

Keywords–bioinformatics; alignment-free methods; descriptors

I. Introduction

Alignment-free bioinformatics methods represent a rapidly developing field and provide an efficient alternative to traditional alignment-based approaches [1]–[6]. Among alignment-free methods are Graphical Representations of Biological Sequences, which focus on the graphical and numerical analysis as well as the classification of biological sequences, with comprehensive reviews available in [7]–[10]. Each approach highlights different aspects of sequence similarity. This document describes the 4D-Dynamic Representation of DNA/RNA Sequences, an alignment-free method developed and published by the authors [11] [12]. While multidimensional in nature, the projections of its four-dimensional graphs into two- and three-dimensional spaces serve as graphical tools for analyzing sequence similarity. This method extends our earlier two-dimensional [13]–[19] and three-dimensional [20]–[23] approaches. We refer to this method as "dynamic" because it models sequences as clouds of material points, maintaining constant distances between one another, similar to the behavior of a rigid body in classical dynamics. The numerical characteristics of these clouds are designed to be analogous to those used in dynamics. Details of the method are provided in Section II, while the summary and our plans are presented in Section III.

II. Method and Results

In this approach, a DNA/RNA sequence is represented as a *4D-dynamic graph* - a collection of material points with unit masses positioned in four-dimensional space. The methodology for constructing this graph is detailed in [11]. As descriptors of the 4D-Dynamic Representation of DNA/RNA Sequences, we proposed using the coordinates of the centers of mass of these 4D-dynamic graphs

$$\mu^k = \frac{1}{N}\sum_{i=1}^{N} x_i^k, \quad k = 1, 2, 3, 4 \tag{1}$$

and using the normalized principal moments of inertia of the 4D-dynamic graphs

$$r_k^{4D} = \sqrt{\frac{I_k}{N}}. \tag{2}$$

N is the length of the sequence which is equal to the total mass of the 4D-dynamic graph ($m_i = 1$ of each material point):

$$N = \sum_{i=1}^{N} m_i. \tag{3}$$

The moment of inertia tensor in four-dimensional space is:

$$\hat{I} = \begin{pmatrix} I_{1\,1} & I_{1\,2} & I_{1\,3} & I_{1\,4} \\ I_{2\,1} & I_{2\,2} & I_{2\,3} & I_{2\,4} \\ I_{3\,1} & I_{3\,2} & I_{3\,3} & I_{3\,4} \\ I_{4\,1} & I_{4\,2} & I_{4\,3} & I_{4\,4} \end{pmatrix} \tag{4}$$

with the matrix elements

$$I_{j\,j} = \sum_{i=1}^{N} m_i \sum_{k=1}^{4} \left[\hat{x}_i^k (1 - \delta_{jk})\right]^2, \tag{5}$$

$$I_{j\,k} = I_{k\,j} = -\sum_{i=1}^{N} m_i \hat{x}_i^j \hat{x}_i^k. \tag{6}$$

The m_i coordinates in the new coordinate system are $\hat{x}_i^k = x_i^k - \mu^k$ and δ_{jk} is the Kronecker-Delta. The new system is a Cartesian coordinate system in which the origin is chosen at the center of mass of the graph. The eigenvalues I_k used in (2), called the principal moments of inertia, are obtained by solving the fourth-order secular equation:

$$\det(\hat{I} - I\hat{E}) = 0, \tag{7}$$

where $\hat{E}$ is the unit matrix.

The proposed descriptors have been applied to the construction of the similarity maps. These descriptors are represented on the axes of the maps. Examples of the similarity maps obtained using different descriptors are shown in Figures 1–3.

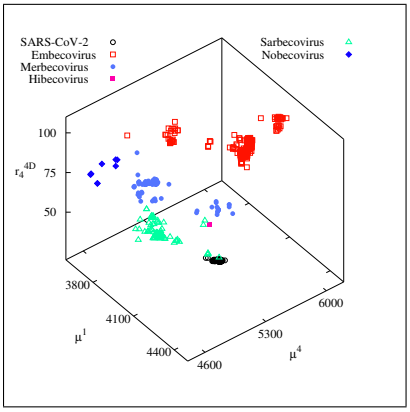

Figure 1. Classification map $\mu^1 - \mu^4 - r_4^{4D}$.

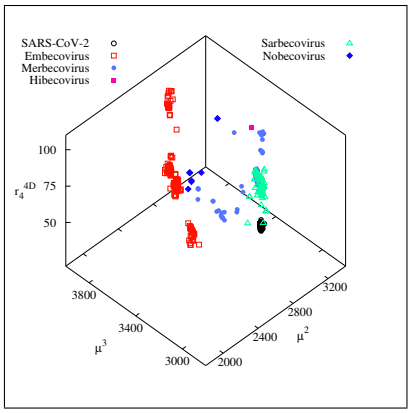

Figure 2. Classification map $\mu^2 - \mu^3 - r_4^{4D}$.

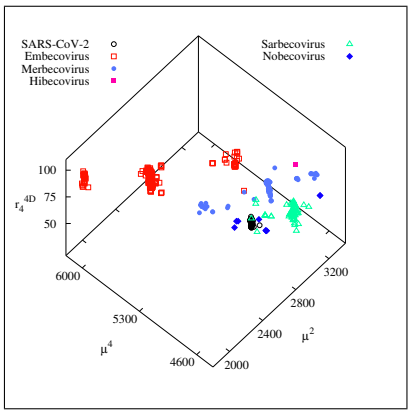

Figure 3. Classification map $\mu^2 - \mu^4 - r_4^{4D}$.

III. Conclusion

We applied this approach to the bioinformatics characterization of the SARS-CoV-2 virus [11] and to studies on the genetic diversity of *Echinococcus multilocularis* in red foxes in Poland [12]. Specifically, the distribution of clusters in the classification maps generated using the 4D-Dynamic Representation of DNA/RNA Sequences supports the hypothesis that SARS-CoV-2 may have originated in bats and pangolins [11]. Our results align with those obtained using standard bioinformatics methods for the *Echinococcus multilocularis* [24] and SARS-CoV-2 [25] datasets. In our future work, we plan to introduce new graph descriptors and apply them to the characterization of other viruses.

References

[1] M.K. Gupta, R. Niyogi, and M. Misra, "An alignment-free method to find similarity among protein sequences via the general form of Chou's pseudo amino acid composition", SAR QSAR Environ. Res. vol. 24, pp. 597–609, 2013.

[2] Y.S. Li, T. Song, J.S. Yang, Y. Zhang, and J.L. Yang, "An Alignment-Free Algorithm in Comparing the Similarity of Protein Sequences Based on Pseudo-Markov Transition Probabilities among Amino Acids", PLOS ONE vol. 11, Art. No. e0167430, 2016.

[3] A.K. Saw, B.C Tripathy, and S. Nandi, "Alignment-free similarity analysis for protein sequences based on fuzzy integral", Sci. Rep. vol. 9, Art. No. 2775, 2019.

[4] Y. Zhao, X. Xue, and X. Xie, "An alignment-free measure based on physicochemical properties of amino acids for protein sequence comparison", Comput. Biol. Chem. vol. 80, pp. 10–15, 2019.

[5] N. Ramanathan, J. Ramamurthy, and G. Natarajan, "Numerical Characterization of DNA Sequences for Alignment-free Sequence Comparison - A Review", Comb. Chem. High Throughput Screen. vol. 25, pp. 365–380, 2022.

[6] K.E. Wade, L. Chen, C. Deng, G. Zhou, and P. Hu, "Investigating alignment-free machine learning methods for HIV-1 subtype classification", Bioinform. adv. vol. 4, Art. No. vbae108, 2024.

[7] A. Nandy, M. Harle, and S. C. Basak, "Mathematical descriptors of DNA sequences: development and applications", Arkivoc vol. ix, pp. 211–238, 2006.

[8] D. Bielińska-Wąż, "Graphical and numerical representations of DNA sequences: Statistical aspects of similarity", J. Math. Chem. vol. 49, pp. 2345–2407, 2011.

[9] M. Randić, M. Novič, and D. Plavšić, "Milestones in Graphical Bioinformatics", Int. J. Quant. Chem. vol. 113, pp. 2413–2446, 2013.

[10] S. Mizuta, "Graphical Representation of Biological Sequences", In Bioinformatics in the Era of Post Genomics and Big Data; I.Y. Abdurakhmonov, Ed.; IntechOpen: London, UK, 2018.

[11] D. Bielińska-Wąż and P. Wąż, "Non-standard bioinformatics characterization of SARS-CoV-2", Comput. Biol. Med. vol. 131, Art. No. 104247, 2021.

[12] D. Bielińska-Wąż, P. Wąż, A. Lass, and J. Karamon, "4D-Dynamic Representation of DNA/RNA Sequences: Studies on Genetic Diversity of *Echinococcus Multilocularis* in Red Foxes in Poland", Life vol. 12, Art. No. 877, 2022.

[13] D. Bielińska-Wąż, T. Clark, P. Wąż, W. Nowak, and A. Nandy, "2D-dynamic representation of DNA sequences", Chem. Phys. Lett. vol. 442, pp. 140–144, 2007.

[14] D. Bielińska-Wąż, W. Nowak, P. Wąż, A. Nandy, and T. Clark, "Distribution moments of 2D-graphs as descriptors of DNA sequences", Chem. Phys. Lett. vol. 443, pp. 408–413, 2007.

[15] D. Bielińska-Wąż, P. Wąż, and T. Clark, "Similarity studies of DNA sequences using genetic methods", Chem. Phys. Lett. vol. 445, pp. 68–73, 2007.

[16] P. Wąż, D. Bielińska-Wąż, and A. Nandy, "Descriptors of 2D-dynamic graphs as a classification tool of DNA sequences", J. Math. Chem. vol. 52, pp. 132–140, 2014.

[17] A. Nandy, S. Dey, S.C. Basak, Bielińska-Wąż, and P. Wąż, "Characterizing the Zika Virus Genome - A Bioinformatics Study", Curr. Comput. Aided Drug Des. vol. 12, pp. 87–97, 2016.

[18] D. Panas, P. Wąż, D. Bielińska-Wąż, A. Nandy, and S.C. Basak, "2D-Dynamic Representation of DNA/RNA Sequences as a Characterization Tool of the Zika Virus Genome", MATCH Commun. Math. Comput. Chem. vol. 77, pp. 321–332, 2017.

[19] D. Panas, P. Wąż, D. Bielińska–Wąż, A. Nandy, and S.C. Basak. "An Application of the 2D-Dynamic Representation of DNA/RNA Sequences to the Prediction of Influenza A Virus Subtypes", MATCH Commun. Math. Comput. Chem. vol. 80, pp. 295–310, 2018.

[20] P. Wąż and D. Bielińska-Wąż, "3D-dynamic representation of DNA sequences", J. Mol. Model. vol. 20, Art. No. 2141, 2014.

[21] P. Wąż and D. Bielińska-Wąż, "Non-standard similarity/dissimilarity analysis of DNA sequences", Genomics vol. 104, pp. 464–471, 2014.

[22] D. Bielińska-Wąż, D. Panas, and P. Wąż, "Dynamic representations of biological sequences", MATCH Commun. Math. Comput. Chem. vol. 82, pp. 205–218, 2019.

[23] D. Bielińska-Wąż, P. Wąż, and D. Panas, "Applications of 2D and 3D-Dynamic Representations of DNA/RNA Sequences for a Description of Genome Sequences of Viruses", Comb. Chem. High Throughput Screen. vol. 25, pp. 429–438, 2022.

[24] J. Karamon, K. Stojecki, M. Samorek-Pieróg, E. Bilska-Zając, M. Różycki, E. Chmurzyńska, J. Sroka, J. Zdybel, and T. Cencek, "Genetic diversity of *Echinococcus multilocularis* in red foxes in Poland: The first report of a haplotype of probable Asian origin", Folia Parasitol. vol. 64, Art. No. 007, 2017.

[25] S.K. Gupta, R. Minocha, P.J. Thapa, M. Srivastava, and T. Dandekar, "Role of the Pangolin in Origin of SARS-CoV-2: An Evolutionary Perspective", Int. J. Mol. Sci. vol. 23, Art. No. 9115, 2022.

A Computational Tool Supporting the Diagnosis of Age-Related Macular Degeneration

Piotr Wąż
Department of Nuclear Medicine
Medical University of Gdańsk
80-210 Gdańsk, Poland
email: phwaz@gumed.edu.pl

Dorota Bielińska-Wąż
Department of Radiological Informatics and Statistics
Medical University of Gdańsk
80-210 Gdańsk, Poland
email: djwaz@gumed.edu.pl

Katarzyna Zorena
Department of Immunobiology and Environment Microbiology
Medical University of Gdańsk
80-210 Gdańsk, Poland
email: kzorena@gumed.edu.pl

Anna Murawska
Department of Immunobiology and Environment Microbiology
Medical University of Gdańsk
80-210 Gdańsk, Poland
email: murawska.anna@gmail.com

***Abstract*—This work describes a diagnostic graphical tool (classification maps) recently developed and published by the authors to support the detection of Age-Related Macular Degeneration (AMD). These maps, constructed using an ordinal regression model, visually represent the progression of AMD. In this model, the degree of AMD advancement serves as the ordinal dependent variable. Independent variables, such as Central Retinal Thickness (CRT), Ganglion Cell Complex thickness (GCC), Macular Pigment Optical Density (MPOD), Early Treatment Diabetic Retinopathy Study (ETDRS) scores, Snellen visual acuity, and patient age are incorporated into the analysis and displayed on the axes of the maps.**

Keywords–medical informatics; biostatistics; mathematical modeling

I. INTRODUCTION

Central vision loss caused by Age-Related Macular Degeneration (AMD) is a significant global health concern, particularly among individuals over 60 years of age. AMD accounts for nearly 50% of cases of blindness as defined by legal standards [1] [2]. This chronic, progressive condition affects the outer layers of the central retina and the choroid. The prevalence of AMD is expected to grow due to increasing life expectancy and heightened exposure to risk factors contributing to degenerative changes in the macula [1] [3].

In the United States, approximately 8 million individuals are diagnosed with early AMD, and more than 1 million are predicted to develop advanced AMD within the next five years [2]. By 2050, one in ten Americans over the age of 50 is projected to have AMD [3]. In Poland, an estimated 1.5 million people suffer from macular degeneration, including 130000 cases of the more severe exudative form of AMD [4]. Globally, the number of AMD cases is expected to rise from 196 million in 2020 to 288 million by 2040 [5]. Europe currently has the highest AMD incidence, with future case numbers anticipated to be surpassed only by Asia [6]. AMD prevalence rates vary across regions, with 12.33% in Europe, 7.38% in Asia, and 7.53% in Africa [5] [6].

The risk of AMD increases with age. Among individuals aged 60 and older, 13.4% are affected, compared to 2.8% among those aged 40–59. Advanced AMD is present in 1.4% of individuals at age 70, increasing to 5.6% at age 80, and 20% at age 90 [7]. The study of AMD has deep historical roots. The advent of ophthalmoscopy, the examination of the eye's fundus, marked the beginning of modern retinology. Hermann von Helmholtz's invention of the ophthalmoscope in 1851 was a pivotal milestone. The macula was first identified in detail in the late 18th century by Samuel Thomas Soemmerring, who described the yellowish area in the posterior retina. In 1875, Jonathan Hutchinson and Warren Tay documented symmetrical fundus changes, later classified as "senile macular degeneration" by Otto Haab in 1885 [8].

While progress in AMD research was initially gradual, it focused on identifying risk factors and developing early classifications. A major advancement occurred in 1967 when J. Donald Gass from Florida described the pathogenesis of central vision loss and delineated the stages of AMD progression [9] [10].

In this work, we review a new computational method, recently developed and published by us [11], for supporting the detection of AMD. Details of this method are outlined in Section II, whereas Section III provides a summary and discusses our future plans.

II. METHOD AND RESULTS

The study involved the examination of 132 eyes from 66 patients, classified according to AMD progression using the four-point Age-Related Eye Disease Scale (AREDS) [12]. As the most widely utilized system for categorizing AMD, the AREDS scale divides AMD progression into the following stages:

1) AREDS 1 (control group): Absence of AMD or presence of only a few small drusen ($< 63\mu m$ in diameter).
2) AREDS 2: Early AMD, characterized by numerous small drusen ($> 15\mu m$), several intermediate-sized drusen ($63 - 125\mu m$), or Retinal Pigment Epithelium (RPE) abnormalities, such as increased pigmentation or depigmentation.
3) AREDS 3: Intermediate AMD, including numerous medium-sized drusen, at least one large druse ($> 125\mu m$), or geographic atrophy not involving the central macula.
4) AREDS 4: Advanced AMD, involving geographic atrophy of the RPE affecting the macula or neovascular maculopathy. This includes Choroidal Neovascularization (CNV), serous or hemorrhagic retinal

or RPE detachment, exudates and hard fibrovascular proliferations beneath the retina and RPE, and discoid scars (choroidal fibrosis) [13] [14].

Of the examined eyes, 32 were classified as AREDS 1, showing no AMD features and serving as the control group due to the absence of visible fundus changes. The remaining eyes were distributed among the AREDS 2 (37 eyes), AREDS 3 (33 eyes), and AREDS 4 (30 eyes) groups.

In certain patients, AMD-related changes were present in only one eye, or the severity varied between eyes, as determined by the AREDS classification.

Outpatient examinations were conducted in 2016–2017 at the Ophthalmology Clinic of UCK in Gdańsk, Poland. Diagnoses were established following the standards and recommendations of the Polish Society of Ophthalmology, consistent with the American Academy of Ophthalmology (AAO) guidelines [15]. Data regarding disease stage and overall health were gathered through medical history, the Simplified Théa AMD Risk-assessment Scale (STARS), resting blood pressure and pulse measurements, and detailed ophthalmological assessments. These assessments included Spectral Optical Coherence Tomography (SOCT) macula scans and Macular Pigment Optical Density (MPOD) measurements. Patients were fully informed about the tests conducted.

Inclusion criteria for the AMD and control groups were an age of over 55 years and an AMD diagnosis aligned with the AREDS scale and Polish Society of Ophthalmology guidelines.

Statistical analyses were conducted using the R programming language [16]. Quantitative variables were summarized using median, minimum, and maximum values. The Kruskal-Wallis test was applied to compare variables across AMD advancement groups, with post-hoc tests following significant results.

Ordinal regression, implemented via a generalized linear model, was utilized to predict ordinal variables, where only relative ordering is important. The dependent variable was the AMD advancement level. Independent variables included Central Retinal Thickness (CRT), average Ganglion Cell Complex (GCC) thickness, MPOD, ETDRS visual acuity, Snellen visual acuity, and patient age. Measurements were performed using the Zeiss Cirrus HD-OCT model 400.

Logit functions were used in the analysis. Since, in this case, the maximization of probability (or the logarithm of probability) does not have an analytical solution, the iteratively reweighted least squares (IRLS) technique was employed to estimate the regression coefficients. The models included all the collected values for each variable. The quality of the models was assessed by evaluating the statistical significance of the coefficients, the -2 log-likelihood value, and the frequency of correctly predicted categories based on the values of the independent variables.

This modeling resulted in a tool for estimating AMD progression. Due to the limited sample size, models were restricted to pairs of independent variables. The article highlights pairs of variables with statistically significant regression coefficients, alongside their Odds Ratios (OR), Confidence Intervals (CI), and thresholds for ordinal categories.

The statistical significance level was set at $\alpha = 0.05$.

Using the proposed model, classification maps were generated, serving as a graphical tool to support the diagnosis of AMD [11]. These maps enable the classification of patients' eyes into specific groups (control group, AREDS 2, AREDS 3, or AREDS 4) based on the values of variables represented on their axes. Classification maps, based on these models, visually represent the predicted probabilities for each AMD stage, using a color-coded scheme. This alternative computational approach facilitates the accurate diagnosis of all stages of AMD with high or good precision.

III. Conclusion

The method described enables the classification of a patient's eyes into specific stages of AMD. The variable pairs displayed on the map axes serve as diagnostic markers essential for determining the stage of AMD.

In our future work, we plan to adapt this method to describe other diseases, such as dementia severity, degrees of obesity, and more.

References

[1] M. Fleckenstein et al., "Age-related macular degeneration", Nat. Rev. Dis. Primers. vol. 7, Art. No. 31, 2021.

[2] Age-related Eye Disease Study Research Group, "Potential public health impact of age-related disease study results: AREDS Report No. 11", Arch. Ophthalmol. vol. 121, pp. 1621–1624, 2003.

[3] K. Pennington and M. DeAngelis, "Epidemiology of age-related macular degeneration (AMD): asssociations with cardiovascular disease phenotypes and lipid factors", Eye and Vision vol. 3, Art. No. 34, 2016.

[4] O. Partyka and M.J. Wysocki, "Epidemiology of eye diseases and infrastructure of ophthalmology in Poland", Przegl. Epidemiol. vol. 69, pp. 905–908, 2015.

[5] Y. Deng, L. Qiao, M. Du, C. Qu, L. Wan, J. Li, and L. Huang, "Age-related macular degeneration: Epidemiology, genetics, pathophysiology, diagnosis, and targeted therapy", Genes Dis. vol. 9, pp. 62–79, 2021.

[6] W.L. Wong et al. "Global prevalence of age-related macular degeneration and disease burden projection for 2020 and 2040: a systematic review and meta-analysis", Lancet Glob Health. vol. 2, pp. 106–116, 2014.

[7] A.R. Rudnicka, Z. Jarrar, R. Wormald, D.G. Cook, A. Fletcher, and C.G. Owen, "Age and gender variations in age-related macular degeneration prevalence in populations of European ancestry: a meta-analysis", Ophthalmology vol. 119, pp. 571–580, 2012.

[8] V. McConnell and G. Silvestri, "Age-related macular degeneration", Ulster Med. J. vol. 74, pp. 82–92, 2005.

[9] J.D. Gass, "Drusen and disciform macular detachment and degeneration", Arch. Ophthalmol. vol. 90, pp. 206–217, 1973.

[10] J.D. Gass, "Pathogenesis of disciform detachment of the neuroepithelium", Am. J. Ophthalmol. vol. 63, pp. 573–585, 1967.

[11] P. Wąż, K. Zorena, A. Murawska, and D. Bielińska-Wąż, "Classification Maps: A New Mathematical Tool Supporting the Diagnosis of Age-Related Macular Degeneration", J. Pers. Med. vol. 13, Art. No. 1074, 2023.

[12] F.L. Ferris et al., "Age-Related Eye Disease Study (AREDS) Research Group. A simplified severity scale for age-related macular degeneration: AREDS Report No. 18", Arch Ophthalmol. vol. 123, pp. 1570–1574, 2005.

[13] Age-Related Eye Disease Study Research Group, "The Age-Related Eye Disease Study system for classifying age-related macular degeneration from stereoscopic color fundus photographs: the age-related eye disease study report number 6", Am. J. Ophthalmol. vol. 132, pp. 668–681, 2001.

[14] M.D. Davis et al., "The Age-Related Eye Disease Study Severity Scale for Age-Related Macular Degeneration: AREDS Report No. 17", Arch. Ophthalmol. vol. 123, pp. 1484–1498, 2005.

[15] Z. Zagórski, "Podsumowanie założeń dla zalecanych algorytmów postępowania w praktyce klinicznej (Preferred Practice Pattern® – PPP)", American Academy of Ophthalmology, 2016, Polish translation 2017.

[16] R Core Team. R: A language and environment for statistical computing. In R Foundation for Statistical Computing; R Core Team: Vienna, Austria, 2023. Available online January, 2025: https://www.R-project.org/