SHORT THESIS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY (PHD)

New approaches for the analysis of health care data

by Szilvia Harsányi

Supervisor: Dr. Judit Zsuga, full professor

UNIVERSITY OF DEBRECEN

DOCTORAL SCHOOL OF HEALTH SCIENCES

DEBRECEN, 2022

New approaches for the analysis of health care data

by Szilvia Harsányi (MSc in Health Policy)

Supervisor: Dr. Judit Zsuga, PhD

Doctoral School of Health Sciences, University of Debrecen

Head of the **Examination Committee**: György Paragh, PhD, DSc

Members of the Examination Committee: Veronika Fenyves, PhD
Ajándék Eőry, PhD

The examination takes place in lecture hall of Department of Emergency Care and Oxyology Faculty of Medicine, University of Debrecen on 14th of December 2022, 13.00

Head of the **Defense Committee**: György Paragh, PhD, DSc

Reviewers: Klára Boruzs, PhD
Krisztián Vörös, PhD

Members of the Defense Committee: Fenyves Veronika, PhD
Ajándék Eőry, PhD

The PhD Defense takes place in the lecture hall of Department of Emergency Care and Oxyology, Faculty of Medicine, University of Debrecen on 14th of December 2022, 14.00

LIST OF CONTENT

1. INTRODUCTION

The collection and analysis of health data has been an important area of health research for centuries. In his world-famous study ("Grand Experiment"), John Snow with the collection and analysis of health data during the London cholera epidemic in the 1850s discovered that geographic representation of cholera-contaminated and clean wells could help to curb the epidemic and improve health of the population. This innovative method of data representation was a milestone and in today's digital world, we often use visualization methods to explore and better understand the contexts.

The research results obtained with new methods can improve the quality and effectiveness of patient care in our time as well. In today's fast-changing world, more and more approaches are being used in health care to analyze health data. New methods are important for evaluating and developing health care systems, which are constantly facing new challenges. The methods used for the quantitative and qualitative data analysis can also be used in healthcare and these methods and the technology undewent great developments in recent decades.

The quality of health care has been defined in several ways and quality has different meanings for the members of society, decision-makers and health professionals. In Donabedian's classical model, he evaluates the quality of health care based on *structure, process* and *outcome* dimensions.

Although the definitions of the concepts of quality of care are not universal, there are some common components (despite their different definitions) that are decisive for characterizing the quality of health care. These are *accessibility, efficacy, efficiency, safety and effectiveness*.

When organizing *patient pathways,* as many dimensions of quality as possible should be considered, the most important are *accessibility, efficiency and effectiveness.* In hospital units, patient pathways need to be organized with *efficiency, safety and effectiveness*.

Achieving positive health outcomes at the population level require translating the scientific evidence into clinical practice. One of the most important steps to extend the interventions when clinical practice guidelines are being formulated to create new recommendations that this should be supported by clinical evidence, which are generally applicable in the medical field.

Stroke has been one of the three most common causes of death in both sexes for decades and a leading cause of long-term disability. There are 40-50 thousand new stroke cases every year, about 180.000 people have some kind of residual symptom of this disease in Hungary.

There has seen a paradigm shift in the treatment of acute stroke in the last decades, with stroke being recognized as an emergency. This was embodied in the formulation of the “time is brain” care guideline, in parallel with the development of the intravenous thrombolysis care protocol. Intravenous administration of recombinant tissue plasminogen activator (rtPa) can be used under certain conditions in ischemic stroke, one of the most important limiting factors being the 4.5-hour time window. This triggered the development of patient pathway management protocols focusing on pre-hospital care, emergency medical care, and hospital management to minimize stroke-to-center and door-to-needle times at the center.

The emergency approach in the case of most diseases, including stroke, requires a relatively new triage approach, the identification of new methods and their transposition into everyday patient care. Important elements of this are the optimization of patient pathways, the study of the network of patient pathways and their improvement and efficiency in order to improve the quality of patient care.

The organization of care between primary care and specialist care, as well as inpatient care is crucial in the design of patients' pathways in order to ensure access to more efficient and effective care. Access to health care is influenced by several factors, such as health education, the gatekeeping function of primary care and patient education.

Social-level campaigns for early detection of stroke symptoms have been improved with the timely recognition of stroke symptoms and the call for appropriate help, but it is important to organize ongoing, targeted campaigns and educate physicians and patients. The basic patient education can be achieved by the education of the patients and health care professionals, who should have the up-to-date knowledge of guidelines and the practical application of this knowledge. One of the cornerstones of acute stroke care is that the patient or their relative should call an ambulance immediately when symptoms appear.

One possible area of the information transfer between doctors and patients is in the primary care setting, where patients maintain regular and often lifelong contact with their family doctor (general practitioner- GP, family physician). Stroke prevention might be possible with the care of pre-stroke cerebrovascular risk factors and chronic diseases (eg hypertension, diabetes), GP can share this important information with patients. Translating GPs' new theoretical knowledge about stroke care into everyday practice and using it in pre-hospital care for acute stroke can greatly improve patient organization, which can improve the outcome of the disease and improve patients' quality of life.

During my research we investigated the translation of theoretical knowledge regarding general practitioners (family physicians) with respect to acute stroke care into practice through qualitative text analysis. Qualitative text analysis using digitized data can be faster and more reliable than surveys using traditional methods and free-text responses can better reflect to the respondents' tacit knowledge.

Professional and funding protocols also could influence the patient pathways for the treatment of certain diseases, which pathways may also determine the flow of patients within emergency and inpatient specialist care.

The other main part of my research was the study of patient pathway optimization using a modeling system. With the help of algorithms, we modeled and examined the models of patient transfer between a campus-based university clinics and the possibilities of the development. Analyzing the health care system and then making it more efficient can optimize patient flow processes in inpatient settings, this helps to transfer the patients to the right place for care faster and more efficiently.

Analyzes using newer methods can be important and forward-looking in assessing and improving the knowledge of health care workers and in organizing patient pathways. The methodology for analyzing health data has been expanded with many new possibilities in recent decades and the digital collection and storage of data has made very large amounts of data readily available. There also been many advances in computing and analysis of larger amounts of data. The applicability of these new sources have also generated more and more useful data in recent years. The role of the data from the social media or websites in health data analysis require further research in the future.

2. LITERATURE REVIEW

2.1. Prehospital care for acute stroke

The concept and designation of patient pathways are very diverse in the literature the terms clinical pathway, integrated care pathway and care pathway are also often used.

The process of patient care is an important part of the patient pathway and it is important for the organization and operation of health care systems. The patient's journey usually begins at the GP and ends at the health care institution that might provide the definitive care. Care routes and patient pathways aim to promote evidence-based and guided care, improve the organization and efficiency of care and reduce costs.

80-85% of stroke cases are of ischemic origin (major arterial atherothrombosis, lacunar infarction, cerebral embolization, haemodynamic stroke) and about 15-20% are haemorrhagic (intracerebral or subarachnoid haemorrhage). However, stroke caused by bleeding or non-bleeding cannot be differentiated during pre-hospital care, so the patient should be treated as a candidate for thrombolysis before the results of medical imaging (scans) are available.

According to the emergency approach of acute stroke, it is important to recognize the symptoms early, notify the stroke center, refer and transfer the patients to the stroke center within the time window. The guideline defines the organization of patient pathway in acute stroke, where it is extremely important to call an ambulance and ensure adequate care capacity when acute symptoms are detected. GPs may be among the first pre-hospital health care providers in the acute stroke care, especially in smaller settlements where ambulance services are more difficult to access. The ambulance services usually transport patients to stroke centers for diagnostic impatient care.

Educating the population is important for early detection of stroke symptoms because it is the basis for seeking help at the right time and place. In 2008, the Hungarian Stroke Company organized its first public information and education campaign. “Stroke Day” is also held annually, and with the help of the National Ambulance Service and Hungarian professional societies (the Hungarian Society of Cardiologists, the Hungarian Society of Hypertension and the Hungarian Dietitians Association) to raise awareness of the importance of stroke prevention, early treatment and follow-up.

Although the guidelines do not give GPs an active role in the treatment of acute stroke, they have a responsibility for early detection and educating the population. It is important to shorten the patient pathway so that the pathway from symptom detection to the first specialist is well organized and the indication for intravenous thrombolysis can be established within the time window.

2.2. Investigating the GPs' knowledge of acute stroke - translating guidelines into practice

Given that one of the main tasks of GPs is to treat chronic diseases, many of which are risk factors for stroke, it is important to increase the knowledge of the patients and GPs about acute stroke.

The education of both medical professionals and the public is important and in the latter the role of GPs can be prominent, as they most often care for and recognize stroke risk factor diseases (diabetes, hypertension, etc.) and can play a key role in the proper organization of patient pathways.

The basis for optimizing acute stroke care is stroke awareness, which involves expanding knowledge about stroke, as one of the main causes of delay is the failure to recognize symptoms and the lack of appropriate measures to seek medical attention. Evidence shows that help is usually sought by family members and the first contact is often with the GP, who notifies the ambulance service. Recent guidelines continue to emphasize the need to educate the population about acute stroke care. The main points of the education are the rapid recognition of the symptoms and the need for immediate medical attention and notification of the ambulance service, as well as immediate the transport to the stroke center.

There are also many examples of “good practice” in the implementation of acute stroke thrombolysis internationally, according to which those working in primary stroke centers should perform their work in close coordination with the emergency medical services (EMS).

In acute stroke cases appropriate pre-hospital care improves the effectiveness of subsequent treatment and in patients who do not recognize the symptoms of stroke early, the hospitalization for stroke may be less effective.

In Hungary, the emergency treatment of stroke was first defined in 2004 by guidelines developed by the Hungarian Stroke Society. Subsequently, acute stroke triage was formalized in several service areas, including the extended catchment area of the Clinic of Neurology of the University of Debrecen. Accordingly, a partnership agreement was signed (in 2007) with the National Ambulance Service to transport all thrombolysis candidates directly to the Neurological Clinic in the 90 km area of Debrecen. The program was complemented by ongoing education and training for ambulance service personnel, general practitioners, and stroke ward nurses and physicians.

2.3. Assessing patient flow within the hospital with lean thinking

The essence of lean thinking is to identify the unnecessary elements of the processes and to reduce them, which results in a reduction of the cost-effectiveness and time-consuming nature of the processes. Of the many management concepts used for process redesign, the "lean" thinking transformation approach monitors the usefulness of processes in terms of whether or not they can add value to results in order to eliminate non-value-added activities.

With respect to the lean concept for patient flow logistics, optimal care can be determined by providing it to the right patient in the right place by the right provider and the right information is used at the right time. According to this, the physical distance made by patients is not a value-adding attribute of patient care, so an attempt should be made to minimize it at the system level. In summary, "lean" thinking projected onto our studies is a theoretical framework that can improve processes related to both acute and subacute care. Previous studies have shown that the application of lean thinking improves patient flow in emergency departments and may result in better through the reduced waiting time.

Hospital transformation initiatives to meet the requirements of patient-centered care for patients referred to a hospital or clinics often face financial, cultural, and structural barriers. The potential for optimizing patient flow logistics is often overlooked because patient transport services are seen as complementary, despite the fact that patient-flow logistics have a fundamental impact on the quality of care and hospital costs. Delays in transporting a patient to an operating room or for an medical imaging examination also reduce the utilization of medical equipment and operating rooms, prevent timely execution, and increase patient waiting times. In addition, shortcomings in the organization of patient transport may underlie waiting lists, as well as the unpredictability of the workload, which increases the stress on healthcare workers, leading to errors and inappropriate placement of patients.

Given that the logistics issues of patient transport have a profound impact on many very important areas of care, e.g., on the quality of care, disease outcomes, patient satisfaction and their financial implications, there is a clear need for systemic approaches to address patient transport issues. Improving patient flow will contribute to reducing waiting times, better access to healthcare, efficiency and lower costs, as well as improving indicators directly related to the number of patient transfers and delivery times, the distance traveled by patients.

2.4. OBJECTIVES

1. We aimed to investigate using new methods (qualitative text analysis, word clouds) the extent new guidelines changed the knowledge and clinical practice of general practitioners (GPs) with regards to acute stroke care, whether knowledge of the theory has been translated into clinical practice.

2. We investigated the knowledge, care practices and organization patient pathways of GPs regarding acute stroke.

3. In our studies, with application of algorithms we analyzed the possibilities for optimizing patient pathways with respect to the physical organization of care, with the goal of minimizing non-value adding processes.

4. We also analyzed patient flow using graph theory and spectral graph theory in case of both the original and optimized settings. We investigated which network parameters are relevant for patient pathway organization.

3. MATERIALS AND METHODS

3.1. Putting GPs' knowledge about acute stroke into practice

3.1.1. Study design and protocol

The study protocol was approved by the Regional Ethics Committee of the University of Debrecen (496-2018) and the Ethics Committee of the Health Research Council (51672/2018 / EKU). In our study, we obtained data by taking supervised self-administered questionnaires during the mandatory training courses of the GPs or the residency training program between February 1, 2018, and July 31, 2018. Participation in the study was voluntary and anonymous, and respondents were GPs or GP-family medicine resident doctors (residents). Demographic variables (age, gender), characteristics of the medical practice (general practitioner, resident doctor, other professional specializations, number of years spent in practice, geographical location of the practice (county, postcode, type of settlement), type of practice (adult or mixed practice) were recorded with the respondent connection.

Knowledge of acute stroke care was investigated by analysing responses to two open-ended questions related to clinical case reporting. The case report described the patient's comorbidities, symptoms and time of onset of the symptoms. The case reports were developed by a neurologist with a neurological examination and experienced in the treatment of acute stroke, adapting the methodology used by Fernandes et al. With regard to the investigation of the translation of the emergency approach, the two cases differed only in respect of that time window. The two clinical case reports included symptoms of acute circulatory stroke in the left cerebral artery media area as follows.

Clinical case 1:

Peter, a 54-year-old male patient, has a history of treated hypertension, type 2 diabetes and atrial fibrillation.

Accompanied by his wife, he arrived at his family doctor at 8:30 a.m., with speech problems (difficulty finding words), paralysis of the right face (central facial paresis) and weakness of the upper extremities. The wife says she was still fine last night, the symptoms that morning were about. The symptoms started at 7.30am.

Clinical case 2:

Peter, a 54-year-old male patient, has a history of treated hypertension, type 2 diabetes, and atrial fibrillation.

Accompanied by his wife, he arrived at his family doctor at 8:30 a.m., with speech problems (difficulty finding words), paralysis of the right face (central facial paresis), and weakness of the upper extremities. The wife says that before he gone to bed yesterday, her husband's speech was confused, and the toothbrush cup slipped out of his hand. He then went to bed and was noticing the above symptoms since waking up in the morning.

The following two open-ended questions were used in connection with both case descriptions:

Question 1: What is your diagnosis for the case?

Question 2: What would you say to the patient about his condition and what will happen to him?

In a pilot study, the questionnaires were analyzed on a sample of 20 resident doctors and 20 GPs and then finalized based on the results obtained.

3.1.2. Qualitative text analysis

Free-text answers were entered into a Microsoft Excel spreadsheet and then imported into NVivo12 program to evaluate the frequency of words used for each question and each case individually. NVivo was used to visualize qualitative data. This method allows the evaluation of language use, interpretation and reflection of textual content. Synonyms were merged, filler words and conjunctions were omitted. After translating the list of words into English, the word clouds were compiled using NVivo12. Font sizes were proportional to the frequency of occurrence of the word.

3.1.3. Quantitative analysis

For quantitative analysis, responses were categorized as correct or incorrect, as follows.

The answer to question 1 of clinical case 1 is correct if the following terms are used: stroke, stroke / embolism, intraocular stroke, stroke (ischemic) in the left middle cerebral artery, left stroke, acute stroke, insufficient blood flow to the left middle cerebral artery (MCA), apoplexy, insufficient circulation in the left MCA, cerebral infarction, cerebral embolism in the left hemisphere, cerebrovascular event within 3 hours, cerebral embolism.

The answer to question 1 of clinical case 2 is correct in case of the use of the following terms: stroke, out-of-time stroke (4.5 hours), out-of-time ischemic stroke, left stroke, ischemic stroke in the MCA area, out-of-lysis stroke, insufficient cerebral artery circulation, apoplexy (ischemic), stroke in the MCA

area, cerebral infarction / cerebral embolism (outside the time window), subacute left stroke.
The answers to question 2 were evaluated on the basis of two dimensions, one was the presence or absence of communication of urgency and the other was an appropriate description of the suitability for thrombolysis. Suitability for thrombolysis was considered correct in the first case and incorrect in the second case. The answer was considered accurate if both dimensions were correctly included in the answer to question 2.
In terms of the geographical location of the respondents' practice, we categorized the practices according to whether they fall within or outside the 90 km radius of the University of Debrecen. We have chosen Debrecen as the geographical reference center, as there has been a joint effort in and around Debrecen since 2007 to maintain an effective acute stroke triage system and, according to the agreement, EMS will transport the acute stroke patients within 90 km directly to the Neurology Clinic University of Debrecen.
When assessing the significant predictors required for the correct diagnosis of clinical case 1, a priori variables (age and gender) and significant predictors identified by simple logistic regression were included in the initial model (degree obtained before 2005, number of years of practice, other specialisations, practice within 90 km from the University of Debrecen, communication with the patient).
For the evaluation of the significant predictors required for the correct diagnosis of clinical case 2, the a priori variables (age and gender) and the significant predictors identified by simple logistic regression were included in the initial model (GP specialization, number of years of practice, correct diagnosis, practice within 90 km from the University of Debrecen).
Significant predictors of accurate information for clinical case 1 (e.g., disclosure of urgency and relevant information about possible thrombolysis for question 2) were included in a multiple logistic regression model, so that a priori variables (age and gender) and significant predictors were presented in a simple logistic identified by regression and initially introduced (GP specialization, number of years of practice, correct diagnosis)
Significant predictors of providing accurate information for clinical case 2 (e.g., disclosure of urgency and relevant information regarding possible thrombolysis to question 2) were included in a multiple logistic regression model, so that a priori parameters (age and gender) and significant predictors were simplified. We identified them by logistic regression and initially introduced them (GP specialization, degree obtained before 2005, number of years of practice, practice within 90 km from Debrecen). The variables were entered into the model simultaneously and then variables that did not contribute significantly to

the model were deleted. To assess the adequacy of the fit, a $\chi 2$ test was performed on the final models.
Statistical analysis was performed using Stata 13.0 software (Stata Corporation - USA). Parameters showing a normal distribution were expressed as mean ± standard deviation (SD), and non-normally distributed parameters were expressed as medians with interquartile ranges (IQR), except for odds ratios (OR), which were presented with a 95% confidence interval (CI).

3.2. Evaluation of patient flow within the clinic

3.2.1. To investigate the efficiency and effectiveness of several models for organizing patient trips at a campus-structured university clinic

In our study, we analyzed the potential benefits of different organization of patient trips by investigating the patient care models of a campus-based university clinic, the Clinical Center of the University of Debrecen, modeled with mathematical algorithms.
Our study consists of a retrospective analysis of anonymized patient-level care data at the Clinical Center of the University of Debrecen generated between January 1 and December 31, 2013. Patient-level data included the inpatient ward(s) and outpatient ward(s) identification code for the care event and the primary diagnosis code associated with the care (according to ICD-10). The codes of the inpatient and outpatient units corresponded to the codes of the State Public Health and Medical Officer Service, which contain information on the medical specialty and subspecialties of the provided service pursuant to Decree 2/2004 (XI. 17.) of the Ministry of Health. This specialty / subspecialty code was further characterized by the ICD-10 code set, which also characterizes the scope of services of patient care units. Diagnosis codes were available for all cases, as they form the basis for funding benefits based on homogeneous disease groups.
In our study, we used the distances traveled/transferred between outpatient and inpatient units, and the patient pathways and distances were optimized.
The distances between the patient care units located on campus were calculated according to the GPS coordinates of the centers of the buildings accommodating the wards, the vertical distances were not taken into account. All outpatient movements were also considered to be patient flows from the inpatient unit to the outpatient unit and from the outpatient unit to the inpatient unit. Only benefits involving the physical movement of patients are included in the analysis, and non-appearance blood, tissue, etc. outpatient benefits related to sampling were excluded. As for the councils, the GPS coordinates of the

outpatient facility hosting the consulting physician were used for the calculations, and the distance to and from the inpatient unit was included in the analysis.

The current practice was used as a starting point for two algorithm-based optimizations.

The first algorithm aimed at rearranging and optimizing the location of inpatients, while the second algorithm optimized the distances involved in moving patients by relocating outpatient units. We investigated the minimization of the non-value-adding component of patient transport — physical distance — according to two algorithms. The optimization considered the professional decisions (the patient's diagnosis) to be given, e.g. the patient with the given diagnosis could only be placed in the inpatient department with the professional code corresponding to the diagnosis. The algorithms targeted the reorganization of the inpatient ward (e.g. patients were reassigned to new inpatient wards) and the outpatient wards were relocated to minimize the need to transfer patients (e.g. number of patient movements and physical distance) while maintaining clinical care. processes remained unchanged.

Following the optimization, we analyzed the characteristics of the initial graph and the two optimized graphs using graph theory. The graph was constructed so that the nodes were the physical building, and the edges were the patient traffic data.

3.2.2. Measurements of patient flow distribution

To compare the equality of patient flow generated by the original and the two optimized algorithms, the Gini coefficient (G) and the Lorenz asymmetry coefficient (S) were calculated as follows.

$$G = \frac{\sum_r \quad r k_r}{N \sum_r \quad k_r} - \frac{N+1}{N},$$

where N is the number of nodes (number of physical buildings) and k_r is the r-edik lowest degree of the nodes in ascending order of degrees.

The Gini coefficient expresses the inequality of the system, its possible values are between 0 and 1. The value of the coefficient is 0 if the whole system is "democratic," that is, there are no inequalities, if the value of the coefficient approaches 1, it means that there are few nodes with a large number of degrees in the system and relatively many smaller degrees.
The Lorenz asymmetry coefficient can be plotted using the Lorenz curve and indicates the degree of asymmetry in the curve. The Lorenz curve is used to describe the inequality of a quantity distribution. One of the summary statistics of the Lorenz curve is the Gini coefficient, which is a comprehensive measure of inequality within the study population. The Lorenz asymmetry coefficient is a useful addition to the Gini coefficient.

3.2.3. Algorithm used for optimization

Using the algorithm, we assumed that all outpatient visits and consultations were medically justified. Movements between inpatient units could not be circular, i.e., returning to the same location, only patient movements that occurred between two buildings were considered. The patient was transferred to an alternative inpatient ward if the sum of outpatient visits and medical consultation distances was less than original, and the care profile of the alternative ward was consistent with the patient's diagnosis (this match was indicated by the patient's BNO- Between the 10 codes and the BNO-10 codes of the units). Patients who underwent surgery were transferred to another inpatient unit for the sole purpose of surgery.
The outpatient units were relocated as follows: the algorithm allowed the unit to be relocated if the sum of the distances associated with patient flow was reduced. Two different algorithms were used, depending on the order of rearrangement of inpatients and the relocation of outpatient units.

3.2.4. Networking

To represent the graphs, we use the so-called neighbourhood matrix of the graph. Let G be a simple undirected graph of order n and denote the vertices $v_1, \ldots, v_n$. Then the matrix $A = (a_{ij})\ n \times n$ is determined as follows

$$a_{ij} := \begin{cases} 1 & ha\ \{v_i, v_j\} \in V \\ 0 & otherwise, \end{cases}$$

where V is the set containing the vertices. Then the matrix A is called the neighborhood matrix of the graph G.
If G is a simple, undirected weighted graph, then its neighborhood matrix $A = (a_{ij})$ can be defined as follows:

$$a_{ij} := \begin{cases} w_{ij} & ha\ \{v_i, v_j\} \in V \\ 0 & otherwise, \end{cases}$$

where w_{ij} denotes the weights.
The neighborhood matrix of undirected simple graphs is symmetric. Denote by $D = diag(d_w(v_i))$ the diagonal matrix containing the weighted degrees of the vertices, where:

$$d_w(v_i) = \sum_{\substack{j=1 \\ i \neq j}}^{n} w_{ij}$$

The normalized neighbourhood matrix for the graph G:

$$N := D^{-\frac{1}{2}} \cdot A \cdot D^{-\frac{1}{2}}$$

and the normalized Laplacian matrix for the graph G can be defined as follows:

$$L := I - N = D^{-\frac{1}{2}} \cdot (D - A) \cdot D^{-\frac{1}{2}}$$

where I is the unit matrix and $L_G := D - A$ is the non-normalized Laplacian matrix of the graph G.

The formula used to calculate the weighted edge density is:

$$\frac{2}{N(N-1)} \sum_{i=1}^{N} \sum_{j=i+1}^{N} w_{ij}$$

According to the mathematical formulas defined above, we created a weighted undirected neighborhood matrix (adjacency matrix) in which the nodes in the graph are the buildings. There is a connection between two nodes if there is a patient / consultation between the two buildings. The weight of the edge belonging to the two nodes (w) is the annual number of patients / doctors commuting between the two nodes (building). We created three separate networks that reflect the two optimization schemes, thus creating an original, inpatient and outpatient network. In addition, 100 randomized and 100 lattice null grids were generated so that the number of nodes, total edges, and degree distributions remained unchanged. We used the Rubinov and Sporns algorithm to generate random graphs by identifying reversible edge pairs in the original network.

3.2.5. Network statistics

Basic network statistics were calculated for all three networks. The number of nodes (N), the degree (n) of the node, and the average node strength (S) were given:

$$s_i = \sum_{\substack{j=1 \\ i \neq j}}^{N} w_{ij}$$

$$S = \frac{1}{N} \sum_{i=1}^{N} s_i$$

Lowercase uppercase letters indicate the node, uppercase letters indicate the global parameter, lowercase i and lowercase j denote individual nodes, where s_i is the weight of the i-th node and S is the average node strength.
The length of the weighted characteristic path was calculated by first normalizing the neighborhood matrix and then using the Floyd-Warshall algorithm. Defining the shortest path:

$$L = \frac{1}{N}\sum_{i=1}^{N} L_i \text{ and } L_i \equiv L(i) = \frac{1}{N-1}\sum_{\substack{j=1 \\ j\neq i}}^{N} \frac{1}{F_{i,j}}$$

where L and F denote the path length and the maximum flow function. The normalized diameter (D), defined as the longest shortest path, was also calculated:

$$D \coloneqq \max_{i,j} d_{ij}$$

where d_{ij} is the shortest path between nodes *I* and *j*.

The so-called spacing centrality of the graph was determined by the number of shortest paths passing through the given node. We also calculated the clustering coefficient.

$$C_i \equiv C(i) = \frac{3\sum_{j,h=1\ j,h=i}^{N} w_{ij}w_{jh}w_{hi}}{\sum_{j,h=1\ j,h=i}^{N} w_{ih}w_{hj} + \sum_{j,h=1\ j,h=i}^{N} w_{ij}w_{jh}w_{hi}}$$

and $C = \frac{1}{N}\sum_{i=1}^{N} C_i$

To compare the optimization procedures, the clustering coefficient was normalized to the average values of the equivalent random networks, while normalization with the average values of the lattice network was used to calculate the small world index (ω). The small-world properties were evaluated as previously described:

$$\omega = \frac{L^{rnd}}{L} - \frac{C}{C^{lat}} \equiv \frac{1}{nL^{rnd}} - nC^{lat}$$

If the value of ω between -1 and 1 is close to 0, it indicates small properties of the world, with positive and negative values showing a greater tendency towards a random and lattice network, respectively.
The small world was further characterized by the global and local efficiency of the network.

3.2.6. Smallworld properties

The problem of the small world has been studied since the 1960s, when Milgram and colleagues first coined the concept of how many personal connections can be made from one person to another.
In a traditional network, the average length of the shortest path between arbitrary node pairs increases rapidly as the system size increases. In contrast, for systems with small-world properties, there is a logarithmic relationship between the average shortest path length and the number of nodes. In other words, you can get from any node to any other node along only a series of edges that touch a few intermediate nodes, so the distances are small relative to the system size.
Further research has shown that people-to-people social networks are characterized by how many connections are needed to get from one person to another, which seems to be an important pre-age issue in today's fast-paced world using the Internet and social media, but graphs can still be used today. and also in the world of matrices.
Among the matrices and graphs, the average distance of the vertices seems to be small compared to the number of vertices, in the case of random graphs the size of the graphs decreases greatly in the case of randomly recorded edges.

3.2.7 The Laplacian matrix

3.2.7.1. Spectral analysis

Understanding the complexity of patient flow using graphs is being investigated by both clinicians and mathematicians.
Models based on parallel interactions are considered by some to be incapable of capturing the relationships between nodes in the network. Beyond these, higher-order network models offer new perspectives for understanding complex systems. Using a unified mathematical language makes it possible to select the optimal model from different modeling approaches.
By transforming the normalized weighted neighborhood matrix, we determined the normalized Laplacian matrix associated with the graph for further analysis of the dynamic interactions between the global network structure and the

network parts. In addition, the feasibility of system-level optimization was characterized by spectral analysis of the weighted neighborhood and Laplacian matrices.

$$L \coloneqq I - N = D^{-\frac{1}{2}} \cdot (D - A) \cdot D^{-\frac{1}{2}}$$

In the dissertation, the spectrum of the graph is defined as the spectrum of the normalized Laplacian matrix, where the normalized Laplacian matrix of the graph G can be defined as follows:

The Laplacian matrix describes aspects of the network related to the global network structure and also provides information on the dynamic interactions between parts of the network. The advantage of using a normalized Laplacian matrix is that all its eigenvalues fall in the [0.2] range, which helps to compare networks of different sizes.
The eigenvalues of the weighted neighborhood matrices (λi $i = 1 \ldots n$) were analyzed by calculating the spectral radius, the spectral gap, and the eigenvector centrality. Spectral diagrams were constructed from the smoothed eigenvalue distribution as previously described. The spectra were compared by quantitative spectral distance measurement, the average Euclidean distance between the two spectral diagrams as described by others.
We calculated the algebraic connectivity, the community structure of the network, the two-sidedness and the inverse Eigenratio.

4. RESULTS

4.1. Investigating the knowledge of GPs about acute stroke - putting the guidelines into practice, organizing patient referrals

4.1.1. The respondents

Of the 127 respondents in the study, 69 persons (54.3%) were women, and the median age of the sample was 49 (34–62) years. The median number of years spent in employment after graduation was 14.5 (2-22.5) years; 98 (77.2%) physicians had a general practitioner board exam, and the remaining 29 respondents (22.8%) enrolled in the general practitioner training program. Of the GPs, 58 also had other medical specialisations (e.g. internal medicine, neurology, anesthesiology etc.), while only 1 resident doctors had a previous internal medicine exam. Regarding the location of the practice in terms of distance from Debrecen, the practice of 74 respondents (51 GPs and 23 residents) was located within 90 km of the city and 44 respondents (40 GPs and 4 residents) were located outside this area (relevant geographical data 9 were missing in some cases).

4.1.2. Qualitative results

According to a qualitative analysis of the responses to the open-ended questions, both case reports referred to strokes. Several respondents indicated spatial localization of stroke (left middle cerebral artery) and possible etiology (ischemic, emboliform), which is also reflected in the frequency of words in both clinical cases. Interestingly, the word “acute” was only among the 10 most common words in clinical case 1, which describes a patient eligible for thrombolysis, while the diagnosis of clinical case 2 was distinguished by the frequent mention of the term “out of time”. Interestingly, TIA is the 4^{th} and 10^{th} most used word for clinical case 1 and 2, respectively. In addition, it is noteworthy that cerebral infarction was often indicated as a diagnosis in clinical case 2 as well.

Regarding physician-patient communication, respondents clearly distinguished the two clinical scenarios in terms of suitability for thrombolysis. The need for hospital treatment (the most used word in both cases) was articulated in both cases, and the neurology department was also frequently mentioned in this context (8 and 6 most common in clinical cases 1 and 2, respectively). In addition, ‘thrombolysis’ (2nd most common word), ‘within time window’ urgency (6th most common word), ‘immediately’ (7th most common word), ‘should’ (9th most common word) and the words “ambulance” (10 most common words)) were described by respondents in relation to clinical case 1, while the ineligibility for thrombolysis ("out of time window" (3 most common words), "no thrombolysis" (5 most common words)) were clearly stated / described by the respondents in clinical case 2. It should also be noted that 'necessary' (9 most common words), 'referral' (10 most common words), 'examination' (4 most common words), 'CT scan' (7 most common words)) were frequently mentioned in connection with communication with the patient in clinical case 2. However, the word "stroke" was not always mentioned directly, but rather circumscribed (cerebral infarction, cerebrovascular accident, impaired blood supply to the brain, blockage in the brain, cerebral artery clot).

4.1.3. Quantitative results

Simple logistic regression (which identifies factors that significantly influence the chance of answering question 1 of clinical case 1 correctly) revealed that age was a significant predictor (OR 1.05; 1.02-1.09; p = 0.004). After examining the effect of all significant predictors and a priori predictors on the outcome variable using a multiple logistic regression model, the chances of giving an inappropriate response were significantly increased if the practice fell outside the 90 km radius of the Debrecen stroke center (OR 4.06; 1.35-12.24; p = 0.013) and that the physician omitted relevant information about the possible use of thrombolysis when communicating with the patient (OR 4.83; 1.55-15.07; p = 0.007). The final model proved to be significant due to the statistical adequacy of the $\chi2$ test (p = 0.48).
Simple logistic regression (used to identify significant determinants of the correct answer to question 1 in clinical case 2) showed that age was a significant predictor (OR 1.05; 1.02-1.09; p = 0.002). Other significant variables related to clinical case 1: degree obtained before 2005 (OR 0.32; 0.11-0.92; p = 0.034), number of years of practice (OR 1.06; 1.02-1.10) ; p = 0.003), other specialisations (0.29; 0.12-0.71; p = 0.007), thrombolysis mentioned to the patient (question 2 of clinical case 1) (OR 3.98; 1.70-9 , 29; p = 0.001), accurate communication of thrombolysis (Clinical Case 1, Question 2) (OR 2.83; 1.17-6.89; p = 0.022). Other significant variables related to clinical case 2 were: age (OR 1.06; 1.02-1.09, p = 0.002), work (qualified GP) (OR 0.09; 0.01-0.65 ; p = 0.018), degree obtained before 2005 (EH 0.10; 0.02-0.47; p = 0.003), number of years of practice (OR 1.07; 1.03-1.12; p = 0.000), other medical qualification (OR 0.39; 0.16-0.94; p = 0.036), the practice is within a radius of 90 km (OR 3.63; 1.45-9.07; p = 0.006).
Significant predictors for correct diagnosis in clinical case 1 for the final multiple logistic regression model were the location of the practice within 90

km (OR 4.06; 1.35-12.24; p = 0.013) and mentioning the thrombolysis for question 1 (OR 4.83; 1.55-15.08; p = 0.007).

The final multiple logistic regression model showed that the chance of giving an inappropriate response was three times higher if the practice fell outside the 90 km radius (OR 3.03; 1.06–8.61; p = 0.038). The final model was significant (p = 0.22).
In the case of clinical case 1, both age (OR 1.05; 1.02-1.08; p = 0.000) and gender (OR 3.28; 1.54-7.00 (p = 0.002) was significant.
Interestingly, the final model included only these a priori pre-identified variables (age: OR 1.05; 1.02-1.08; p = 0.01 and no: OR 2.73; 1.23-6.06; p = 0.013). The set-up model was significant (p = 0.41).
Through simple logistic regression of clinical case 2 (description of an acute stroke patient outside the time window), significant predictors of accurate physician-patient communication were age (OR 1.08; 1.04-1.13; p = 0.000) , work (general practitioner with other medical specialist exam) (yes / no) (OR 0.19; 0.07-0.54; p = 0.002), diploma obtained before 2005 (yes, no) (OR 0.22; 0, 79-0.62; p = 0.004), the number of practitioners (OR 1.09; 1.02-1.15; p = 0.006) and the practice is within a 90 km radius (yes / no) (OR 3 , 8; 1.03-13.97; p = 0.045). The final multivariate logistic regression model found only one factor, age to be significant (OR 1.15; 1.02–1.29; p = 0.019). The final model was significant (p = 0.094).

4.2. Results obtained by network analysis

At the University of Debrecen, we investigated the efficiency of various patient pathway organization models and algorithms. In the final network the patient flow between 24 clinical buildings was first analysed, and after optimization, the Neurology, Internal Medicine, and Pulmonology Clinics were investigated with statistical analyses.

4.2.1. Patient flow data

When evaluating patient flow data, we found that the distance resulting from moving patients (including patients transferred to outpatient wards for examination and physicians commuting for consultation) was 83,184 km. This value refers to 100,301 episodes when patients had to leave the building that housed their inpatient ward or doctors came from a unit located in a building other than the patient’s building.

Following optimization to minimize total patient-related distance, the sum of patient-related distances was reduced to 57,096 km and 55,070 km for outpatient and outpatient optimization, respectively. Thus, the optimization led to a 31.4% and 33.8% reduction in patient / physician mileage following inpatient and outpatient-to-outpatient optimization. In contrast, the number of episodes associated with exiting the building hosting their inpatient units also decreased to 77,022 (23.2% decrease) and 76,376 (23.8% decrease) episodes for inpatient and outpatient optimization. As a result of inpatient- and outpatient-inpatient optimization, out of the 411 outpatient wards, 97 (23.6%) and 117 (28.5%) outpatient wards were relocated.
Equality indicators (Gini coefficient, Lorenz curve) showed no significant change, indicating node inequality in patient flow. Nevertheless, the Gini coefficient increased after both optimization schemes. The Lorenz curve was virtually symmetric in all three cases.

4.2.2. Summary network statistics

The original network has 24 nodes (e.g. outpatient and inpatient buildings) and 266 connections (edges). After optimization, the number of edges connecting the 24 nodes was reduced to 232 and 237, respectively, in the inpatient-outpatient and outpatient-inpatient networks.
The optimization led to a decrease in the average node degree, as this value was 22.17 in the original network and 19.33 in the inpatient and 19.75 in the outpatient network. Regarding the average strength of nodes reflecting patient flow, we found that optimization reduced the number of patients commuting between buildings in both newly created scenarios. Accordingly, the weighted edge density decreased, and the duality centrality increased following both optimization procedures.

4.2.3. Smallworld properties

Both optimization schemes appear to result in networks that are associated with lower levels of small world properties than the original network. Among the parameters influencing the small world index, the normalized weighted characteristic path length decreased (1.30 vs 0.98 for the original and inpatient networks), while the normalized clustering coefficient for the lattice network increased (0.65 vs 0.89 for the original and inpatient networks). The inpatient-outpatient network Omega (smallworld index) showed a marginal increase (0.11 vs. 0.13 for the original and inpatient-outpatient networks), making the optimized network somewhat more random than the original network. The discrepancy is further supported by the fact that both local and global efficiencies are lower than the corresponding parameters in the original network.

The relocation of outpatient units and then the rearrangement of inpatients resulted in a network that was significantly more regular than the original network, indicated by a significantly lower omega value (small world index) (-0.61). Nevertheless, this is probably due to a disproportionate increase in the normalized characteristic pathway, which reaches 5.26, possibly due to some buildings assuming very high normalized path lengths (rheumatology, gastroenterology, paediatric oncology, oncology and rehabilitation - 7.44; 3.08; 3.08; 2.28; 2.03). Large route distances between nodes indicate that not enough information is coming from one node to the rest of the network. In contrast, both global efficiency and local efficiency were lower in this network, indicating that parallel information flow is less efficient, and nodes have more limited ability to integrate themselves into the network.

4.2.4. Hub status

The global indicator of clustering and hub formation, Eigenvector-centricity, has increased. Using any values of 3 or 2 for the clustering threshold (normalized to random and lattice networks), we found that while in the original network only the Neurology building showed high clustering, reflecting high specialization, additional new, highly specialized buildings after optimization appeared in addition to neurology. Thus, following inpatient-outpatient optimization, we identified the Internal Medicine and Pulmonology buildings as highly specialized buildings. After outpatient inpatient optimization, only one other building, Internal Medicine Building 1, appeared as a specialized building. The assessment of the “hub” status of the buildings yielded similar results using any 0.85 threshold. In the original network, only the Neurology building received hub status and new hubs emerged after optimization (as a result of inpatient-outpatient optimization, the hubs became the Neurology, Internal Medicine and Pulmonology buildings, while the outpatient-inpatient optimization resulted in the formation of “hubs” within the Neurology, Internal Medicine buildings). In addition, after both optimization schemes, the value of the numerical parameters (Normalized clustering coefficient (random), Normalized clustering coefficient (grid), Eigenvector centricity reflecting the importance of the nodes increased for the network as a whole.

4.2.5. Spectral analysis

Spectral analysis using the eigenvalues of the neighbourhood matrix showed that the optimization led to a less coherent network (spectral radii of 2.19 and 2.21 and 2.29, respectively, for outpatient-inpatient, inpatient-outpatient and original network) Synchronization, on the other hand, improved following inpatient-outpatient optimization procedures, which is reflected in an increased spectral gap between 1.49 and 1.20 in the original network, while the rearrangement of outpatients did not affect this parameter (λ = 1, 20). The Eigenvector centrality was higher in both optimized networks than in the original network, indicating that these new networks contain a larger number of nodes.

The metric of the algebraic relationship, the normalized Laplacian matrix spectrum, increased after inpatient-outpatient optimization, while it decreased in the other optimization scheme. The normalized Laplacian spectra of the Laplacian eigenvalues of the original and the two optimized networks were visually compared as described by de Lange et al. We found that the three spectra were similar in that their community structure was weak, showed a high peak around 1, and had a skewed asymmetric distribution. The weak community structure shared by all three of our networks is reflected in the fact that there is only one non-zero eigenvalue in the informative range of the community structure of the networks (λ <0.5). Further evaluation of the community structure by evaluating the largest eigenvalue difference (difference between eigenvalues) showed that none of the three networks could be divided into communities (the largest difference was $\lambda1 - \lambda2 = 0.41$, $\lambda1 - \lambda2 = 0.48$ and $\lambda1 - \lambda2 = 0.39$ was the original network, inpatient-outpatient, and outpatient-inpatient networks). This also means that although there is a degree of modularity, the community structure is weak.

The average Euclidean distance between the spectra was very small. The outpatient-inpatient network is approximately equidistant from the other two networks, while the distance between the original and inpatient-outpatient networks is slightly greater.

It is likely that this distance reflects differences in the density of eigenvalues around one, whereas the Laplacian spectrum of the outpatient network shows the peak intermediate of the other two networks.

The maximum eigenvalue of the Laplacian spectrum decreased significantly after each optimization, reflecting the fact that the newly developed networks show reduced bipartidity ($\lambda24 = 1.65$; $\lambda24 = 1.58$; $\lambda24 = 1.45$, respectively, of the original, inpatient-outpatient and outpatient-inpatient networks). Based on further observations, in the case of the original and inpatient-outpatient networks, the highest eigenvalue was localized to the neurology building, while in the case of the outpatient-inpatient network to the urology building. In addition, visual inspection revealed a more similar network topology for the original and inpatient networks in terms of assigning patient flow to buildings (reflected in the eigenvector of the maximum eigenvalues of the buildings, similar to the node strength distribution of the networks). After further evaluation of the maximum eigenvalues of the Laplacian spectra, symmetric peaks around $\lambda = 1$ were found in each case, reflecting the high number of eigenvalues in all three networks between $0.9 \leq \lambda \leq 1.1$. This effect was most pronounced in the inpatient-outpatient rearrangement scheme and was least pronounced in the original network.

In addition, based on the evaluation of inverse Eigenratio stability, network synchronization was increased (0.255, 0.31, and 0.27 for the original inpatient and outpatient networks).

5. DISCUSSION

Our results suggest that the practical transposition of the clinical guidelines for the treatment of acute stroke (thrombolysis) into general practice was successful in the regions of Hungary we studied. This conclusion was drawn based on how the general practitioners evaluated the two acute stroke clinical cases presented. We found a clear geographical effect that reflects the regional, formalized acute stroke triage system, drawing attention to the beneficial effects of a systematic approach to translational efforts.

The vast majority of respondents clearly defined the two clinical cases of acute stroke as strokes and distinguished them in terms of their suitability for thrombolysis (suitability for thrombolysis or lack of suitability for thrombolysis). It also appears that over the years, acute stroke has become known as an emergency condition among respondents. Our results also implicitly indicated that GPs / GP residents consider atrial fibrillation as a potential risk factor for stroke because they often indicated an embolic origin.

Our results show better practice in the application of acute stroke guidelines than described in previous studies. In a similar Brazilian survey that assessed the stroke knowledge of 149 emergency and rescue professionals, the authors found that although the vast majority of respondents were able to properly diagnose the case presented as a stroke, the time window was mentioned only once, the possibility of thrombolysis four times and the need for CT scan twelve times. In contrast, the results of a recent study in China also showed that GPs do not have sufficient knowledge about prehospital stroke care, a finding supported by low rates of intravenous thrombolysis. The need for pre-hospital recognition of stroke to reduce delays in pre-hospital care has been articulated by other authors. The significance of these findings is very significant, as both countries, China and Brazil, also organize acute stroke care in accordance with international guidelines. A recent prospective study in Hungary analysed data from 250 stroke patients unfit for thrombolysis in a hospital without a dedicated stroke unit. Their results showed that 37.2% of these patients first visited the GP on duty (47 patients) or the GP surgery (46 patients), and only 91 (36.4%) patients called an ambulance immediately. In addition, the authors found a 2.66-fold increase in contact with ambulances directly if the patient's previous atrial fibrillation was known, confirming the suggestion that patients treated with stroke risk factors are more aware of the best outcome for acute stroke. Doctor-patient encounters can provide a good foundation for patient education.

In our research, the quantitative evaluation of the data showed that the location of the practice within 90 km from Debrecen increases the chances of making a proper diagnosis in both clinical cases presented. The significance of this geographical distinction is derived from the organized acute stroke triage system operating in the region, which is maintained by the Department of Neurology of the University of Debrecen. In addition to a well-defined referral regimen supported by a formal agreement between the relevant actors in regional acute stroke care (ambulances, regional hospitals, GPs), the systematic follow-up of thrombolysis candidates, clinical decisions and thrombolysis outcome was complemented by further training of GPs in parallel with the ongoing training of paramedics, emergency doctors and the hospital stroke team. The success of the program is reflected in the high rate of intravenous thrombolysis, which accounts for 16–19% of acute stroke cases and the death rate from lower strokes in Hajdú-Bihar County 7.07 per 100,000 population, compared to the national average 8.2 / 100,000 inhabitants in 2018 and 7.07 / 100,000 inhabitants in Debrecen, compared to the national average: 8.2 / 100,000 inhabitants.

Age was a significant factor in both clinical cases presented when evaluating the content of urgency and thrombolysis-relevant information in both clinical cases presented. If the age of GPs increased by one year, in clinical cases 1 and 2 the chances of incorrect answers increased 1.04-fold and 1.15-fold, respectively, with an increase in age per unit. Older GPs are less positive about applying the guidelines, which may contribute to a lack of implementation of policy changes.

In addition, our results suggest that in clinical case 2 (where the patient was not a candidate for thrombolysis according to the case report), male physicians were more likely to respond incorrectly regarding communication with the patient. This finding is inconsistent with the results of a previous study examining adherence to practical guidelines involving primary care physicians. The study had to analyse video cases showing coronary heart disease. The results showed no major or interaction effect on the gender of the responding physicians.

Previous reports have also highlighted the significant impact of understanding prognosis and the limitations of recommendations for the treatment of patients with acute intracerebral stroke. In the sample of 742 neurosurgeons and neurologists, the odds of treatment limit increased by 1.61 (95% CI 1.12; 2.33) when the 0% prognosis for functional independence was measured on a prognostic scale. In contrast, the likelihood of treatment limitations was lower when the prognostic score estimated a high chance of independence. In a previous study, findings were found to support an association between knowledge about stroke and expected outcome. Here, deeper knowledge increased the chances of giving more importance to the rapid identification of stroke symptoms (OR: 1.23; 1.002-1.51).

In interpreting the results, the circumstances of the data collected in the present study should be borne in mind, according to which respondents were approached at an in-service training course or a seminar held for GP resident doctors, so participants were presumably more motivated to learn. The topic of these trainings was not acute stroke care, so we presumably evaluated the existing knowledge of the participants in our survey. In addition, the use of open-ended questions to assess clinical cases allowed for the assessment of clinical competencies, as this presupposes a higher level of thinking and knowledge building. The sample size was relatively small; however, robust results were obtained.

Summarizing our results, it can be said that the geographical location of the GP practice and the age of the GPs showed a correlation with their knowledge of the organization and treatment of acute stroke.

The algorithms used in the clinical patient pathway organization to improve the initial model have been shown to improve care efficiency and reduce the costs associated with patient movements.

The training of GPs has proved important, and they seem to have put their theoretical knowledge into practice. However, further research would be needed to improve knowledge about the organization of patient trips and acute stroke care, possibly using GIS (geographic information systems) knowledge.

Our main findings showed that changes in guidelines for the treatment of acute stroke are being successfully applied by GPs in their general practice; therefore, GPs can be good targets for raising awareness about acute stroke issues, including timely referral of patients to an appropriate specialist setting.

In the other part of our study, we used individual patient-level data to evaluate the organization of patient flow in a campus-based university hospital to examine the possibility of reorganizing patient transport, one of the necessary but not physically added moments of patient care, by minimizing physical journeys. This concept can be derived from a "lean" approach to process organization where the lack of physical proximity between inpatient wards and outpatient units is not considered to add value. In view of this, an algorithm-based optimization scheme was used to minimize physical distance, with the patient being transferred to another inpatient class, followed by reassignment of outpatient units, or vice versa. Supporting the work of others, this concept has allowed us to design patient transport in a way that simplifies non-value-added steps and thus limits their potential impact while keeping sales outcomes for patient care unchanged (e.g., the need for consultations and outpatient visits). Using graph theory, we performed a network analysis to characterize the service delivery system before and after optimization. While we used the concepts of graph theory to address management issues such as scheduling or logistics, to the best of our knowledge, this is the first implementation of graph theory to evaluate the performance of a health care provider.

Feasible alternatives to the current practice have been identified using an algorithm that reduces patient / physician distances by more than 30% and avoids patient movements between forms of care in 23,279 and 23,925 cases, inpatient-outpatient and outpatient- inpatient patient flow after optimization activities. This was achieved through a system-wide, top-down approach, reallocating outpatient units, and reassigning patients to departments that receive patients with similar diagnoses. After optimization, we compared the old and the new supply system by network analysis. We found that the rearrangement of inpatients and the relocation of outpatient units led to a decrease in the number of patient pathways (reflected in mean nodes, mean nodule strength, mean weighted edge density). The decrease in weighted edge density also points to some form of specialization in buildings, as fewer buildings are connected by fewer roads, meaning that patients receive more services in their building, so these buildings offer more patient-specific care by offering a variable number of outpatient services, in line with the needs of inpatients. These findings are further reinforced by the increase in clustering (indicated by an increase in the clustering coefficient) and the appearance of nodes (reflected by an increase in the centrality of the eigenvector). In addition, the emergence of major nodes was indicated by an increase in the centrality of the duality. This means that the number of high-impact nodes is increasing, so more buildings are generating high traffic. More specifically, topological segregation (or local grouping) in this information processing of a university-centred hospital refers to special care provided in separate buildings. In our original and optimized networks, we were able to identify node regions that are in a position to make a strong contribution to global network communication by assuming hub status. In contrast, the connectivity of networks also decreases, which is reflected in the spectral radius and algebraic relationship of the indices (only in the case of outpatient-inpatient optimization). On the other hand, synchronization was also increased by neighbourhood and Laplacian matrix analysis (see spectral gap and normalized inverse Eigenratio).

In evaluating small world properties, we found that optimization led to a decrease in smallworldness and a decrease in local and global efficiency, thereby reducing fault tolerance and parallel information flow in these networks. Although these parameters appear to be unfavourable from a graph theory point of view, it is suggested that changing these indicators could be beneficial for the health care environment, as it reflects a reduction in one form of specialization and parallel processes. When organizing health services, the best way to organize care is to organize in a physical location with sufficient expertise and infrastructure. This observation also suggests that different parameters should be interpreted for systems and their networks, as previously suggested by others.

However, when examining the Laplacian spectra of the three networks, we found that they were quite similar and that the characteristic parameters did not change (e.g., the maximum of the difference in normalized eigenvalues) or changed favourably (inverse Eigenratio).

Examination of assortability showed that all three networks are disortative, optimized networks are significantly more disorientative than the original networks. This finding reinforces our initial idea that patient flow is organized along medical needs that reflect biological processes and is not based solely on social relationships.

A number of patient flow models have been set up to analyse health networks and examine patient flow. These were most often modelled using different algorithms and efficacy was analysed based on patient movement patterns.

Reducing patient delivery requires significant resources from the patient, the health care provider, and the care system, both in terms of the number of nursing deliveries and the distances travelled. In addition to the inconveniences that have already arisen, patient safety issues need to be addressed. Previous research has shown that relocating patients from one location to another increases the risk of adverse events resulting from inadequate communication. Based on this, it can be argued that proper management of patient flow logistics contributes to increasing the value of health care, which is often defined by a framework for improving performance. It may also reflect outcomes based on the outcome of the treatment of the disease, which may contribute to the outcome of patient care (in terms of quality of care, patient satisfaction and safety) and the direct and indirect costs of care.

In summary, in our current study we examined the possibility of reorganizing patient care in a campus-based hospital through simple, top-down algorithm-based optimization so that care is reduced by the distance physically travelled by patients and their physicians without medical perspectives. The optimization resulted in a system that reduced the need for multi-building patient transfers, increased building specialization, and reduced parallel processes. These effects were identified using concepts related to graph theory. Based on our results, it can be suggested that network analysis can be a useful tool in planning patient flow capacity.

6. NEW RESULTS

1. Qualitative text analysis and the use of word clouds adequately reflected GPs' knowledge and practice of acute stroke patient care. Based on our results, we suggest that qualitative text analysis is a suitable method to analyse the translation of new knowledge into clinical practice.
2. Geographical location was a significant factor regarding the successful translation of knowledge, as if the GP practice was within 90 km of the stroke centre regularly offering trainings about acute stroke care, the assessment of suitability for thrombolysis and the organization of the appropriate patient pathway were significantly higher.
3. The application of algorithms in the organization of care is an effective tool in the planning of the structure of optimal outpatient and inpatient specialist care and was able to decrease a non-value adding step in patient care (e.g. the distance travelled by patients/doctors). Minimizing number of patient transport episodes and their physical length is preferred as it increases patient safety and capacity utilization.
4. Keeping every medical decision unchanged the algorithm-based optimization yielded an altered physical structure for patient care. Using graph theory and spectral graph theory the patient flow was analysed with conclusions emerging regarding the network underlying the patient care (both pre-optimization and post-optimization).

7. SUMMARY

The processing of healthcare data has been an important area of health research for centuries and its methodology is constantly changing and evolving. The application of new approaches is important and forward-looking for the development of health care. In my dissertation, I presented two studies with a new approach.

In the first study, we investigated the transfer of clinical guidelines for acute stroke into everyday clinical practice among GPs, in which we performed quantitative and qualitative text analysis using NVivo software and a world cloud. Our qualitative analysis revealed stroke several times as a diagnosis correctly established by GPs. Examining the differences between areas and the possible aetiology, the respondents adequately identified their suitability for thrombolysis. Quantitative evaluation has shown that the closer the GP district is to the stroke centre, the greater the likelihood of a proper diagnosis of acute stroke.

Our results show that the majority of GPs properly diagnose acute stroke and identify patients eligible for intravenous thrombolysis, but there may be differences between the practices of individual physicians. The geographical location of the practice may influence the system of acute stroke triage and the referral order of patients, which also draws attention to the beneficial effects of a systematic approach to translational efforts and ongoing training.

In our other study, we examined more efficient patient pathway reorganization models for patients who have already been hospitalized to meet the requirements of patient-centred care. One of our goals was to reorganize patient pathways by optimizing the relocation of inpatient and outpatient units. Our analysis was performed in a campus-based hospital (University of Debrecen, Clinical Center). Patient flow data were used for algorithm-based optimization to minimize the sum of the distance between outpatient visits and visits for physician consultation. In our optimized model, we reorganized inpatient care and relocated outpatient units to minimize transport demand. The optimized schemes were analysed using graph and spectral graph theory.

Our results show that the optimization process changed the hospital layout and reduced the need for patient transfers. Systems developed with algorithm-based optimization methods can reduce the need for patient transport between buildings, and network analysis has proven to be an effective tool for planning patient flow capacity within a hospital.

In summary, newer approaches may be suitable for analysing health systems and contributing to their development.

8. ACKNOWLEDGMENTS

I would like to express my gratitude to Prof. Dr. Judit Zsuga, my supervisor, for the participation in the planning, implementation and evaluation of the scientific research underpinning the Ph.D. dissertation and for allowing me to publish it as a first and co-author. Thank you for your continued help in preparing this dissertation, for your support and encouragement. I would like to thank the members of our research team for their cooperation in helping me with my scientific work.

I am grateful for the support of my husband, my family, who, by providing a calm, peaceful background, created the conditions for me to carry out my work and research and for enduring the hardships and resignations that come with it.

9. APPENDIX

UNIVERSITY of DEBRECEN

UNIVERSITY AND NATIONAL LIBRARY
UNIVERSITY OF DEBRECEN
H-4002 Egyetem tér 1, Debrecen
Phone: +3652/410-443, email: publikaciok@lib.unideb.hu

Registry number: DEENK/131/2022.PL
Subject: PhD Publication List

Candidate: Szilvia Kolozsváriné Harsányi
Doctoral School: Doctoral School of Health Sciences

List of publications related to the dissertation

1. Papp, C., **Kolozsváriné Harsányi, S.**, Gesztelyi, R., Emri, M., Zsuga, J.: Assessment of patient flow and optimized use of lean thinking transformation from the perspective of graph theory and spectral graph theory: a case study.
   Technol. Health Care. 29 (2), 199-211, 2021.
   DOI: http://dx.doi.org/10.3233/THC-191782
   IF: 1.285 (2020)

2. **Kolozsváriné Harsányi, S.**, Balogh, N., Kolozsvári, L. R., Mézes, L., Papp, C., Zsuga, J.: Acute stroke awareness of family physicians: translation of policy to practice.
   Health Res Policy Sys. 18 (1), 1-9, 2020.
   DOI: http://dx.doi.org/10.1186/s12961-020-00642-5
   IF: 3.318

List of other publications

3. Nánási, A., Ungvári, T., Kolozsvári, L. R., **Kolozsváriné Harsányi, S.**, Jancsó, Z., Lánczi, L., Mester, L., Móczár, C., Semánová, C., Schmidt, P., Szidor, J., Torzsa, P., Rurik, I.: Expectations, values, preferences and experiences of Hungarian primary care population when accessing services.
   Prim. Health Care Res. Dev. 22, 1-7, 2021.
   IF: 1.458 (2020)

4. Garbóczy, S., Szemán-Nagy, A., Ahmad, M. S., **Kolozsváriné Harsányi, S.**, Ocsenás, D., Rekenyi, V., Al-Tammemi, A. B., Kolozsvári, L. R.: Health anxiety, perceived stress, and coping styles in the shadow of the COVID-19.
   BMC Psychol. 9 (1), 1-13, 2021.
   DOI: http://dx.doi.org/10.1186/s40359-021-00560-3

UNIVERSITY AND NATIONAL LIBRARY
UNIVERSITY OF DEBRECEN
H-4002 Egyetem tér 1, Debrecen
Phone: +3652/410-443, email: publikaciok@lib.unideb.hu

5. Kolozsvári, L. R., Bérczes, T., Hajdu, A., Gesztelyi, R., Tiba, A., Varga, I., Al-Tammemi, A. B., Szőllősi, G. J., **Kolozsváriné Harsányi, S.**, Garbóczy, S., Zsuga, J.: Predicting the epidemic curve of the coronavirus (SARS-CoV-2) disease (COVID-19) using artificial intelligence: an application on the first and second waves.
Informatics in Medicine Unlocked. 25, 1-13, 2021.
DOI: http://dx.doi.org/10.1016/j.imu.2021.100691

6. Garbóczy, S., Szemán-Nagy, A., Ahmad, M. S., **Kolozsváriné Harsányi, S.**, Ocsenás, D., Rekenyi, V., Tischler, P., Al-Tammemi, A. B., Kolozsvári, L. R.: The Emergence of Dissociative Experiences as a Function of Perceived Stress Among University Students During the COVID-19 Lockdown.
Research Square. 2021, 1-18, 2021.
DOI: http://dx.doi.org/10.21203/rs.3.rs-145432/v1

7. Tele-Héri, B., Dobos, K., **Kolozsváriné Harsányi, S.**, Pálinkás, J., Fenyősi, F., Gesztelyi, R., Móré, E. C., Zsuga, J.: Vestibular Stimulation May Drive Multisensory Processing: principles for Targeted Sensorimotor Therapy (TSMT).
Brain Sci. 11 (8), 1-14, 2021.
IF: 3.394 (2020)

8. Kolozsvári, L. R., Eörsi, D., Busa, C., Csikós, Á., Hargittay, C., Kalabay, L., **Kolozsváriné Harsányi, S.**, Mohos, A., Radványi, I., Rurik, I., Szabó, J., Török, K., Varga, A., Várnai, R., Gutási, É., Korolovszky, J.: A háziorvosképzés helyzete és fejlesztési lehetőségei. Nemzeti Népegészségügyi Központ, Budapest, 132 p., 2020.

9. Busa, C., Csikós, Á., Eörsi, D., Hargittay, C., Kolozsvári, L. R., **Kolozsváriné Harsányi, S.**, Mohos, A., Radványi, I., Szabó, J., Várnai, R.: Elemzés az alapellátásban dolgozó ápolók képzésének, kompetenciáinak nemzetközi gyakorlatáról. Akademiai Kiado Rt, Budapest, 100 p., 2020.

10. Barczi, E., Breitenbach, Z., Busa, C., Czető, Á., Csikós, Á., Eörsi, D., Galvács, H., Gyetvai, G., Hanka, K., Hargittay, C., Hegedűs, O., Járomi, M., Kalmár, Z. J., Karamánné Pakai, A., Kárpáti, Z., Kerti, E., Kiss, L., Kiss, N., Kolozsvári, L. R., **Kolozsváriné Harsányi, S.**, Korolovszky, J., Markó-Kucsera, M., Máté, O., Mohos, A., Oláh, A., Pállay-Kovács, S., Szabó, E., Szabóné Tamás, H., Tománé Mészáros, A., Ujváriné Siket, A., Várnai, R., Zrínyi, M.: Fejlesztési terv az egészségügyi szakdolgozók alapellátási és népegészségügyi ismereteinek, kompetenciáinak bővítésére. Akademiai Kiado Rt, Budapest, 86 p., 2020.

11. Barczi, E., Breitenbach, Z., Busa, C., Czető, Á., Csikós, Á., Eörsi, D., Galvács, H., Gyetvai, G., Hanka, K., Hargittay, C., Hegedűs, O., Járomi, M., Kalmár, Z. J., Karamánné Pakai, A., Kárpáti, Z., Kerti, E., Kiss, L., Kiss, N., Kolozsvári, L. R., **Kolozsváriné Harsányi, S.**, Korolovszky, J., Markó-Kucsera, M., Máté, O., Mohos, A., Oláh, A., Pállay-Kovács, S., Pátri, L., Szabó, E., Szabóné Tamás, H., Tománé Mészáros, A., Várnai, R., Zrínyi, M.: Koncepció az alapellátási kompetenciabővítő képzésekre. Akademiai Kiado Rt, Budapest, 76 p., 2020.

UNIVERSITY AND NATIONAL LIBRARY
UNIVERSITY OF DEBRECEN
H-4002 Egyetem tér 1, Debrecen
Phone: +3652/410-443, email: publikaciok@lib.unideb.hu

12. Kolozsvári, L. R., Bérczes, T., Hajdu, A., Gesztelyi, R., Tiba, A., Varga, I., Szőllősi, G. J., **Kolozsváriné Harsányi, S.**, Garbóczy, S., Zsuga, J.: Predicting the epidemic curve of the coronavirus (SARS-CoV-2) disease (COVID-19) using artificial intelligence.
medRxiv 2020, 1-16, 2020.
DOI: http://dx.doi.org/10.1101/2020.04.17.20069666

13. Égerházi, A., Kovács, Z. G., Magócs, É., Szőlősi, G. J., **Kolozsváriné Harsányi, S.**, Garbóczy, S., Kolozsvári, L. R.: Teszteld a memóriád (Test Your Memory) kérdőív magyar változatának (TYM - HUN) alkalmazhatósága.
Magy. gerontol. 11 (37-38), 25-34, 2020.

14. Garbóczy, S., Magócs, É., Szőllősi, G. J., **Kolozsváriné Harsányi, S.**, Égerházi, A., Kolozsvári, L. R.: The use of the Hungarian Test Your Memory (TYM-HUN), MMSE, and ADAS-Cog tests for patients with mild cognitive impairment (MCI) in a Hungarian population: a cross-sectional study.
BMC Psychiatry. 20 (1), 571-577, 2020.
DOI: http://dx.doi.org/10.1186/s12888-020-02982-6
IF: 3.63

15. Móré, E. C., Papp, C., **Kolozsváriné Harsányi, S.**, Gesztelyi, R., Mikáczó, A., Tajti, G., Kardos, L., Seres, I., Lőrincz, H., Csapó, K., Zsuga, J.: Altered irisin/BDNF axis parallels excessive daytime sleepiness in obstructive sleep apnea patients.
Respir. Res. 20 (1), 1-15, 2019.
DOI: http://dx.doi.org/10.1186/s12931-019-1033-y
IF: 3.924

16. Kolozsvári, L. R., Kónya, J., Paget, J., Schellevis, F. G., Sándor, J., Szőllősi, G. J., **Kolozsváriné Harsányi, S.**, Jancsó, Z., Rurik, I.: Patient-related factors, antibiotic prescribing and antimicrobial resistance of the commensal Staphylococcus aureus and Streptococcus pneumoniae in a healthy population - Hungarian results of the APRES study.
BMC Infect Dis. 19, 1-8, 2019.
DOI: https://doi.org/10.1186/s12879-019-3889-3
IF: 2.688

17. Zsuga, J., Móré, E. C., Erdei, T. D., Papp, C., **Kolozsváriné Harsányi, S.**, Gesztelyi, R.: Blind spot for sedentarism: redefining the diseasome of physical inactivity in view of circadian system and the irisin/BDNF axis.
Front. Neurol. 9, 1-13, 2018.
DOI: http://dx.doi.org/10.3389/fneur.2018.00818
IF: 2.635

UNIVERSITY of DEBRECEN

UNIVERSITY AND NATIONAL LIBRARY
UNIVERSITY OF DEBRECEN
H-4002 Egyetem tér 1, Debrecen
Phone: +3652/410-443, email: publikaciok@lib.unideb.hu

18. Kolozsvári, L. R., Kovács, Z. G., Szőllősi, G. J., **Kolozsváriné Harsányi, S.**, Frecska, E., Égerházi, A.: Validation of the Hungarian version of the Test Your Memory = a Teszteld a memóriádat (Test Your Memory) magyar változatának validálása.
Ideggyogy. Szle. 70 (7-8), 267-272, 2017.
DOI: http://dx.doi.org/10.18071/isz.70.0267
IF: 0.252

Total IF of journals (all publications): 22,584
Total IF of journals (publications related to the dissertation): 4,603

The Candidate's publication data submitted to the iDEa Tudóstér have been validated by DEENK on the basis of the Journal Citation Report (Impact Factor) database.

01 April, 2022