

### AI-driven Optimization in Healthcare: the Diagnostic Process

Lyon, Jérôme Yves; Bogodistov, Yevgen; Moormann, Jürgen

Veröffentlichungsversion / Published Version

Zeitschriftenartikel / journal article

**Empfohlene Zitierung / Suggested Citation:**

Lyon, J. Y., Bogodistov, Y., & Moormann, J. (2021). AI-driven Optimization in Healthcare: the Diagnostic Process. *European Journal of Management Issues*, 29(4), 218-231. <https://doi.org/10.15421/192121>

**Nutzungsbedingungen:**

Dieser Text wird unter einer CC BY Lizenz (Namensnennung) zur Verfügung gestellt. Nähere Auskünfte zu den CC-Lizenzen finden Sie hier: <https://creativecommons.org/licenses/by/4.0/deed.de>

**Terms of use:**

This document is made available under a CC BY Licence (Attribution). For more information see: <https://creativecommons.org/licenses/by/4.0>

UDC classification: 004.08:614.23

JEL Classification: I10, O33, C63

## AI-driven Optimization in Healthcare: the Diagnostic Process

J. Y. Lyon<sup>‡</sup>,  
Y. Bogodistov<sup>#</sup>,  
J. Moormann<sup>##</sup>

**Purpose:** Process optimization in healthcare using artificial intelligence (AI) is still in its infancy. In this study, we address the research question “To what extent can an AI-driven chatbot help to optimize the diagnostic process?”

**Design / Method / Approach:** First, we developed a mathematical model for the utility (i.e., total satisfaction received from consuming a good or service) resulting from the diagnostic process in primary healthcare. We calculated this model using MS Excel. Second, after identifying the main pain points for optimization (e.g., waiting time in the queue), we ran a small experiment ( $n = 25$ ) in which we looked at time to diagnosis, average waiting time, and their standard deviations. In addition, we used a questionnaire to examine patient perceptions of the interaction with an AI-driven chatbot.

**Findings:** Our results show that scheduling is the main factor causing issues in a physician’s work. An AI-driven chatbot may help to optimize waiting time as well as provide data for faster and more accurate diagnosis. We found that patients trust AI-driven solutions primarily when a real (not virtual) physician is also involved in the diagnostic process.

**Practical Implications:** AI-driven chatbots may indeed help to optimize diagnostic processes. Nevertheless, physicians need to remain involved in the process in order to establish patient trust in the diagnosis.

**Originality / Value:** We analyze the utility to physicians and patients of a diagnostic process and show that, while scheduling may reduce the overall process utility, AI-based solutions may increase the overall process utility.

**Research Limitations / Future Research:** First, our simulation includes a number of assumptions with regard to the distribution of mean times for encounter and treatment. Second, the data we used for our model were obtained from different papers, and thus from different healthcare systems. Third, our experimental study has a very small sample size and only one test-physician.

<sup>‡</sup> Jérôme Yves Lyon,  
Junior IT-Consultant,  
GOB Software & Systeme GmbH & Co. KG, Germany

<sup>#</sup> Yevgen Bogodistov,  
Senior Lecturer in Process Management,  
MCI Management Center Innsbruck, Austria,  
e-mail: yevgen.bogodistov@mci.edu,  
<https://orcid.org/0000-0002-5904-8362>

<sup>##</sup> Jürgen Moormann,  
Professor of Bank and Process Management,  
Frankfurt School of Finance & Management, Germany,  
e-mail: j.moormann@fs.de,  
<https://orcid.org/0000-0002-8682-2747>

**Paper type:** Empirical

**Keywords:** Artificial Intelligence, Process Optimization, Chatbot, Healthcare, Diagnosing.

**Reference** to this paper should be made as follows:

Lyon, J. Y., Bogodistov, Y., Moormann J. (2021). AI-driven Optimization in Healthcare: The Diagnostic Process. *European Journal of Management Issues*, 29(4), 218-231. doi:10.15421/192121.

## Штучний інтелект та оптимізація в галузі охорони здоров'я: процес встановлення діагнозу

Жером Ів Ліон<sup>†</sup>  
Євген Богодістов<sup>#</sup>  
Юрген Морманн<sup>##</sup>

<sup>†</sup> GOB Software & Systeme GmbH & Co. KG, Німеччина  
<sup>#</sup> MCI Центр менеджменту м. Інсбрук, Австрія  
<sup>##</sup> Франкфуртська школа фінансів та менеджменту, Німеччина

**Мета роботи:** Оптимізація процесів у охороні здоров'я з використанням штучного інтелекту все ще перебуває у початковому стані. У цьому дослідженні ми вирішуємо дослідницький питання: «Якою мірою чат-бот, керований штучним інтелектом, може допомогти оптимізувати діагностичний процес?»

**Дизайн / Метод / Підхід дослідження:** По-перше, ми розробили математичну модель для корисності (тобто загального задоволення, що отримується від споживання товару чи послуги) внаслідок діагностичного процесу у первинній охороні здоров'я. Ми розрахували цю модель за допомогою MS Excel. По-друге, визначивши основні болючі точки для оптимізації (наприклад, час очікування в черзі), ми провели невеликий експеримент ( $n = 25$ ), в якому вивчили час до встановлення діагнозу, середній час очікування та їх стандартні відхилення. Крім того, ми використовували анкету для вивчення сприйняття пацієнтами взаємодії з чат-ботом, керованим штучним інтелектом.

**Результати дослідження:** Наші результати показують, що складання розкладу є основним фактором, який викликає проблеми у роботі лікаря. Керований штучним інтелектом чат-бот може допомогти оптимізувати час очікування, а також надати дані для більш швидкої та точної діагностики. Ми виявили, що пацієнти довіряють рішенням на основі штучного інтелекту насамперед тоді, коли у діагностичному процесі бере участь реальний (не віртуальний) лікар.

**Практична цінність дослідження:** керовані штучним інтелектом чат-боти дійсно можуть допомогти оптимізувати діагностичні процеси. Тим не менш, лікарі повинні залишатися залученими до процесу, щоб пацієнти довіряли діагнозу.

**Оригінальність / Цінність дослідження:** Ми проаналізували корисність діагностичного процесу для лікарів та пацієнтів і показали, що хоча складання розкладу може знизити загальну корисність процесу, рішення на основі штучного інтелекту можуть підвищити загальну корисність процесу.

**Обмеження дослідження / Майбутні дослідження:** По-перше, наше моделювання включає низку припущень щодо розподілу середнього часу зустрічі та лікування. По-друге, дані, які ми використовували для нашої моделі, були отримані з різних документів, а отже, із різних систем охорони здоров'я. По-третє, наше експериментальне дослідження має дуже маленький розмір вибірки і лише одного лікаря-випробувача.

**Тип статті:** Емпіричний

**Ключові слова:** штучний інтелект, оптимізація процесів, чат-бот, охорона здоров'я, діагностика.

## Искусственный интеллект и оптимизация в области здравоохранения: процесс постановки диагноза

Жером Ив Лион<sup>†</sup>  
Евгений Богодистов<sup>#</sup>  
Юрген Морманн<sup>##</sup>

<sup>†</sup> GOB Software & Systeme GmbH & Co. KG, Германия  
<sup>#</sup> MCI Центр менеджмента г. Инсбрук, Австрия  
<sup>##</sup> Франкфуртская школа финансов и менеджмента, Германия

**Цель работы:** Оптимизация процессов в здравоохранении с использованием искусственного интеллекта (ИИ) все еще находится в зачаточном состоянии. В данном исследовании мы решаем исследовательский вопрос: «В какой степени чат-бот, управляемый ИИ, может помочь оптимизировать диагностический процесс?»

**Дизайн / Метод / Подход исследования:** Во-первых, мы разработали математическую модель для полезности (т.е. общего удовлетворения, получаемого от потребления товара или услуги) в результате диагностического процесса в первичном здравоохранении. Мы рассчитали эту модель с помощью MS Excel. Во-вторых, определив основные болевые точки для оптимизации (например, время ожидания в очереди), мы провели небольшой эксперимент ( $n = 25$ ), в котором изучили время до постановки диагноза, среднее время ожидания и их стандартные отклонения. Кроме того, мы использовали анкету для изучения восприятия пациентами взаимодействия с чат-ботом, управляемым искусственным интеллектом.

**Результаты исследования:** Наши результаты показывают, что составление расписания является основным фактором, вызывающим проблемы в работе врача. Управляемый ИИ чат-бот может помочь оптимизировать время ожидания, а также предоставит данные для более быстрой и точной диагностики. Мы обнаружили, что пациенты доверяют решениям на основе ИИ в первую очередь тогда, когда в диагностическом процессе участвует реальный (не виртуальный) врач.

**Практическая ценность исследования:** Управляемые ИИ чат-боты действительно могут помочь оптимизировать диагностические процессы. Тем не менее, врачи должны оставаться вовлеченными в процесс, чтобы пациенты доверяли диагнозу.

**Оригинальность / Ценность исследования:** Мы проанализировали полезность диагностического процесса для врачей и пациентов и показали, что, хотя составление расписания может снизить общую полезность процесса, решения на основе ИИ могут повысить общую полезность процесса.

**Ограничения исследования / Будущие исследования:** Во-первых, наше моделирование включает ряд допущений в отношении распределения среднего времени встречи и лечения. Во-вторых, данные, которые мы использовали для нашей модели, были получены из разных документов, а значит, из разных систем здравоохранения. В-третьих, наше экспериментальное исследование имеет очень маленький размер выборки и только одного врача-испытателя.

**Тип статьи:** Эмпирический

**Ключевые слова:** искусственный интеллект, оптимизация процессов, чат-бот, здравоохранение, диагностика.

## 1. Introduction

The world of healthcare has changed enormously in recent years, and a patient-centric approach is increasingly important for modern healthcare business practice. In 2018, on average 71% of the world's population had visited a primary care physician at least once a year, while 28% had consulted a physician three times or more a year (Advisor, 2018). Patients surveyed stated “access to treatment and long waiting times” as the top issue in their healthcare system, followed by issues of “not enough staff,” “too high costs of accessing treatment,” and “bureaucracy” (Advisor, 2018: 44).

The world's population is getting older, and there is a lack of medical capacity to cope with the resulting demand for treatment. Besides capacity problems, physicians struggle with new technology and have to cope with changing and increasing regulation (Fuchs, 1996; Saltman & Figueras, 1997; Haimi, Brammli-Greenberg, Waisman, & Baron-Epel, 2018; Carayon & Hoonakker, 2019). In addition, patients can choose from a variety of physicians and hospitals, making the healthcare market even more competitive (Ettinger, 1998; Varkevissar, van der Geest, & Schut, 2012). Consequently, medical professionals are under constant pressure to offer cost- and time-efficient treatment while at the same time satisfying the individual needs and expectations of their patients. Technology is often regarded as an approach that can improve cost-effectiveness and scheduling (Cutler, 2007; Rau et al., 2013).

Recent developments in the field of healthcare require each physician to have not only up-to-date professional knowledge but also the capability to process vast amounts of information (Moreira, Rodrigues, Korotaev, Al-Muhtadi, & Kumar, 2019). Digitalization allows central storage of patient-related data as well as opportunities for collecting additional data (e.g., using smartwatches or smartphone-connected pill bottles) and applying advanced data analysis strategies (Bhavnani, Narula, & Sengupta, 2016). At the same time, the burden of learning, predicting, and diagnosing grows accordingly. This growth requires more sophisticated AI technologies such as Machine Learning and Deep Learning to allow physicians to extract useful information from data (Bohr & Memarzadeh, 2020).

One specific new technology has attracted great attention and is expected to revolutionize the healthcare sector in the future: artificial intelligence (AI). Using machine learning, computers can learn from experience, recognize causal connections in the recorded data, execute tasks based on these learnings, and further improve their knowledge. Thus, implementation of AI in healthcare information systems is expected to assist or even partly replace medical professionals in the future. This study contributes to the literature by proposing a new approach to using AI to reduce the workload of medical professionals as well as costs for patients while ensuring proper care and patient satisfaction. A holistic view of primary diagnosis in ambulatory care is used to examine the effects of an AI-based decision support tool that is incorporated into a standard primary care process.

## 2. Research Question

In this study, we address the research question: To what extent can an AI-driven chatbot help to optimize the diagnostic process?

## 3. Theoretical Background

The healthcare market is evolving continuously and becoming more complex as a result (Plsek & Greenhalgh, 2001). Despite the improvements already made, researchers are developing strategies, concepts, and tools to advance the healthcare system further. Suggestions in the literature are focused on four areas. The first area is concerned with policy-related topics, including how to improve policymaking and regulate or deregulate

the healthcare sector (Fuchs, 1996; Marmor & Wendt, 2012). Some researchers estimate the quality and performance of healthcare, while others try to determine the utility derived from treatment. Researchers also use the process utility derived from screening procedures to operationalize measures in preventive care using a range of measurement methods such as standard gamble techniques, time trade-off techniques, and conjoint analysis (Brennan & Dixon, 2013).

Nevertheless, physicians and patients may perceive the quality of the process differently, and this implies different process utilities. Indeed, the perception of service quality by a physician deviates to some extent from the perception by the patient (Levine et al., 2012), a fact that should be taken into consideration when service quality is evaluated. Results and findings vary as much as the approaches taken. Some researchers have analyzed best practice in diagnosing patients and minimizing medical errors. The literature in this stream suggests that healthcare is far from being accurate and that error rates are unacceptably high (Herzlinger, 2006; Graber, 2013).

Another stream of literature analyzes queuing techniques and utilization planning. A variety of modeling methods and heuristic models have been used to determine which effects occur if appointment-making is altered (Ahmadi-Javid, Jalali, & Klassen, 2017). Papers in this area mostly deal with the uncertainty of different determinants of everyday healthcare practices, and the uncertainty that might be related to the perception of process quality.

Current efforts to use advanced technology focus mainly on the application of telehealth systems to specific medical conditions, for example, telehealth monitoring devices for managing congestive heart failure patients (Lehmann, Mintz, & Giacini, 2006), and telehealth approaches to chronic obstructive pulmonary disease (Polisena et al., 2010) and diabetes management (Polisena et al., 2009). Other approaches analyze the impact of digital health assistants that are not directly connected to human healthcare professionals. These assistants include apps that remind a patient to take his or her medicine (Dayer, Heldenbrand, Anderson, Gubbins, & Martin, 2013) or enable a patient to perform a self-diagnosis (Semigran, Linder, Gidengil, & Mehrotra, 2015). Other systems give general advice on how to improve a patient's general health by, for example, losing weight (Kamel Boulos, Brewer, Karimkhani, Buller, & Dellavalle, 2014). However, according to the evaluation of a symptom-checker compared with a real practitioner (Semigran et al., 2015), self-diagnosing tools have lower accuracy rates than real-life physicians and are currently not accurate enough to represent a viable alternative to physician visits.

Only a few studies focus on a combined approach, i.e., evaluating a shared solution where the patient's use of technology at the front end is managed by a healthcare professional at the back end. There are also some decision support systems that rely on AI to propose a diagnosis or make recommendations (Krittawong, Zhang, Wang, Aydar, & Kitai, 2017; Miller & Brown, 2018). However, most of these tools are focused on specific conditions or symptoms and, more importantly, are designed for specialists only. The question therefore remains: Can these approaches help to optimize processes in healthcare facilities? In this paper, we address this question from the viewpoint of process management.

We root our study in the notion of system welfare, proposed by Allon & Kremer (2018), who suggested looking at the welfare of a system from the perspective of an individual. For hospital management, welfare is present if the system runs as planned in terms of service value  $v$  (taken as interchangeable with service quality), cost (i.e., disutility) of waiting in a queue  $c_w$ , cost due to time spent in the patient encounter  $c_s$ , waiting time in the queue  $T_w$ , and the time required for each patient to be processed by a physician  $T_s$ . The processing time  $T_s$  can be described as a relation between  $\mu_j$  (time spent per unit of work  $j$ , e.g., time to measure body temperature) and  $w_j$  (units of work  $j$ , i.e., the activity of measuring body temperature).  $R$  represents the system

throughput, i.e., the number of patients treated during a time period. Following Allon & Kremer (2018), we can therefore formalize system welfare in the following manner:

$$SW = (v - c_w T_w - c_s T_s) \cdot R = (v - c_w T_w - c_s \sum_j \frac{w_j}{\mu_j}) \cdot R$$

In this study, we focus on two types of service value, as  $v$  represents both the utility of the encounter for a patient and the utility of the encounter for a physician. Perceptions of value may differ; whereas patients may enjoy longer conversation time with the physician, perceiving it as a sign of respect and necessary attention, physicians may perceive the time spent as a missed opportunity to encounter more patients and, thus, as low efficiency and effectiveness of their work. The welfare of the system is, therefore, a balance between the utilities of both the patient and the physician.

However, as Allon & Kremer (2018) argued, it is not only the waiting time that has an impact on the perception of utility, but also the work context. AI-driven technologies may reduce waiting time thanks to fast and precise information processing. However, if the system is not trusted, the utility and, consequently, the welfare of the system may decrease. In our investigation, we shed light not only on process optimization through the use of chatbots but also on patient perceptions of AI-driven technology. We also consider the content variables (i.e., the encounter-specific diagnostic activities) necessary for the welfare of a health system.

## 4. Methodology

To analyze the factors that can improve patient and doctor satisfaction, we develop a model that calculates the overall utility of a patient–doctor interaction in primary care. To calculate the utility, we identify a list of determinants of utility for the patient received during a standard encounter. We also include the utility for the physician in order to simulate the process of patient–physician interaction.

### 4.1. Determinants of Utility

The primary goal of doctor–patient interaction is to increase the patient’s welfare. The fact that the satisfaction of patients contributes to the positive outcome of treatments is widely acknowledged (Hall, Ferreira, Maher, Latimer, & Ferreira, 2010; Hudak, Hogg-Johnson, Bombardier, McKeever, & Wright, 2004; Rubel, Bar-Kalifa, Atzil-Slonim, Schmidt, & Lutz, 2018). It follows that the actual medical treatment is not the only determinant of a patient’s perceived utility of the healthcare system.

#### 4.1.1. Patient’s Perceived Utility

Additional determinants of patient satisfaction include the attitude of medical personnel, prompt service, the ability to share information with patients, the patience of the doctor in doing so, and the availability and use of the latest equipment (Carlucci, Renna, & Schiuma, 2013; Hassin & Haviv, 2003; Levine et al., 2012; Peprah, 2013; Teke et al., 2010). In contrast, long waiting times, unfriendliness, and incorrect diagnoses reduce a patient’s welfare. The utility function of the patient can be set up on this basis. Let  $U_t$  be the utility of the treatment outcome,  $U_a$  the utility from the attitude of the personnel,  $U_i$  the utility from the physician’s ability to share information,  $U_p$  the utility derived from the patience of the doctor,  $U_e$  the utility derived from new equipment, and  $U_w$  the utility resulting from waiting time. Thus:

$$U_{patients} = U_t + U_a + U_i + U_p + U_e + U_w \quad (1)$$

Some factors cannot be influenced by the implementation of a technical solution such as an AI-driven chatbot. The specificity of treatments, personalities of medical professionals, and money spent on the interior of a hospital are beyond the scope of technological improvements. This fact leads to a restriction on our utility function (1), where  $U_a$ ,  $U_i$ , and  $U_e$  become irrelevant for the further analysis of an AI-based system and are set to zero. This leaves us with a restricted utility equation for the patient:

$$U_{patients} = U_t + U_p + U_w \quad (2)$$

To calculate the utility from waiting time ( $U_w$ ), the waiting time ( $T_w$ ) should be multiplied by a utility factor for every time unit spent on waiting. Note that the benefit from low waiting time ( $b_{waiting}$ ) and the cost of waiting ( $c_{waiting}$ ) because of long waiting time have been separated. This separation is necessary to account for the endowment effect by which people overvalue the loss of goods they already possess in comparison to the goods they gain. In our case, this is the loss of time that patients could have spent on other activities in comparison to having completed the physician visit earlier than expected. Consequently, it can be assumed that  $|b| < |c|$ .

$$U_w = b_w \times |T_w|, \quad \text{if } T_w \leq 0 \quad (3a)$$

$$U_w = c_w \times T_w, \quad \text{if } T_w > 0 \quad (3b)$$

If medical treatment is effective, the patient will receive a positive utility ( $b_{treatment}$ , or  $b_t$ ). However, if the treatment is not effective or if it worsens the condition of the patient, s/he will incur an even higher negative utility ( $c_{treatment}$ , or  $c_t$ ). The utility gain of an effective treatment is calculated based on the benefit of an effective cure ( $b_t$ ) multiplied by the expected time needed for the cure minus the actual time needed for the cure ( $T_c^{expected} - T_c^{actual}$ ). If a patient is cured earlier than expected, s/he will receive a higher benefit from such a process. If a patient is cured in time, i.e., exactly as expected, s/he might still derive a benefit, which means that the multiplier will always be 1 or higher. If a treatment is not effective or the physician has diagnosed the patient incorrectly, the cost of the wrong treatment ( $c_t$ ) will be multiplied by the actual time to the cure and also by a wrong-treatment impact factor ( $i_w$ ). The longer a patient receives an incorrect treatment, the higher the utility loss. In addition, some consequences of illness are worse than others; a false diagnosis of a severe illness will cause a higher utility loss than, for example, the common flu. Therefore, we introduce a wrong-treatment impact factor as follows:

$$U_t = b_t \times \max(1, T_c^{expected} - T_c^{actual}), \quad \text{if treatment was effective} \quad (4a),$$

$$U_t = c_t \times T_c^{actual} \times i_w, \quad \text{if treatment was ineffective} \quad (4b)$$

The utility derived from the patience of the physician ( $pU_p$ ) is determined by the individual perception of the patient of the benefit or cost derived from the physician’s patience ( $b_{patience}/c_{patience}$ ) but also by the time that is actually “freed” for the physician, as freed time reflects reduced workload ( $p+T_b$ ). If the actual processing time of the patient ( $p^*$ ) is less than the planned time slot ( $p+T_b$ ) for the diagnosis, the physician has enough time for the patient and is not in a hurry. S/he can engage in a more personal conversation, which results in an extended and better explanation of the illness and of the proposed treatment, as well as helping to build a better relationship by answering the patient’s questions. Consequently, the patient will derive a benefit (Coulter & Jenkinson, 2005; Thompson, Yarnold, Williams, & Adams, 1996). If the physician is under time pressure because the diagnosis took the planned time or longer ( $(p+T_b) \leq p^*$ ), every additional minute will reduce the patience of the physician, and the patient will derive a negative utility:

$$U_p = b_p \times (p + T_b - p^*), \quad \text{if } (p + T_b) > p^* \quad (5a)$$

$$U_p = c_p \times (p + T_b - p^*), \quad \text{if } (p + T_b) \leq p^* \quad (5b)$$

Given these determinants, we can set up an equation for the patient’s utility. Parts of this equation can be changed against the benefit or cost equations if applicable.

$$U_{patient} = b_t \times \max(1, T_c^{expected} - T_c^{actual}) + b_p \times (p + T_b - p^*) + b_w \times |T_w| \quad (6a),$$

$$U_{patient} = c_t \times T_c^{actual} \times i_w + c_p \times (p + T_b - p^*) + c_w \times T_w \quad (6b)$$

#### 4.1.2. Physician's Utility

A physician also derives utility from interacting with a patient. Clearly, the physician's income depends on the diagnostic process; the more efficiently the doctor can diagnose patients, the higher his/her income. It is therefore in the interest of the healthcare professional to diagnose the patient correctly and offer the best treatment for the diagnosed illness. To define the utility function of the physician, let  $U_{tr}$  be the benefit from the treatment,  $U_{pr}$  the utility from processing the patient in time, and  $U_d$  the utility derived from the accompanying administrative work.

$$U_{doctor} = U_{tr} + U_{pr} + U_d \quad (7)$$

For simplicity, we assume that a physician derives utility from each treatment regardless of whether s/he diagnosed the patient correctly<sup>1</sup>. Consequently, the physician derives a benefit from the treatment of each patient ( $b_{tr}$ ).

$$U_{tr} = b_{tr} \quad (8)$$

An encounter that takes longer than expected might cause delays in the schedule and create additional stress for the medical personnel. The benefit from processing time depends, therefore, on the deviation from the planned diagnosis time slot ( $p+T_b$ ). If diagnosing the patient takes less time than expected ( $p+T_b > p^*$ ), the physician will receive an additional positive utility from this specific treatment. If the actual processing time ( $p^*$ ) is exactly as planned, the physician will still be satisfied and will derive a benefit, which means that the multiplier is always 1 or higher. However, if the processing time is longer than expected ( $p+T_b < p^*$ ), the physician will derive a negative utility, as s/he will be working overtime and will be faced with more time pressure and stress. This, in turn, has a negative impact on his/her accuracy (Williams, Manwell, Konrad, & Linzer, 2007). Given these considerations, costs are calculated as costs per minute of overtime.

$$U_{pr} = b_{pr} \times \max(1, p + T_b - p^*), \text{ if } (p + T_b) \geq p^* \quad (9a)$$

$$U_{pr} = c_{pr} \times (p + T_b - p^*), \text{ if } (p + T_b) < p^* \quad (9b)$$

Another determinant of the physician's utility is the administrative work related to the patient-physician interaction. The burden of administrative work is high and time-consuming. As administrative work should not be the core activity of a physician, for every minute spent on administrative tasks, the physician will derive a negative utility.

$$U_d = T_d \times c_d \quad (10)$$

Given these determinants, the overall utility function of a physician can be set up as follows and is again interchangeable with the respective positive or negative utility parts:

$$U_{doctor} = b_{tr} + b_{pr} \times \max(1, (p + T_b - p^*)) + T_d \times c_d \quad (11a)$$

$$U_{doctor} = b_{tr} + c_{pr} \times (p + T_b - p^*) + T_d \times c_d \quad (11b)$$

Combining the utility equations for the patient and the physician results in the overall system utility:

$$U_{overall} = b_t \times \max(1, T_c^{expected} - T_c^{actual}) + b_p \times (p + T_b - p^*) + b_w \times |T_w| + b_{tr} + b_{pr} \times \max(1, (p + T_b - p^*)) + T_d \times c_d$$

<sup>1</sup> We assume per capita financing, as this is the case in Ukraine, where part of our data comes from. Each family doctor is assumed to supervise approximately 1,800 patients. Ukrainian family doctors are not paid according to the number of patients treated or treatments assigned, but according to a collective agreement signed with the hospital. Some

#### 4.2. Study 1: Standard Simulation Setup

In our first study, we developed a model that incorporates the utility equations from the previous section. The model was built in MS Excel and based on a computer-generated data set that calculates the arrival time, processing time, and administrative work time needed per patient. For each patient, we also took into account whether his/her treatment is effective, how long his/her cure takes, and for how long s/he expects to follow the prescribed treatment. Using the utility equation and the generated data, we calculated the physician and patient utility for each patient during one work week. The model operates on the basis of a five-day work week and nine work hours per day. Per day, it is assumed that a fixed number of 27 patients will be treated. Finally, we calculated the sum of the utility for all patients within one work week. This procedure was executed 100 times, and the trimmed average (with a cutoff of 10%) of the summed utility of all 100 runs was derived.

##### 4.2.1. Arrival Time

The calculation of the patient arrival time was based on the findings of Alexopoulos and colleagues (2008), who tested different distributions of their data using the ExpertFit automatic fitting procedure. They found that the best fit method for modeling the unpunctuality of patients was a Johnson  $S_U$  distribution with the following parameters estimated using quantile matching:  $\gamma = -0.576$ ,  $\delta = 1.548$ ,  $\lambda = 21.741$ , and  $\xi = -0.775$ , where  $\gamma$  and  $\delta$  are the shape parameters of their model,  $\lambda$  is a scale parameter, and  $\xi$  is a location parameter (Alexopoulos et al., 2008). Using this distribution, our model estimated the deviation from the planned appointment time (in other words, the actual arrival time).

##### 4.2.2. Processing Time

The processing time for each patient was calculated in a similar way. We used data collected within a research project that analyzed the impact of process changes on healthcare providers in two cities in Ukraine and had 179 participating family doctors (Bogodistov, Moormann, & Sibbel, 2018). The measured processing times of those physicians were analyzed using EasyFit (MathWave Technologies, 2019). The fitted distributions were tested using Kolmogorow-Smirnow, Anderson-Darling, and Chi-square tests of goodness of fit (GoF). The results of the distributions that might be relevant for analysis of the diagnostic process are given in Tab. 1. Based on their performance in the GoF tests, the fitted distributions were ranked from poorest fit (i.e., approximating 1) to best fit (i.e., approximating 0). The tests found that a normal distribution on average performed best in all three tests and consequently had the best fit to the dataset.

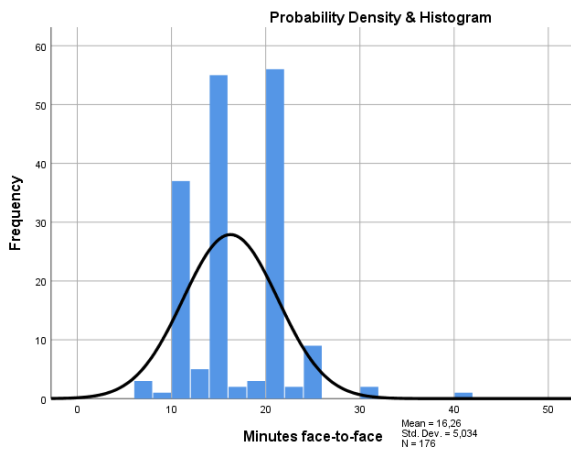
The calculated normal distribution had the determinants of  $\sigma = 5.0336$  and  $\mu = 16.261$  (Fig. 1). Skewness (.689) and kurtosis (1.919) were in the acceptable range (George & Mallery, 2019; Hair, Black, Babin, & Anderson, 2010). Based on an average processing time (which was varied during the simulation) and a standard deviation from the dataset above, we calculated the actual processing time for each patient as a normal distribution of the average processing time. As the processing time cannot be zero, the calculated processing time is always  $\geq 1$ .

hospitals develop sophisticated collective agreements that take account of the types of patients and the number of treatments or patient visits. However, in most cases, the quality of treatment is less relevant. Medical errors may have legal consequences and/or may lead to the hospital management deciding to reduce the agreed payment.

**Table 1:** Test statistics for processing time distribution generated with EasyFit

#	Distribution	Kolmogorow–Smirnow		Anderson–Darling		Chi-square	
		Statistic	Rank	Statistic	Rank	Statistic	Rank
1	Exponential	.43661	55	41.324	54	365.31	51
2	Logistic	.1914	24	7.2334	22	186.46	18
3	Lognormal	.19794	28	7.7726	28	197.65	39
4	Normal	.1728	1	6.5043	2	45.371	2
5	Uniform	.20604	36	50.802	56		N/A

Source: developed by the authors



**Figure 1:** Distribution of patient encounter times

Source: developed by the authors

**4.2.3. Starting Time**

The calculation of the starting time for each treatment recognizes that patients can be late (i.e., they arrive after the appointment time of the next patient) or that their diagnosis can take longer than expected, which will cause delays in subsequent appointments. As patients cannot be rejected because of lateness or because of a complicated condition, rescheduling has to be executed. The model checks whether each patient and the previous patients are too late or whether the next scheduled patient (who might have arrived early) can be treated first. If an encounter takes longer than expected, the next encounter will be delayed.

If a patient arrives early, the waiting time is calculated only from the time of the planned appointment. We assume that waiting caused by being too early will not affect the patient’s utility, as s/he knows the time of appointment and does not expect any additional utility from being early. Nevertheless, if the patient is encountered earlier than the planned appointment time (e.g., because the previous patient arrived too late), s/he will derive a positive utility from the reduction in waiting time. If a patient is late, the waiting time is calculated from the point of arrival.

**4.2.4. Duration of Administrative Work**

The time needed for administrative work related to each patient was calculated as 37% of the total processing time, based on the findings of Sinsky and colleagues (2016) that physicians spend on average 37% of the time they spend in the examination room on administrative tasks.

**4.2.5. Healing Time**

To calculate the time needed for healing, this simulation assumes a standard distribution based on a mean of seven days with a standard deviation of four days<sup>2</sup>. As the time needed for curing cannot be zero, the number will always be  $\geq 1$  and is rounded up to a full day. The same calculation was used to generate the patient’s expectation of the time needed for curing. This setting assumes that the curing time expected by the patient is realistic, although in some cases this time will deviate because of uncertainty associated with the course of the disease. Therefore, the curing time can be equal to, greater than, or less than expected. This assumption might not be accurate, as in reality the curing time can differ markedly depending on the disease. Nevertheless, for the sake of simplicity, we retained this assumption to generate a random deviation of actual curing time and expectations and to estimate the utility losses based on the assumed deviations.

**4.2.6. Incorrect Treatment**

According to the results of a study conducted by Singh and colleagues (2014), the likelihood of incorrect treatment or false diagnosis is around 5.08%. Our model therefore includes the possibility of incorrect treatment and the resulting utility losses, assuming that the effectiveness of a treatment has a probability of 94.92% (100–5.08%).

**4.2.7. Appointment Scheduling**

The patient encounters (appointments) in our model were planned with respect to the average processing time plus the time buffer ( $p+T_b$ ). For example, if on average the processing of one patient takes 20 minutes and the physician plans a time buffer of five minutes before the next patient arrives, appointments will be scheduled every 25 minutes.

**4.2.8. Utility Factors**

First, we assume that  $b_t > b_p > b_w$  and  $c_t > c_p > c_w$ , which means that patients primarily want to be treated effectively. As this is the main reason for their visit to the physician in the first place, it seems reasonable to assume that this is the most important factor. Second, patients want to be informed about their condition at their own pace (i.e., the doctor’s patience is required), and this is more important than a short waiting time. To make this order of priorities effective in the utility calculation, the factors are explored ( $b_t=9, b_p=5, b_w=1$ ). Utility factors are also set in respect to  $b_x < c_x$  (the endowment effect), so  $|c_x| = |b_x| \times 2$ . To make this simulation a patient-centric model, all utility factors of the physician are set to 1, meaning that the utility of the patient is in focus, whereas the physician is indifferent with regard to the determinants important to him or her. This is also a reasonable assumption, as the physician’s job is to remain profitable without stress or unsatisfactory work. Consequently, we hold all determinants as equally important.

<sup>2</sup> A normal flu is usually cured in about 3–4 days, whereas curing a broken ankle might take weeks. Although the assumption is not

realistic and actual outcomes differ massively from disease to disease, for modeling purposes we simplify.

#### 4.2.9. Simulation Results

The results of the model are shown in Fig. 2 and Fig. 3. The graphs show the utility in relation to the applied average processing time and standard deviation of processing time, with a time buffer ( $T_b$ ) of 1 minute (Fig. 2) and 5 minutes (Fig. 3). A decrease in the average processing time results in an

increase in utility. If the average processing time ( $p$ ) is equal to or higher than the standard deviation of processing time ( $SDp$ ), the utility starts to decrease again. An increase in the time buffer shifts the utility curve, meaning that the utility is generally higher when a larger time buffer is applied. The lower the  $SDp$ , the higher the utility achieved.

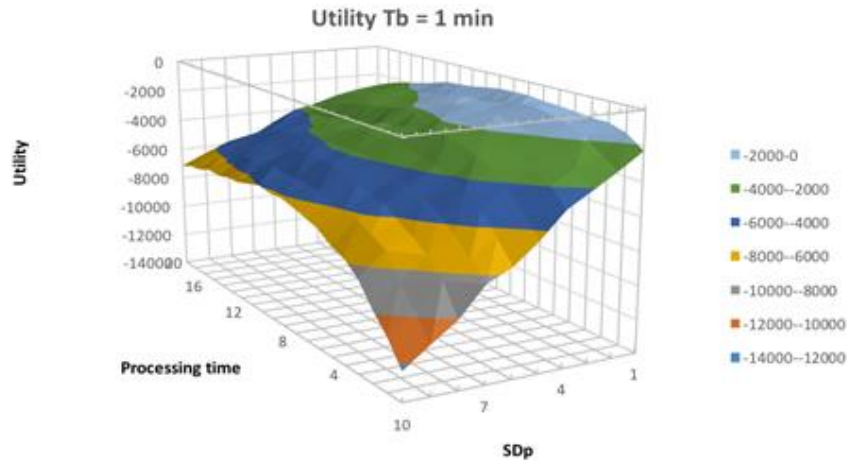


Figure 2. Simulation result, utility with time buffer of 1 minute

Source: developed by the authors

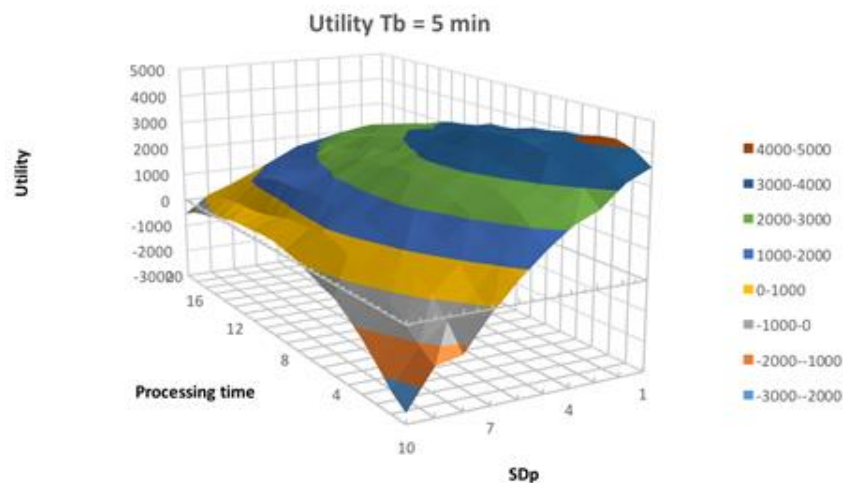


Figure 3. Simulation result, utility with time buffer of 5 minute

Source: developed by the authors

As the results of our simulation show, utility increases substantially if the system manages to reduce standard deviations for the patient encounter time. The higher the standard deviation, the lower the overall utility. This is likely to be due to a shift in encounter times, as each deviation causes a chain reaction where the other appointments have to be delayed or rescheduled. A standard deviation of only 10 minutes drives patient utility (and, thus, the overall system welfare) into negative values. Consequently, process standardization is of the highest priority (Bogodistov, Moormann, Sibbel, Krupskiy, & Hromtseva, 2021). The overall utility becomes negative if the processing time is greater than 15 minutes or less than about 7 minutes. This holds true in combination with a standard deviation of encounter time of about 7 minutes or higher. We conclude that not only processing times that are too long but also processing times that are too short are perceived by patients as negative outcomes of the process. This corresponds to measurements in Six Sigma projects where results

should always be within the boundaries of a lower and an upper specification limit (George, Maxey, Rowlands, & Upton, 2005). Indeed, patients value communication, and if the encounter is too short (e.g., due to longer processing time of a previous patient's encounter), the patient may perceive the communication as unsatisfactory (Greene, Adelman, Friedmann, & Charon, 1994; Like & Zyzanski, 1987).

#### 4.3. Study 2: AI-based Optimization

To analyze whether the two crucial determinants identified in Study 1, namely processing time and standard deviation of processing time, can be influenced by an AI-powered solution, we developed and tested a sample software solution. We used a chatbot system that relies on AI. The system was programmed and designed to assist a physician in primary diagnosis. The diagnostic AI was created by Infermedica<sup>3</sup> and

<sup>3</sup> We used the 2019 version of Infermedica-API. At the time of publication, a new version with new features is available: <https://developer.infermedica.com/docs/changelog>.



trained on well-founded medical literature and millions of patient records. The gained knowledge was checked and revised by a team of medical professionals and continues to be improved regularly. Entered symptoms are checked against the knowledge database of the AI, a list of possible diagnoses is generated, and further diagnosis questions are asked automatically.

#### 4.3.1. Patients' Interface

In our version of the chatbot, patients are able to enter their symptoms in a chat (for an example, see Fig. 4), and their chat opponent is AI-driven. The solution was placed on a university

server and was accessible by smartphone, tablet, personal computer, and laptop. The physician interface was accessed via tablet. The patient was guided through the diagnosis process by asking questions or telling him/her which data should be entered. The patient entered responses as free text. Afterwards, the chatbot asked diagnostic questions with predefined answer buttons “Yes,” “No,” and “Don't know”.

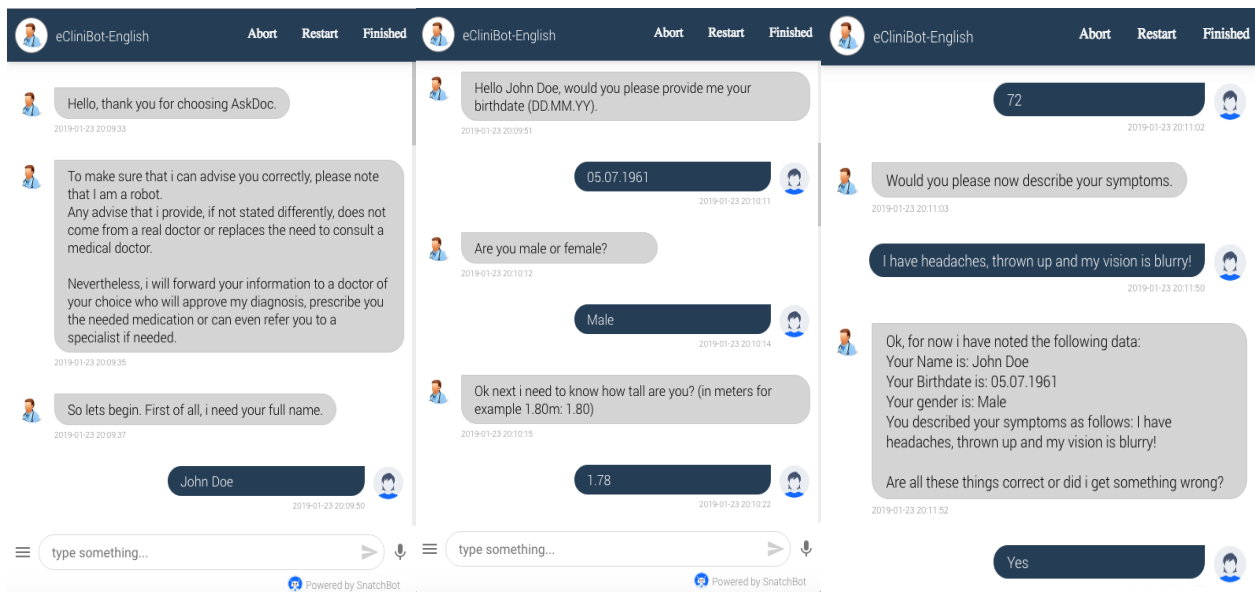


Figure 4. Simulation result, utility with time buffer of 5 minute

Source: developed by the authors

#### 4.3.2. Procedure

In order to test the system, we invited participants who remembered their last visit to a physician. Although the participants were not ill at the time of participation in the study, they recalled their symptoms and entered them as “actual” symptoms.

After the initial symptom description, the AI generated a list of possible diagnoses. By asking diagnostic questions, the AI excluded options from the list of possible diseases and proceeded with the most probable remaining diagnoses. When a specified probability threshold or question limit was reached, the diagnostic process was stopped. As a result, each patient was ascribed to one of the three categories:

- Self-care, i.e., the patient can cure the disease himself/herself and there is no need to see a physician;
- Consultation, i.e., the patient is advised to see a physician for approval; or
- Emergency, i.e., immediate medical care is needed.

All data, including the diagnoses with their respective probabilities, the triage scale (immediate, urgent, non-urgent), the patient's

description of symptoms, and the questions asked by the AI and answered by the patient, were forwarded to the physician. Until the physician had reviewed the data, no diagnosis was sent or shown to the participant.

#### 4.3.3. Physician's Interface

The physician role was played by a medical student who is in the final year of his studies. The physician used a Web-based interface where he could see all the patient records, i.e., the records of those who had recently entered their symptoms and those whom he had already diagnosed (Fig. 5).

This overview contained the name of the patient, the triage level, and the most likely diagnosis. Additionally, the status of the patient case was shown: pending, i.e., the diagnosis has not been reviewed by the physician, or diagnosed for already diagnosed patients.

By clicking on “View,” the physician was able to see all the data entered by the patient, including age, body mass index (BMI), allergies, blood pressure (optional), pulse (optional), the patient's description of his/her symptoms, and all the diagnostic questions with their respective answers (Fig. 6).

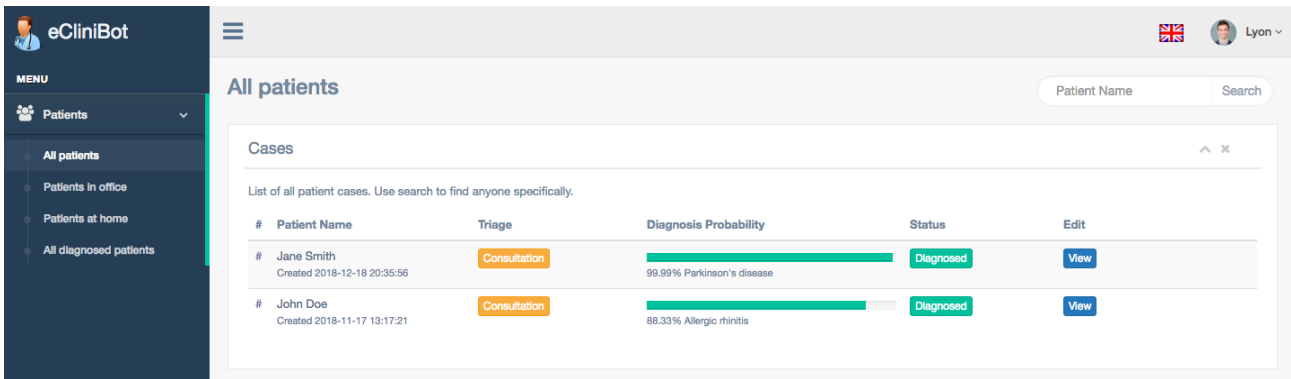


Figure 5. Physician's interface

Source: developed by the authors

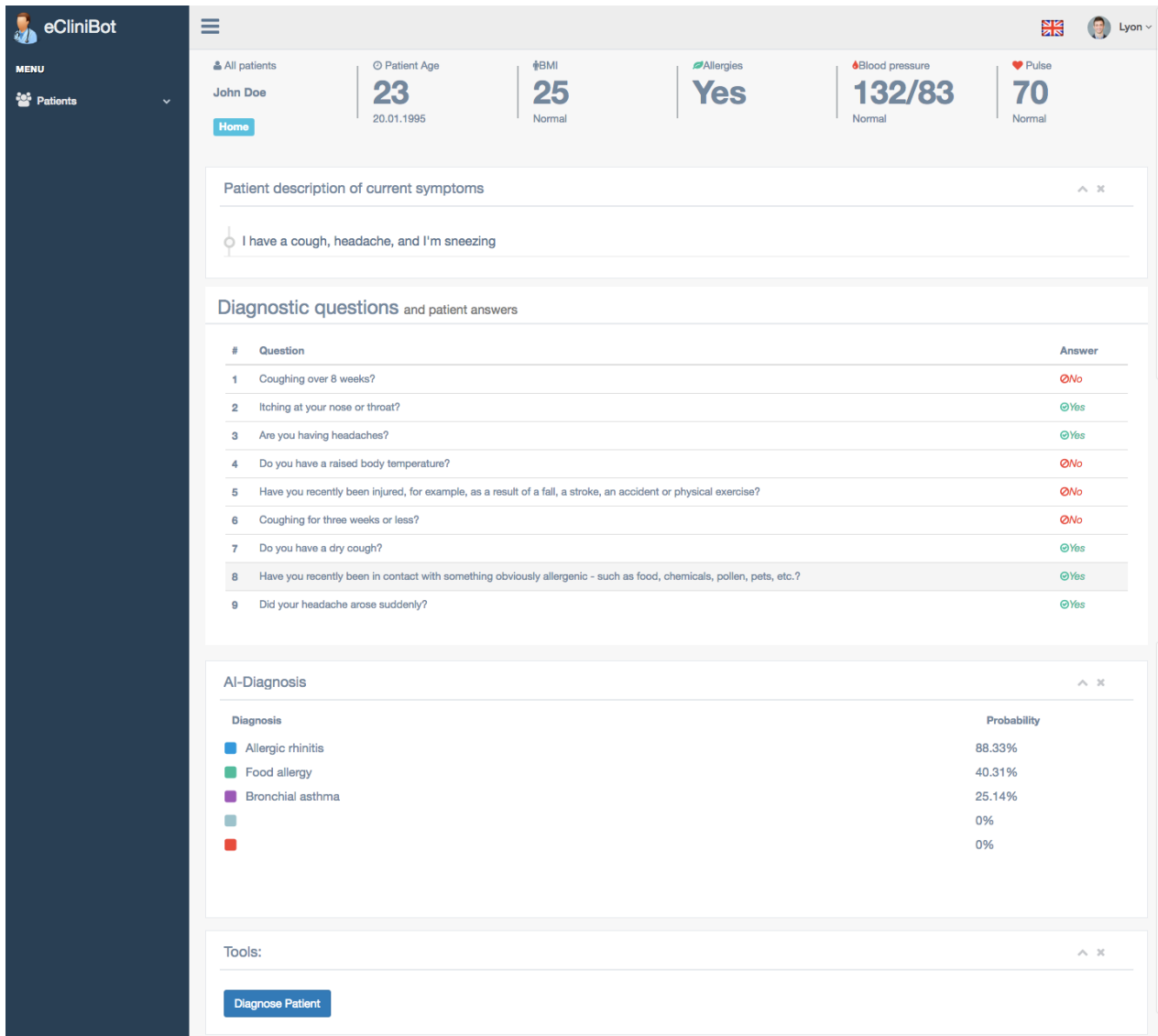
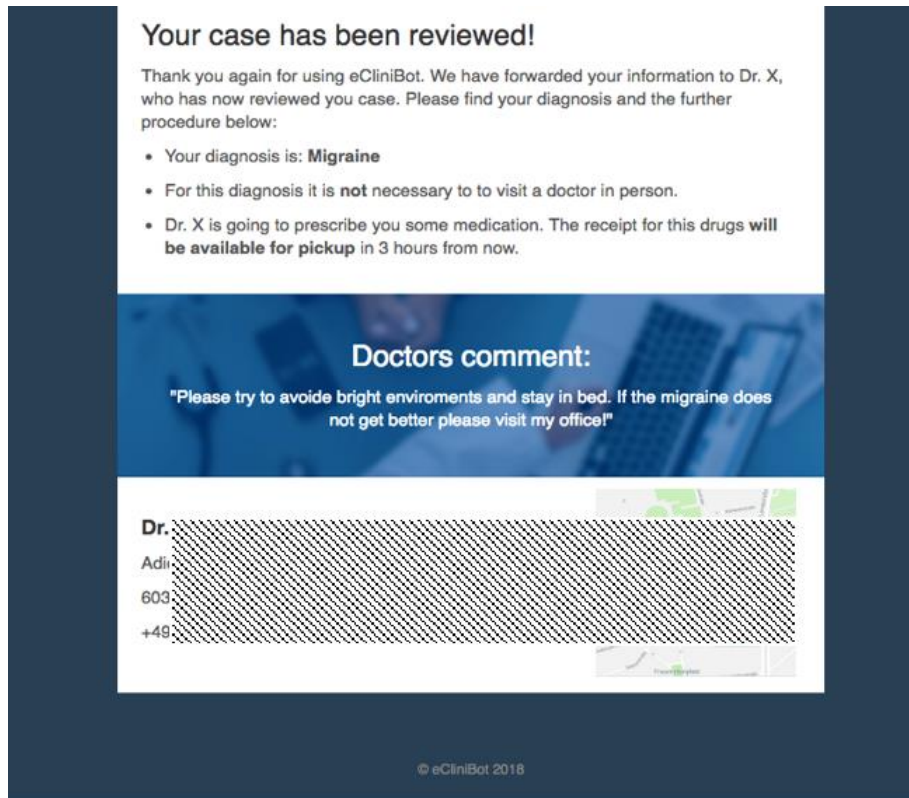


Figure 6. Detailed view of the patient's case

Source: developed by the authors

In terms of action opportunities, the physician could choose between the following options: stay at home, visit the doctor, emergency (i.e., call the emergency service). Furthermore, the physician was expected to add free text explaining the diagnosis and next steps to the patient. With a click on the diagnosis button,

these selections were saved in the database, and an e-mail was sent to the patient (Fig. 7).



Please note, that you have not visited a doctor in person. In case that your situation is getting worse and the advice of the doctor does not help, please consult a doctor in person. Our service aims to reduce waiting time for patients and workload of doctors. Nevertheless in some circumstances a doctor visit in person might be necessary to diagnose your condition correctly. In case of unclarities please contact the doctor. You can answer to this email or call the doctor directly. In case of emergency never use our service or write an e-mail to the doctor! Call your local emergency number!

Figure 7. E-mail sent to the patient

Source: developed by the authors

#### 4.3.4. Process Optimization

Within this example system, our participants were able to enter their symptoms from home. Thus, the staff knew about the condition of the patient in advance of his/her visit to the physician. In such circumstances, appointment scheduling becomes more reliable and straightforward, as the possible diagnosis is known (or can be assumed with a certain level of probability). Moreover, the physician can better organize visits in order to avoid cross-contamination between patients with different illnesses. Consequently, the deviation from the planned processing time should reduce. Likewise, waiting times for patients should be reduced, as fewer deviations from the scheduling plan are likely.

The average processing time is also assumed to be lower, as the standard diagnostic questions have been answered prior to the face-to-face interaction with the physician. Theoretically, this should save time and make the diagnosis more accurate, as the medical professional can then focus on the more sophisticated questions required. The burden of administrative work may remain unchanged. However, if a hospital managed to incorporate documentation from the chatbot system into their existing IT system, the workload would be reduced, as a crucial part of the documentation would be typed by the patient and/or processed by the AI-based system. With improved scheduling, less administrative work, and more time for each patient, stress levels of doctors should decrease. Accordingly, both the physician and the patient would benefit from improved accuracy and a faster, more reliable diagnostic process.

#### 4.3.5. System Testing

We tested our system with a group of 25 participants. Their mean age was 30 with a median of 24. Thirteen of the participants (52%) were male and twelve (48%) were female.

The participants were asked to enter symptoms that they had experienced in the past and that had prompted them to consult a medical professional. After they had used the system, the participants were asked to fill out a questionnaire that asked about their perception of the chatbot, whether they could imagine using a similar tool in the future, and what costs and benefits they would attribute to such a system. The participants were also asked to estimate the time to diagnosis when they visited the doctor and to enter the diagnosis of the physician. While the participants were using the chatbot, the duration of each chat was measured.

The medical student in the final year of his studies acted as the physician. He was asked to review the medical cases. He diagnosed all the patients on the basis of the data collected by the chatbot and the system's recommendations. This diagnostic procedure was executed using the Web interface of the developed system. We also measured the time from the point of opening a patient record to the finalization of that patient's diagnosis. The diagnosis of the physician and the diagnosis of the AI-based system were then compared with the actual diagnosis the patient received when s/he visited a physician. The AI-proposed diagnosis takes the form of a list of possible illnesses, sorted by their degree of probability. For our analysis, we took the first five diagnoses from the list and checked whether they corresponded to the patient's condition. As some conditions can only be diagnosed through physical

examination of the patient, the physician was asked to name the diagnostic test he would execute to confirm a provisional diagnosis. We used this explanation to determine whether the physician would have been able to diagnose the patient correctly.

#### 4.3.6. Process Optimization Results

All of the participants completed the chatbot diagnosis successfully. On average, the chat with the diagnosis AI took each patient 5 minutes and 12 seconds. The mean chat time was 5:02 minutes for male participants and 5:23 minutes for female participants. Our physician needed on average 4 minutes and 40 seconds to review each case. The chatbot was accurate in 76% of cases (meaning that the correct diagnosis was listed among the top five AI diagnoses). Male participants were diagnosed with an accuracy rate of 92%, whereas female participants were diagnosed with an accuracy rate of only 58%. With one exception, the diagnosis of the chatbot was always in the correct area. In other words, even if the correct diagnosis was not included in the list, the diagnoses listed by the system pointed in the direction of the body part or organ affected. The diagnostic accuracy of our physician was 100%, meaning that in each case his diagnosis was accurate or he named a diagnostic procedure that he would execute in a physical examination that would have resulted in the correct diagnosis. The physician asked for an additional personal consultation in 84% of cases. Only four patients were not required to see the doctor in person.

Most of the participants stated that they did not have the feeling that they had spoken to a human (64%). Nevertheless, most of them (88%) felt understood by the AI, and the questions asked were related to their disease or its symptoms (84%). In a small percentage of cases (16%), questions were asked repeatedly. A majority of the test-patients (68%) stated that they could imagine using a similar system for physician consultation in the future. If a real physician were definitely to be included in the diagnostic process, that figure rose to 76%. Nevertheless, only 64% would use such a system instead of visiting a physician in person. Participants believed that the main advantage of an AI-based system would be “reduced waiting time” for the patient (76%), followed by “time savings” from patients not always having to see a doctor in person (72%), “reduced workload” for the physician (72%), “faster treatment” (48%) and “fewer failures” of diagnosis (4%). When asked about the disadvantages, 88% of the participants stated that they thought such a system would result in “less personal contact” with the physician, while 28% believed that they would have “no personal contact” at all. Most participants assumed that AI-based systems would result in “more diagnostic errors” (68%). Increased workload for patients or physicians was attributed only once (4%). In reality, based on the answers of the test participants, the average time needed for diagnosing a patient, without waiting time, was 15:26 minutes, with the longest diagnosis needing 50 minutes and the shortest 2 minutes. The latter finding indicates an enormous growth in utility for both patients and physicians.

## 5. General Discussion

The simulation suggests that a shortened average processing time can increase the overall utility of the system. This is because less time spent on diagnosing a patient face to face means fewer actions are performed, which implies less administrative work. Moreover, if the processing time is generally decreased, the deviations from the processing time will also decrease.

Nevertheless, if appointments are scheduled without buffer time, the results will be deviations in waiting time, impatience for patients, and stress for the physician, which will cause negative utilities. This is the reason why utility starts to decrease again when  $p \leq SDp$ , i.e., when processing time is lower than or equal to the standard deviation in processing time. In contrast, when  $SDp$  becomes lower, treatments do not deviate much from the planned diagnosis time. Fewer delays occur, which reduces waiting time and stress for the physician and therefore increases utility. As a

patient’s disease is usually not known before s/he visits the physician, appointment scheduling is susceptible to uncertainty. If physicians rely in their planning on a medium processing time, they risk generating high waiting times and dissatisfaction due to deviations from their planned appointment slots. As the simulation results show, high utility can be achieved, but only with a low standard deviation of processing time.

From the test results, it appears that the diagnosis chatbot system saves time compared to the face-to-face diagnosis process. The physician spent an average of 16 minutes on diagnosing a patient face to face. On the other hand, the AI needed just 5 minutes for a diagnosis of 76% accuracy. Adding in the physician’s review time, the overall diagnosis time using the system took 10 minutes, which is a reduction of 37.5%. Nevertheless, in most cases, our physician needed a personal consultation for confirmation of the provisional diagnosis, and so additional diagnosis time was needed. Overall, patients perceived the chatbot-driven diagnosis as efficient and accurate and could imagine using such a system, especially if a real physician was also involved in the diagnostic process. We would like to emphasize here that we ran our model with the older version of 2019-API. The newer versions of the software incorporate COVID-19 symptoms along with several other updates. Researchers seeking to replicate our study should therefore use older versions of the API.

Participants mentioned the same advantages as those assumed previously in this paper (time savings, shorter waiting times, and less work for the physician). We can therefore conclude that patients recognize the benefits of such a system and would be willing to use it. Nevertheless, the questionnaire results indicated that patients are concerned about losing contact with their physician and that they fear being incorrectly diagnosed in the absence of a personal consultation. However, in most cases the physician called for a physical examination, which showed the diagnosis to be accurate; thus, these concerns would mostly be rejected.

As the computer model results show, a higher overall utility can be achieved by simultaneously decreasing the processing time and the standard deviation of the processing time. Our AI solution has been shown to decrease diagnosis time, as well as providing other opportunities for improvement. For instance, patients can enter their symptoms on a tablet provided by the hospital while waiting in the queue, or on a smartphone on the way to an appointment with a physician. This can enable medical professionals to plan an appropriate time slot for each patient, thereby reducing deviations from the planned processing time. Consequently, AI-driven systems in combination with a face-to-face visit to a physician may reduce waiting time and stress while also increasing the time available for more sophisticated analyses and tests. The resulting higher rates of accuracy and reduced waiting times could benefit both physicians and patients.

The results of our study will be of great interest for practitioners as well as researchers. First, as the standard deviations of time during patient encounters play a crucial role in patients’ utility and, thus, in the system welfare, reducing these deviations should be the main focus of hospitals. We recommend the (Lean) Six Sigma methodology, as it focuses on standard deviations in processes (Coleman, 2012; Corn, 2009; Proudlove, Moxham, & Boaden, 2008). We also emphasize that organizations such as hospitals should develop an organizational capability with regard to (Lean) Six Sigma instead of conducting ad hoc process optimizations. It is important to measure the capability permanently and rigorously (Bogodistov & Moormann, 2019) and to report results in an easily understandable manner (Bogodistov, 2017; Moormann, Antony, Chakraborty, Bogodistov, & Does, 2017). Research in the application of (Lean) Six Sigma in combination with AI-based systems seems to be very promising. Second, the application of AI-driven tools may reduce processing time. Moreover, part of the diagnosis procedure can be transferred from the encounter itself to other parts of the process, e.g., while the patient is waiting in a queue or sitting on public transport on the way to the encounter. Further research

should consider the linear and non-linear effects of AI-driven solutions and introduce both into queuing theory. We hold it especially important to investigate not only the process throughput but also the affective states of patients and physicians, as the quality of their interaction will change.

## 6. Limitations

Our model was run using a normal distribution for the treatment time. Although the normal distribution was found to be the best fit to the data, it performed poorly on the Kolmogorow–Smirnow, Anderson–Darling, and Chi-square tests, all of which indicated that the hypothesis that these data follow a normal distribution should be rejected. Using a different dataset might therefore result in different modeling results. Furthermore, the utility-based model assumes a fixed number of patients. If processing time can be reduced, the physician might be able to see more patients in a day, which would increase the system's utility. Future studies should include this option in the model.

Because of the structure of the model, the influence of each part of the utility calculation is determined by the utility determinants. If the utility determinants are set differently, the results will vary. Nevertheless, even if the applied utility assumptions do not hold, the improvements in utility will still be present. The influence on the overall utility depends, of course, on the relative (assumed) weight of the assumptions. The same applies to the different distribution of processing times; if the distribution deviates slightly from the one used in this paper, the results will vary, but the improvements should still be observable.

Our test of the system was executed on a small scale and with a medical student who does not diagnose primary care patients on a regular basis. The results of a test under everyday ambulatory care conditions with an experienced physician might vary from those reported in this paper. To determine the effects under real-life conditions, the system will have to be retested in future research. Nevertheless, the goal of this study (to establish whether even a simple AI-driven chatbot could help to reduce diagnosing time) has been achieved. Moreover, we have uncovered further issues for investigation, including trust in AI and willingness to have personal contact with the physician. We hope that further research on a larger scale will shed light on the fascinating topic of the use of AI in healthcare from the point of view of process management.

## 7. Funding

This study received no specific financial support.

## 8. Competing interests

The authors declare that they have no competing interests.

## References

- Advisor, I. G. (2018, July 24). Global Views on Healthcare in 2018. Ipsos Global Advisor. Retrieved from <https://www.ipsos.com/en-be/global-views-healthcare>
- Ahmadi-Javid, A., Jalali, Z., & Klassen, K. J. (2017). Outpatient appointment systems in healthcare: A review of optimization studies. *European Journal of Operational Research*, 258(1): 3–34. [doi:10.1016/j.ejor.2016.06.064](https://doi.org/10.1016/j.ejor.2016.06.064).
- Alexopoulos, C., Goldsman, D., Fontanesi, J., Kopald, D., & Wilson, J. R. (2008). Modeling patient arrivals in community clinics. *Omega*, 36(1), 33–43. [doi:10.1016/j.omega.2005.07.013](https://doi.org/10.1016/j.omega.2005.07.013).
- Allon, G., & Kremer, M. (2018). Behavioral Foundations of Queueing Systems. *The Handbook of Behavioral Operations*, 323–366. [doi:10.1002/9781119138341.ch9](https://doi.org/10.1002/9781119138341.ch9).
- Bhavnani, S. P., Narula, J., & Sengupta, P. P. (2016). Mobile technology and the digitization of healthcare. *European Heart Journal*, 37(18), 1428–1438. [doi:10.1093/eurheartj/ehv770](https://doi.org/10.1093/eurheartj/ehv770).
- Bogodistov, Y. (2017). Example of an individual report on Lean Six Sigma. no. 2017-1, ProcessLab. Retrieved from [https://www.researchgate.net/profile/Yevgen-Bogodistov/publication/320298556\\_Example\\_of\\_an\\_individual\\_report\\_on\\_Lean\\_Six\\_Sigma/links/59dc8384458515e9ab4c67d2/Example-of-an-individual-report-on-Lean-Six-Sigma.pdf](https://www.researchgate.net/profile/Yevgen-Bogodistov/publication/320298556_Example_of_an_individual_report_on_Lean_Six_Sigma/links/59dc8384458515e9ab4c67d2/Example-of-an-individual-report-on-Lean-Six-Sigma.pdf).
- Bogodistov, Y., & Moormann, J. (2019). Theorizing on Operational Excellence: A Capability-Based Approach. *Academy of Management Proceedings*, 2019(1), 12174. [doi:10.5465/ambpp.2019.12174abstract](https://doi.org/10.5465/ambpp.2019.12174abstract).
- Bogodistov, Y., Moormann, J., & Sibbel, R. (2018). Beyond Health Care Reform: How Process Management Can Alter Patients' Experience. *Academy of Management Proceedings*, 2018(1), 12544. [doi:10.5465/ambpp.2018.12544abstract](https://doi.org/10.5465/ambpp.2018.12544abstract).
- Bogodistov, Y., Moormann, J., Sibbel, R., Krupskiy, O. P., & Hromtseva, O. (2021). Process maturity and patient orientation in times of a health system reform. *Business Process Management Journal*, ahead-of-print (ahead-of-print). [doi:10.1108/BPMJ-09-2020-0428](https://doi.org/10.1108/BPMJ-09-2020-0428).
- Bohr, A., & Memarzadeh, K. (2020). Current healthcare, big data, and machine learning. *Artificial Intelligence in Healthcare*, 1–24. [doi:10.1016/b978-0-12-818438-7.00001-0](https://doi.org/10.1016/b978-0-12-818438-7.00001-0).
- Kamel Boulos, M. N., Brewer, A. C., Karimkhani, C., Buller, D. B., & Dellavalle, R. P. (2014). Mobile medical and health apps: state of the art, concerns, regulatory control and certification. *Online Journal of Public Health Informatics*, 5(3). e229. [doi:10.5210/ijphi.v5i3.4814](https://doi.org/10.5210/ijphi.v5i3.4814).
- Brennan, V. K., & Dixon, S. (2013). Incorporating Process Utility into Quality Adjusted Life Years: A Systematic Review of Empirical Studies. *Pharmacoeconomics*, 31(8), 677–691. [doi:10.1007/s40273-013-0066-1](https://doi.org/10.1007/s40273-013-0066-1).
- Carayon, P., & Hoonakker, P. (2019). Human Factors and Usability for Health Information Technology: Old and New Challenges. *Yearbook of Medical Informatics*, 28(01), 071–077. [doi:10.1055/s-0039-1677907](https://doi.org/10.1055/s-0039-1677907).
- Carlucci, D., Renna, P., & Schiuma, G. (2012). Evaluating service quality dimensions as antecedents to outpatient satisfaction using back propagation neural network. *Health Care Management Science*, 16(1), 37–44. [doi:10.1007/s10729-012-9211-1](https://doi.org/10.1007/s10729-012-9211-1).
- Coleman, S. Y. (2012). Six Sigma in Healthcare. *Statistical Methods in Healthcare*, 286–308. [doi:10.1002/9781119940012.ch14](https://doi.org/10.1002/9781119940012.ch14).
- Corn, J. B. (2009). Six sigma in health care. *Radiologic technology*, 81(1), 92–95.
- Coulter, A., & Jenkinson, C. 2005. European patients' views on the responsiveness of health systems and healthcare providers. *European Journal of Public Health*, 15(4), 355–360. [doi:10.1093/eurpub/ckio04](https://doi.org/10.1093/eurpub/ckio04).
- Cutler, D. M. (2007). The lifetime costs and benefits of medical technology. *Journal of Health Economics*, 26(6), 1081–1100. [doi:10.1016/j.jhealeco.2007.09.003](https://doi.org/10.1016/j.jhealeco.2007.09.003).
- Dayer, L., Heldenbrand, S., Anderson, P., Gubbins, P. O., & Martin, B. C. (2013). Smartphone medication adherence apps: Potential benefits to patients and providers: Response to Aungst. *Journal of the American Pharmacists Association*, 53(4), 345. [doi:10.1331/japha.2013.13121](https://doi.org/10.1331/japha.2013.13121).

- Ettinger, W. H. (1998). Consumer-Perceived Value: The Key to a Successful Business Strategy in the Healthcare Marketplace. *Journal of the American Geriatrics Society*, 46(1), 111–113. doi:10.1111/j.1532-5415.1998.tb01024.x.
- Fuchs, V. R. (1996). Economics, values, and health care reform. *American Economic Review*, 86(1), 1–24. Retrieved from [https://web.stanford.edu/~jay/health\\_class/Readings/Lecture\\_01/fuchs\\_health\\_survey.pdf](https://web.stanford.edu/~jay/health_class/Readings/Lecture_01/fuchs_health_survey.pdf).
- George, D., & Mallery, P. (2019). *IBM SPSS statistics 26 step by step: A simple guide and reference* (6th Edition). New York, NY: Routledge. Retrieved from <https://www.taylorfrancis.com/books/mono/10.4324/9780429056765/ibm-spss-statistics-26-step-step-darren-george-paul-mallery>.
- George, M., Maxey, J., Rowlands, D., & Upton, M. (2005). *The Lean Six Sigma Pocket Toolbook: A Quick Reference Guide to Nearly 100 Tools for Improving Process Quality, Speed, and Complexity*. New York, NY: McGraw Hill. Retrieved from <https://play.google.com/store/books/details?id=ssab5W-exREC>.
- Graber, M. L. (2013). The incidence of diagnostic error in medicine. *BMJ Quality & Safety*, 22(Suppl 2), ii21–ii27. doi:10.1136/bmjqs-2012-001615.
- Greene, M. G., Adelman, R. D., Friedmann, E., & Charon, R. (1994). Older patient satisfaction with communication during an initial medical encounter. *Social Science & Medicine*, 38(9), 1279–1288. doi:10.1016/0277-9536(94)90191-0.
- Haimi, M., Brammli-Greenberg, S., Waisman, Y., & Baron-Epel, O. (2018). Physicians' experiences, attitudes and challenges in a Pediatric Telemedicine Service. *Pediatric Research*, 84(5), 650–656. doi:10.1038/s41390-018-0117-6.
- Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (Eds.). (2010). *Multivariate data analysis (7th ed)*. Upper Saddle River, NJ: Prentice Hall.
- Hall, A. M., Ferreira, P. H., Maher, C. G., Latimer, J., & Ferreira, M. L. (2010). The Influence of the Therapist-Patient Relationship on Treatment Outcome in Physical Rehabilitation: A Systematic Review. *Physical Therapy*, 90(8), 1099–1110. doi:10.2522/ptj.20090245.
- Hassin, R., & Haviv, M. (2003). To Queue or Not to Queue: Equilibrium Behavior in Queueing Systems. doi:10.1007/978-1-4615-0359-0.
- Herzlinger, R. E. (2006). Why innovation in health care is so hard. *Harvard business review*, 84(5), 58–66. Retrieved from <https://www.hse.ie/eng/about/who/healthbusinessservices/hbs-news-and-events/why-innovation-in-health-care-is-so-hard.pdf>.
- Hudak, P. L., Hogg-Johnson, S., Bombardier, C., McKeever, P. D., & Wright, J. G. (2004). Testing a New Theory of Patient Satisfaction With Treatment Outcome. *Medical Care*, 42(8), 726–739. doi:10.1097/01.mlr.0000132394.09032.81.
- Krittawong, C., Zhang, H., Wang, Z., Aydar, M., & Kitai, T. (2017). Artificial Intelligence in Precision Cardiovascular Medicine. *Journal of the American College of Cardiology*, 69(21), 2657–2664.
- Krittawong, C., Zhang, H., Wang, Z., Aydar, M., & Kitai, T. (2017). Artificial Intelligence in Precision Cardiovascular Medicine. *Journal of the American College of Cardiology*, 69(21), 2657–2664. doi:10.1016/j.jacc.2017.03.571.
- Lehmann, C. A., Mintz, N., & Giacini, J. M. (2006). Impact of Telehealth on Healthcare Utilization by Congestive Heart Failure Patients. *Disease Management & Health Outcomes*, 14(3), 163–169. doi:10.2165/00115677-200614030-00005.
- Levine, R., Shore, K., Lubalin, J., Garfinkel, S., Hurtado, M., & Carman, K. (2012). Comparing physician and patient perceptions of quality in ambulatory care. *International Journal for Quality in Health Care*, 24(4), 348–356. doi:10.1093/intqhc/mzso23.
- Like, R., & Zyzanski, S. J. (1987). Patient satisfaction with the clinical encounter: Social psychological determinants. *Social Science & Medicine*, 24(4), 351–357. doi:10.1016/0277-9536(87)90153-5.
- Marmor, T., & Wendt, C. (2012). Conceptual frameworks for comparing healthcare politics and policy. *Health Policy*, 107(1), 11–20. doi:10.1016/j.healthpol.2012.06.003.
- MathWave Technologies. (2019). *EasyFit - Software für Verteilungsanpassung*. Retrieved from <http://www.mathwave.com/de/home.html>.
- Miller, D. D., & Brown, E. W. (2018). Artificial Intelligence in Medical Practice: The Question to the Answer? *The American Journal of Medicine*, 131(2), 129–133. doi:10.1016/j.amjmed.2017.10.035.
- Moormann, J., Antony, J., Chakraborty, A., Bogodistov, Y., & Does, R. (2017). *Lean Six Sigma in the Financial Services Industry: Germany*. doi:10.13140/RG.2.2.27000.55043.
- Moreira, M. W. L., Rodrigues, J. J. P. C., Korotaev, V., Al-Muhtadi, J., & Kumar, N. (2019). A Comprehensive Review on Smart Decision Support Systems for Health Care. *IEEE Systems Journal*, 13(3), 3536–3545. doi:10.1109/jsyst.2018.2890121.
- Peprah, A. A. (2013). *Healthcare Delivery in Sub-Saharan Africa: Patients' Satisfaction and Perceived Service Quality, A Case Study of Sunyani Regional Hospital in Ghana*. LAP Lambert Academic Publishing. Retrieved from [https://books.google.com/books/about/Healthcare\\_Delivery\\_in\\_Sub\\_Saharan\\_Afric.html?hl=&id=\\_wO4ngEACAAJ](https://books.google.com/books/about/Healthcare_Delivery_in_Sub_Saharan_Afric.html?hl=&id=_wO4ngEACAAJ).
- Plsek, P. E., & Greenhalgh, T. (2001). Complexity science: The challenge of complexity in health care. *BMJ*, 323(7313), 625–628. doi:10.1136/bmj.323.7313.625.
- Polisena, J., Tran, K., Cimon, K., Hutton, B., McGill, S., & Palmer, K. (2009). Home telehealth for diabetes management: a systematic review and meta-analysis. *Diabetes, Obesity and Metabolism*, 11(10), 913–930. doi:10.1111/j.1463-1326.2009.01057.x.
- Polisena, J., Tran, K., Cimon, K., Hutton, B., McGill, S., Palmer, K., & Scott, R. E. (2010). Home telehealth for chronic obstructive pulmonary disease: a systematic review and meta-analysis. *Journal of Telemedicine and Telecare*, 16(3), 120–127. doi:10.1258/jtt.2009.090812.
- Proudlove, N., Moxham, C., & Boaden, R. (2008). Lessons for Lean in Healthcare from Using Six Sigma in the NHS. *Public Money & Management*, 28(1): 27–34.
- Rau, C.-L., Tsai, P.-F. J., Liang, S.-F. M., Tan, J.-C., Syu, H.-C., Jheng, Y.-L., ... Jaw, F.-S. (2013). Using discrete-event simulation in strategic capacity planning for an outpatient physical therapy service. *Health Care Management Science*, 16(4), 352–365. doi:10.1007/s10729-013-9234-2.
- Rubel, J. A., Bar-Kalifa, E., Atzil-Slonim, D., Schmidt, S., & Lutz, W. (2018). Congruence of therapeutic bond perceptions and its relation to treatment outcome: Within- and between-dyad effects. *Journal of Consulting and Clinical Psychology*, 86(4), 341–353. doi:10.1037/ccp0000280.
- Saltman, R. B., & Figueras, J. (1997). *European health care reform: analysis of current strategies*. WHO Regional Publications. European Series, 72, 5–38.
- Semigran, H. L., Linder, J. A., Gidengil, C., & Mehrotra, A. (2015). Evaluation of symptom checkers for self diagnosis and triage: audit study. *BMJ*, h3480. doi:10.1136/bmj.h3480.

- Singh, H., Meyer, A. N. D., & Thomas, E. J. (2014). The frequency of diagnostic errors in outpatient care: estimations from three large observational studies involving US adult populations. *BMJ Quality & Safety*, 23(9), 727–731. [doi:10.1136/bmjqs-2013-002627](https://doi.org/10.1136/bmjqs-2013-002627).
- Sinsky, C., Colligan, L., Li, L., Prgomet, M., Reynolds, S., Goeders, L., ... Blike, G. (2016). Allocation of Physician Time in Ambulatory Practice: A Time and Motion Study in 4 Specialties. *Annals of Internal Medicine*, 165(11), 753–761. [doi:10.7326/m16-0961](https://doi.org/10.7326/m16-0961).
- Teke, A., Cengiz, E., Çetin, M., Demir, C., Kirkbir, F., & Fedai, T. (2010). Analysis of the Multi-Item Dimensionality of Patients' Perceived Value in Hospital Services. *Journal of Medical Systems*, 36(3), 1301–1307. [doi:10.1007/s10916-010-9590-0](https://doi.org/10.1007/s10916-010-9590-0).
- Thompson, D. A., Yarnold, P. R., Williams, D. R., & Adams, S. L. (1996). Effects of Actual Waiting Time, Perceived Waiting Time, Information Delivery, and Expressive Quality on Patient Satisfaction in the Emergency Department. *Annals of Emergency Medicine*, 28(6), 657–665. [doi:10.1016/s0196-0644\(96\)70090-2](https://doi.org/10.1016/s0196-0644(96)70090-2).
- Varkevisser, M., van der Geest, S. A., & Schut, F. T. (2012). Do patients choose hospitals with high quality ratings? Empirical evidence from the market for angioplasty in the Netherlands. *Journal of Health Economics*, 31(2), 371–378. [doi:10.1016/j.jhealeco.2012.02.001](https://doi.org/10.1016/j.jhealeco.2012.02.001).
- Williams, E. S., Manwell, L. B., Konrad, T. R., & Linzer, M. (2007). The relationship of organizational culture, stress, satisfaction, and burnout with physician-reported error and suboptimal patient care. *Health Care Management Review*, 32(3), 203–212. [doi:10.1097/01.hmr.0000281626.28363.59](https://doi.org/10.1097/01.hmr.0000281626.28363.59).



This is an open access journal and all published articles are licensed under a **Creative Commons «Attribution» 4.0**.