AI-Enabled Decision Support in Health Care

Kumulative Dissertation
an der Wirtschaftswissenschaftlichen Fakultät
der Universität Augsburg
zur Erlangung des Grades eines
Doktors der Wirtschaftswissenschaften
(Dr. rer. pol.)

vorgelegt von
Julian Schiele, M. Sc.

Erstgutachter:	Prof. Dr. Jens O. Brunner
Zweitgutachter:	Prof. Dr. Robert Klein
Drittgutachter:	Prof. Dr. Axel Tuma
Vorsitzende der mündlichen Prüfung:	Prof. Dr. Jennifer Kunz
Tag der mündlichen Prüfung:	30. Juni 2020

Augsburg, 30. Juni 2020

List of contributions

This thesis contains the following contributions published in or submitted to scientific journals. The specified categories relate to the journal ranking VHB-JOURQUAL3 of the Verband der Hochschullehrer für Betriebswirtschaft e.V. [2]. The order of the contributions corresponds to the order of print in this thesis.

Contribution 1

Schiele, J., Brunner, J.O. (2019): A Framework for the Hospital of the Future.

Status: Submitted to Journal of Business Economics, category B.
Date of submission: December 30, 2019.

Contribution 2

Schiele, J., Koperna, T., Brunner, J.O. (2019): Predicting ICU Bed Occupancy for Integrated Operating Room Scheduling via Neural Networks. Naval Research Logistics. `https://doi.org/10.1002/nav.21929`.

Status: Accepted for publication in Naval Research Logistics, category B.
Date of submission: December 28, 2018.
Date of submission (major revision): June 26, 2019.
Date of submission (major revision): September 20, 2019
Date of submission (minor revision): March 31, 2020
Date of submission (minor revision): June 1, 2020
Date of acceptance: June 8, 2020.

Contribution 3

Schiele, J. (2019): Predicting Surgical Durations and Implications at the Operational Level.

Status: Submitted to Medical Decision Making, category A.
Date of submission: December 30, 2019.

Contribution 4

Schiele, J., Rabe, F., Schmitt, M., Glaser, M., Häring, F., Brunner, J.O., Bauer, B., Schuller, B., Traidl-Hoffmann, C., Damialis, A. (2019): Automated Classification of Airborne Pollen using Neural Networks. 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), 4474–4478. `https://doi.org/10.1109/EMBC.2019.8856910`.

Status: Accepted for publication in EMBC, unranked.
Date of submission: February 15, 2019.
Date of acceptance: April 10, 2019.

We know the past
but cannot control it.
We control the future
but cannot know it.

CLAUDE SHANNON

Contents

1 Introduction

Mainly driven by big data and artificial intelligence (AI), the health care system is on the cusp of a new era, offering the potential to mitigate today's imminent challenges. This thesis explores how AI can support decision making in health care with a particular focus on hospitals.

1.1 Motivation

Today's health care systems around the globe face imminent social, organizational, medical, and financial challenges. Among the most critical ones are demographic change, legal regulations, operational inefficiencies, acute staff shortage, unpleasant working conditions, investment backlog, and increasing cost pressure. In most countries, the health care industry accounts for a large share of expenditures and faces ongoing growth. In the United States, 3.3 trillion USD or nearly 18% of its gross domestic product were spent on health care in 2016[1] reflecting an annual growth of 4.3% compared to 2015 [12]. A closer look reveals that hospitals are a key driver accounting for nearly one third of those health expenditures[2] [12]. With rising costs, hospitals are increasingly attracting attention from sponsors in both the governmental and the private sector demanding more cost effectiveness while ensuring the same level of service quality. "Pressures to make operating margins will continue to be at the forefront of most hospital and health system leaders' minds" [38] - particularly, since it seems that nothing "will stop public spending on health

[1] about 9% in the OECD countries [39] and 11% in Germany [19]
[2] nearly 40% in the OECD countries [39]

care from rising" [41]. Within a hospital, the operating room is the most expensive resource followed by the intensive care unit (ICU). A cost break-down for the United States is depicted in Figure 1.1: 1.1 trillion USD were spent on hospitals, mainly driven by operating room expenditures (40%, depicted in red) and ICU expenditures (15%, depicted in orange). Consequently, this thesis studies decision making in hospitals (see Section 2.1) with a particular focus on the operating room and the ICU (see Sections 2.2 and 2.3).

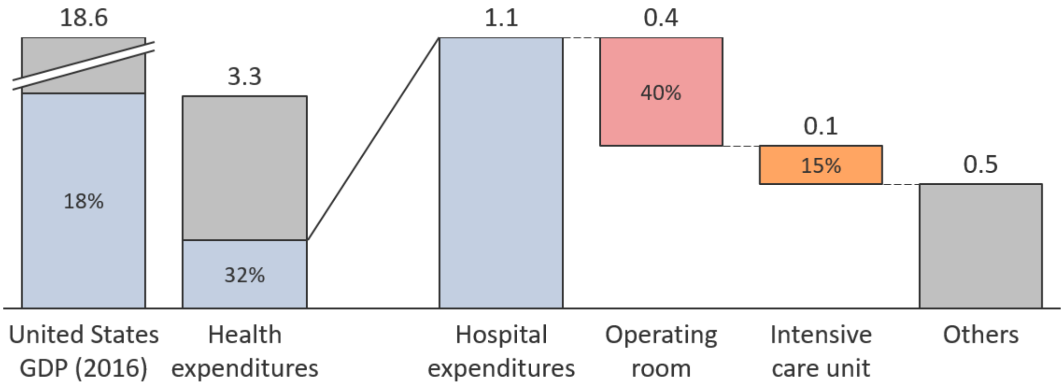

Figure 1.1: United States' health expenditures 2016 in trillion USD. Within a hospital, the operating room (depicted in red) and the intensive care unit (depicted in orange) are the most expensive resources representing around 40% and 15% of total hospital expenditures, respectively.

According to Berwick *et al.* [6], at least 20% of the health care expenditures could be eliminated by addressing overtreatment, failures in coordination and execution of care processes, inefficient pricing, administrative complexity, fraud and abuse. Health care operations management has emerged as a key discipline to increase the operational efficiency founded on a data-driven, mathematical approach [11]. Decision making in operating room management can be categorized into three different hierarchical decision levels [26]: the strategic level, the tactical level, and the operational level. At the strategic level, hospitals decide on the case mix planning and long-term ambitions. At the tactical level, available resources are allocated to medical specialties within the strategic boundaries. At the operational level, the individual surgeries of patients are scheduled on a daily basis. This thesis leverages the potential of big data and AI to support decision making in hospitals, particularly at the tactical (see Section 2.2) and the operational (see Section 2.3) level.

1.2 Introduction to machine learning

Within the last decades, AI has become ubiquitous and is now transforming many industries. Most prominently, machine learning is rapidly advancing and paves the way for numerous “intelligent” applications - from Arthur Samuel’s checkers playing program [43] in the 1960s to today’s autonomous vehicles, drug discovery, and precision medicine. When speaking about AI, one frequently refers to machine learning which represents a branch of AI (see Figure 1.2) that systematically applies algorithms to learn relationships from underlying data. Although the ideas behind

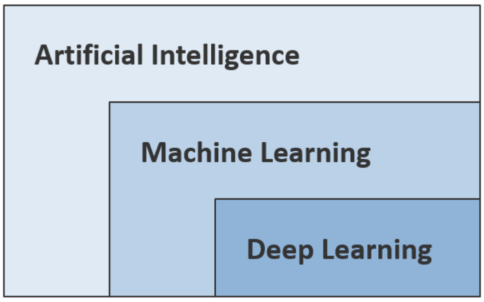

Figure 1.2: Schematic representation of artificial intelligence, machine learning, and deep learning.

machine learning are not new, it was only recently that its potential could be exploited thanks to the availability of big data, unprecedented analytical power, and powerful frameworks. Arthur Samuel defined machine learning as the “field of study that gives computers the ability to learn without being explicitly programmed” [43] and Tom M. Mitchell stated the following definition: “A computer program is said to learn from experience E with respect to some task T and some performance measure P, if its performance on T, as measured by P, improves with experience E” [37]. Nowadays, machine learning is widely used in everyday applications such as spam e-mail detection, face recognition, weather forecasting, movie recommendations on Netflix, and voice recognition systems such as Amazon’s Alexa.

In this thesis, we focus on AI’s branch of machine learning. In general, it can be categorized by learning type, problem type, and algorithm type.

Learning type. Commonly, one distinguishes between *supervised learning*, *unsupervised learning*, and *reinforcement learning*. In supervised learning, the correct labels are known for a set of historical samples such that the underlying relationship can be learned. Unsupervised learning has the focus to identify patterns without the knowledge of labels and reinforcement learning aims to balance between exploration and exploitation. In this thesis, we focus on supervised learning.

Problem type. Commonly, one distinguishes between *regression problems* and *classification problems*. While continuous valued output is predicted in the former, it is discrete valued output in the latter. In this thesis, we study regression (see Section 2.2) as well as classification problems (see Sections 2.3 and 2.4).

Algorithm type. Depending on learning type and problem type, various algorithms can be applied. For example, *linear regression* is suitable for supervised regression problems and *logistic regression* for supervised classification problems. Other powerful algorithms include *support vector machines*, *k-nearest neighbors*, *decision trees*, *random forests*, and most prominently *neural networks*. In this thesis, we explore various algorithms, but mainly focus on neural networks.

In the following, a short description of neural networks shall be provided. More details can be found in [4], [25], and [31]. In supervised learning, it is the aim of the model to find a computable function $\boldsymbol{x} \mapsto f(\boldsymbol{x}) = \boldsymbol{y}$ where input vectors $\boldsymbol{x}$, i.e., so-called *features*, and output vectors $\boldsymbol{y}$, i.e., so-called *labels*, are known for some historical samples $\left(\boldsymbol{x}^{(m)}, \boldsymbol{y}^{(m)}\right), m \in \mathcal{M}$. Note, that the output vector $\boldsymbol{y}$ decays to a scalar y for regression problems. The prediction function $\hat{\boldsymbol{y}}$ is obtained as output of the neural network that is composed of *neurons* as fundamental building blocks. Figure 1.3 depicts a neuron that maps the three inputs x_1, x_2, and x_3 to one output $\hat{y}$ (omitting the index m). In a neuron, three mathematical operations are performed. In Eq. (1.1), the dot product between the input vector $\boldsymbol{x}$ and the corresponding weight vector $\boldsymbol{w}$ is computed and a bias b is added. In Eq. (1.2), an activation function $g(z)$ is applied to the term resulting in the prediction $\hat{y}$.

$$z = \boldsymbol{w}\boldsymbol{x}^{\mathrm{T}} + b \tag{1.1}$$

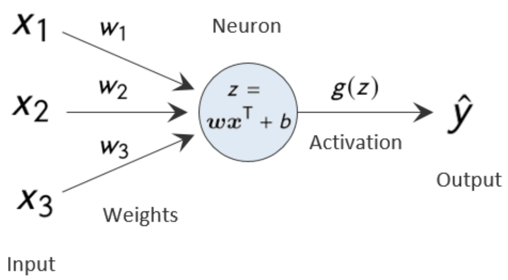

Figure 1.3: Schematic representation of neuron as fundamental building block.

$$\hat{y} = g(z) \tag{1.2}$$

Common activation functions are the sigmoid function, ReLu function, and tanh function (see Chapter 6 in [25]). Neurons are then combined to form larger structures, so-called neural networks. Figure 1.4 depicts a neural network with one *input layer*, several *hidden layers*, and one *output layer*. A neural network's topology is defined

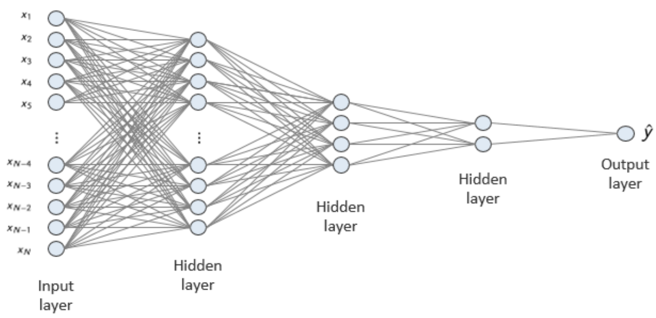

Figure 1.4: Schematic representation of a neural network.

by the number of layers and the number of neurons in each layer. Hereafter, we use the form [10:4:2] to indicate the number of neurons in the hidden layers. *Deep learning* refers to neural networks with multiple hidden layers and is a branch of machine learning that is particularly powerful for image classification (see Figure 1.2). Commonly, the set of historical samples is split into a *training set*, a *validation set*, and a *test set*. The former is used to train the network, i.e., backpropagation adjusts the weights such that the deviation between measured and predicted labels is minimized. Validation and test set are used to tune the hyperparameters and measure the predictive performance, respectively.

1.3 Machine learning in health care

Machine learning has the potential to radically transform health care for three reasons. First, the health care system faces imminent challenges, particularly staff shortage, cost pressure, and pressure to improve outcomes (see Section 1.1). Second, there is no shortage of health care data, including electronic health records, clinical trials, magnetic resonance imaging, and the constantly increasing number of information from monitoring devices such as consumer wearables (see Section 3.3). Third, recent advances in computational power make it possible to fully leverage that data with new, powerful algorithms in order to improve many aspects of drug discovery, preventive care, diagnosis, intervention, medication, decision making, risk assessment, operational and administrative processes.

Already today, machine learning helps to address inefficiencies and create new opportunities in health care. By 2022, spending on AI-related tools is expected to reach USD 8 billion annually across seven areas: remote prevention and care, diagnostics support, treatment pathways and support, drug discovery and development, operations, marketing and sales, and support functions [3]. Most prominently, machine learning is used to advance diagnosis and detection, e.g., to classify skin cancer [18], to predict pneumonia [28], to classify fetal heart rates [32], to predict diseases [13] and colorectal cancer outcome based on tissue samples [9], to identify neurodegenerative diseases [29], to suggest referrals in retinal disease [21], and to detect schizophrenia [36]. Furthermore, it is used to support robot-assisted surgeries [52], advance drug discovery [53], and for numerous other applications [38, 48]. Eventually, machine learning also advances decision support in hospital operations, e.g., to predict surgical durations [20, 49], to assess perioperative cardiac risks [30], to identify patients at risk of postinduction hypotension [33], to predict implantation outcome of individual embryos [50], to predict clinical deterioration [15] and other clinical events [17].

Recently, numerous health tech startups become attracted by the new opportunities offered by machine learning, e.g., to discover new drugs (Atomwise), to improve patient experience (Babylon Health), to enable robot-assisted surgery (Auris Health), and to analyze medical imaging (Zebra) [3]. An overview of the current landscape with more than 90 promising AI startups transforming the health care sector is provided in [1] and [16]. Given their consumer-friendliness, digital sophistication, and record amount of unspent capital, also big tech players have recognized their potential to create lasting impact and recently started to heavily invest in health care, e.g., IBM's Watson helps health care professionals to manage and harvest medical data, Apple builds a platform for healthcare and wellness around its Apple Health Record and Apple Watch, and Google develops an augmented reality microscope to diagnose cancer in real-time [14]. In fact, Google's CEO Sundar Pichai announced that "healthcare is one of the most important fields AI is going to transform" (Google I/O 2018 keynote).

1.4 Organization of this thesis

While there is a large body of research on AI-enabled decision support in health care, there are still a few research gaps. The following open research questions are answered by this thesis.

1. Can ideas from Industry 4.0 help mitigate the imminent challenges in health care?
2. How can hospitals navigate the variety of use cases on the way to the hospital of the future?
3. What role do big data and data analytics play in hospitals?
4. Which aspects should be considered for operating room management?
5. How can machine learning contribute to operating room scheduling at the tactical level?

6. How can machine learning contribute to surgery scheduling at the operational level?

7. Can machine learning contribute to other health care areas besides hospitals?

In this Section 1, we have stated the motivation for this work, explained the key concepts of machine learning, and presented vivid applications of machine learning in health care. The remainder of this work is organized as follows. In Section 2, we summarize each of the four contributions contained in this thesis. Section 3 discusses each of the seven aforementioned research questions and links them to the presented contributions. Eventually, Section 4 concludes this work and points out possible directions for further research. All contributions can be found in their full version in the appendix of this thesis.

2 Summaries of the contributions

This thesis makes several contributions to the current literature highlighting how machine learning can support decision making in health care. All contributions are summarized in this section and can be found in their entirety in the appendix. The order of the contributions does not follow the chronological order of submission to scientific journals. In fact, we start with the overarching conceptional framework for the hospital of the future that also proves useful to categorize the remaining three contributions. The two subsequent contributions focus on selected processes for operating room steering and resources management at the tactical and operational level, respectively. Finally, we show that decision support extends beyond the scope of operating rooms by presenting our contribution about automated monitoring of airborne pollen for personalized allergy management.

2.1 A Framework for the Hospital of the Future

Schiele *et al.* [45] develop a vision for the hospital of the future. Based on scientific findings as well as insights from an extensive survey with hospital experts, they present a conceptual framework for the hospital of the future in 2040 structured in 32 dimensions along seven areas and based on four enablers as foundation. This contribution has been submitted to "Journal of Business Economics", which is ranked in category B in the VHB-JOURQUAL3 ranking [2]. It can be found in its entirety in Appendix A. Moreover, selected excerpts of the contribution have been presented at the 2019 autumn conference of the Bundesverband KH-IT e.V. in Erlangen, Germany, on September 19, 2019, and at the interdisciplinary research seminar in Augsburg, Germany, on November 12, 2019, as well as published in "Krankenhaus-IT Journal", issue 6, 2019.

Today's health care systems around the globe face social, organizational, medical, and financial challenges and must constantly adapt to a changing environment. Demographic change, increasing cost pressure, lack of funding, time-consuming documentation, legal regulations imposed by authorities, data security, and acute shortage of staff in hospitals are just some of the key challenges. However, inspired by the fourth industrial revolution, some of the aforementioned challenges might be mitigated by technical, social, and organizational innovations. The fact that big data and machine learning offer unprecedented analytical power, wearables become omnipresent, and digital health attracts increasing investments indicate that we are on the cusp of a new era in health care. This concept is called *Health Care 4.0* and has been studied extensively in the literature. Schiele *et al.* derive several implications for the four traditional health care sectors, i.e., providers, payers, medical technology companies, and biopharma companies, as well as for patients and technology companies. Since literature dedicated to providers is rare, they focus on hospitals and develop a conceptual framework for the hospital of the future.

According to Schiele *et al.*, the hospital of the future in 2040 will be characterized by digitization, integration, automation, and personalization in both support processes as well as core processes using cyber-physical systems and data analytics. There will be more integration within the hospital and with other stakeholders, less paper-based work and less bureaucracy, efficient and optimized processes, more automation and data-driven solutions, transparency and sustainability, more time spent for value-adding tasks and with the patient, higher quality and safety, and after all better medicine and better treatment leading to higher satisfaction of staff and patients. However, they also state that the human component will remain essential in the future. Not few hospitals have already started to discuss and implement selected use cases such as electronic health records or surgery robots. Even though these individual lighthouses might be innovative and beneficial, most hospitals still lack a holistic concept that guides them on the way towards the hospital of the future.

This contribution proposes a conceptual framework for the hospital of the future that orchestrates the individual use cases in 32 dimensions along seven areas. This so-called *HoF framework* reflects both latest scientific findings as well as hands-on insights experienced in practice. The authors conducted an extensive survey with more than 265 hospital experts such as managers, physicians, nurses, and IT professionals. The survey reveals that hospital managers have high ambitions to enhance their hospitals, but are struggling to build momentum. Most of the seven key benefits were evaluated by the participants as relevant and with high potential. To make the hospital of the future more tangible, the authors develop a target picture that illustrates how they envision the hospital of the future. In fact, they distill seven key *HoF areas*: 'Patient', 'Staff', 'Treatment and intervention', 'Logistics and supply', 'Management and organization', 'Data and control', and 'Infrastructure'. Each area is divided into several dimension that are described in detail and enhanced with vivid examples from scientific literature and practice. Furthermore, they asked the survey participants to evaluate those dimensions by importance and maturity level (on a scale from low (0) to high (4), respectively) leading to one of four recommendations: 'Deprioritize', 'Monitor for changes', 'Need for action', and 'Become master'. The survey results reveal need for action in almost all dimensions. Eventually, they also distill the four most critical enablers that are indispensable prerequisites to become a hospital of the future: 'Employees & skills', 'IT infrastructure & data security', 'Strategy & roadmap', and 'People engagement & governance'. The authors describe all four enablers in detail and present evaluation results obtained from the survey.

Given the pressing challenges and the fast-changing environment, the authors emphasize that hospitals need to move quickly to address all four key enablers, develop a long-term vision, and conduct many pilots. This will help hospitals to familiarize with the new concept, initiate a cultural change, attract qualified employees, build up strong partnerships, and gain first hands-on experiences. In order to support practitioners who want to apply the framework, the authors propose a four-step approach. In fact, they suggest to assess the overall maturity level of a hospital and illustrate it by indicating the score for each dimension in a radar chart. The derived radar chart might be useful to evaluate gaps between the current and the target level and to compare the maturity level with other hospitals.

In conclusion, the proposed high-level target picture, the structured framework, and the four-step approach will facilitate strategic discussion and serve as valuable guidance for practitioners to navigate the multitude of use cases in a structured manner. Moreover, the framework enables researchers to categorize their work into one of the 32 dimensions, develop structured literature overviews, and uncover research gaps. The work opens up several options for future research such as field studies with selected hospitals, a global survey, and development of a benchmarking database.

2.2 Predicting ICU Bed Occupancy for Integrated Operating Room Scheduling via Neural Networks

Schiele *et al.* [47] study the implications of tactical operating room scheduling on the downstream units in a hospital. The work at hand can be categorized into the dimension 'Operating room steering and resources management (T1)' belonging to the area 'Treatment and intervention' of the HoF framework (see Section 2.1). They propose a neural network based approach to predict the ICU bed occupancy level resulting from a given master surgery schedule (MSS). This contribution has been accepted for publication in "Naval Research Logistics", which is ranked in category B in the VHB-JOURQUAL3 ranking [2]. It can be found in its entirety in Appendix B. Moreover, selected excerpts of the contribution have been presented at the 2019 Graduate Program in Operations Management (GPOM) in Munich, Germany, on January 11, 2019, and at the 2019 Wissenschaftstag at Universitätsklinikum Augsburg, Germany, on November 14, 2019, as well as published in "Kongresszeitung 136. Kongress der Deutschen Gesellschaft für Chirurgie (DGCH)", issue 1, 2019.

The ICU is one of a hospital's most expensive resources representing nearly 15% of the United States' total hospital expenditures [27] and an important bottleneck

[34] bearing the risk of blocked operating rooms and inferior patient treatment. Consequently, also the implications on the ICU should be considered when scheduling operating rooms. Most hospitals are challenged by ICU capacity shortages causing overtime costs, unsatisfied staff and patients, and postponed and cancelled surgeries. It is well known that allocating operating room capacity to different medical specialties has an impact on the resulting occupancy levels of downstream units. Hence, the hospital management requires a supporting tool in order to address the ICU shortages already at the tactical level.

At the tactical level, operating room managers are asked to develop a MSS by assigning a hospital's operating room capacity to different medical specialties. The suitability of block scheduling has been analyzed by Van Oostrum *et al.* [51]. Usually, a MSS is constructed cyclical, i.e., repeating after a fixed cycle. Most MSS approaches in the literature focus only on the operating room, but neglect or simplify the impact on downstream units. However, an integrated approach that incorporates downstream units seems more suitable to improve their combined performance. In order to consider the impact on supporting units, hospital managers require a model that predicts the resulting bed occupancy levels for a given MSS. The state-of-the-art approach by Fügener *et al.* [23, 24] considers multiple downstream units and serves as comparison for the study at hand.

In this contribution, the authors present the first prediction model for the integrated operating room scheduling problem that is based on machine learning. The proposed prediction model leverages a large data set with $m \in \mathcal{M}$ samples $\left(\boldsymbol{x}^{(m)}, y^{(m)}\right)$. Instead of explicitly modeling the rather complex relationship between inputs $\boldsymbol{x}_n, n \in \mathcal{N}$, i.e., a given MSS, and the corresponding labels $\boldsymbol{y}$, i.e., number of occupied ICU beds on day m, their proposed neural network learns automatically from the historical data. Unlike previous work, this method enables the authors to reflect all patient paths that have occurred in the past without any simplifications. In fact, they reconstruct the exact location of each patient within the hospital for any given date from admission to discharge. The bed occupancy levels in hospital departments such as ward, ICU, intermediate care unit, emergency room, and operating room are obtained by superimposing all patients in the respective location. The authors use multiple input features, i.e., number of operating room blocks allocated to specialty $j \in \mathcal{J}$ on day $m + d, \forall m \in \mathcal{M}, d \in \mathcal{D}$, and one-hot encoded days of the week $e \in \mathcal{E}$. For each room that has been occupied by a medical specialty on a given day, one operating room block is assigned to the respective specialty. In case of shared operating rooms, the operating room block is allocated to the involved medical specialties according to their accumulated surgery durations. The authors introduce a new parameter called memory depth $d \in \mathcal{D}$ to consider operating room blocks of previous days and subsequent days as well. The objective of the prediction model is to minimize the deviation between the predicted and the realized ICU bed occupancy.

Schiele *et al.* evaluate the model with real-world data retrieved from Universitätsklinikum Augsburg, a 1,700-bed, maximum-care university hospital located in Southern Germany. In fact, surgery records as well as supporting unit records from 2010 to 2016 were retrieved from the hospital information system to derive more than $M = 2{,}500$ samples covering more than $P = 77{,}000$ patients, $F = 1{,}017$ distinct patient paths, and $J = 8$ medical specialties, i.e., Cardiothoracic Surgery, General, Visceral, and Transplant Surgery, Gynecology, Oral and Maxillofacial Surgery, Neurosurgery, Traumatology, Orthopedics, and Plastic Surgery, Urology, and Vascular and Intravascular Surgery. The used data sets are commonly available in most hospitals and usually can be interlinked by an unique patient identifier which is essential to reconstruct the patient paths. The authors propose a 3-step approach to structure pre-processing and training. In step 1, the patient group in focus is selected, e.g., elective patients, outpatients, and selected medical specialties. In step 2, pre-processing is performed to derive features and labels for various memory depths. In step 3, model hyperparameters are selected for the training, e.g., topology, activation function, and optimizer. The proposed model has been implemented in Python. For data processing, `Pandas`, `NumPy`, `SciPy`, `scikit-learn`, and `Tensorflow` are used. Pre-processing and training were performed on a dedicated simulation node equipped with 56 physical Intel(R) Xeon(R) Platinum 8176 cores with enabled hyperthreading.

After discussions with the hospital management, the root mean squared error is used to evaluate the performance since robustness of the model has highest priority. Schiele *et al.* present numerical results for various combinations of patient selection (step 1), data set selection (step 2), and model selection (step 3). The model achieves a root mean squared error of 3.46 for the base case, i.e., for $P = 77k$ patients, memory depth $\mathcal{D} = \{-20, 10\}$, and [200:50] topology. In order to better contrast the performance of the proposed model to the literature, the authors present two types of comparison. First, they keep the pre-processing (steps 1 and 2) as proposed and only vary the machine learning algorithm (step 3). They conclude that most of the studied alternatives are also well-suited for the prediction problem at hand, however, do not outperform the proposed neural network. Second, they compare the entire model with the state-of-the-art model by Fügener *et al.* [23] and conclude that their model achieves an prediction error that is 43% lower in number of predicted beds.

Among other possible applications, the proposed prediction model will be most beneficial for the evaluation of a given MSS. Since the benefits of a new prediction model comes not only from the improvement in prediction quality but rather from the improvement in the quality of the decision that the model informs, the authors present two options to incorporate it into a decision making process in order to inform a better decision. First, the proposed prediction model serves as a valuable tool in regular management discussions to adaptively evaluate a given, feasible MSS with respect to the expected ICU bed occupancy levels. Second, the prediction model can be incorporated as an objective function in an optimization model such as a genetic algorithm. The authors compare the current MSS A that is based on status quo with MSS B that is based on discussions with the hospital management and MSS C that is based on genetic algorithm optimization. The bed occupancy for MSS C shows a better leveling compared to the other ones leading to an estimated reduction of the maximum ICU bed demand by 8.9%.

In summary, the model can be used as a supporting tool for hospital managers or incorporated in an optimization model in order to make operating room scheduling more efficient in the future. Machine learning is well suited for this problem since traditional models struggle to reflect the hospital's real-world complexity and its inherent uncertainty. The work opens up several directions for further research, e.g., enrichment with additional features such as emergency status.

2.3 Predicting Surgical Durations and Implications at the Operational Level

Schiele [44] investigates the consequences of surgery scheduling at the operational level of a hospital. The work at hand can be categorized into the dimension ‘Operating room steering and resources management (T1)’ belonging to the area ‘Treatment and intervention’ of the HoF framework (see Section 2.1). He proposes a multi-objective approach based on neural networks to predict the duration as well as several operational implications of a surgery. He shows that a variety of parameters can be retrieved from commonly available hospital records and incorporated into the prediction model. Numerical results are presented for a case study with data from Universitätsklinikum Augsburg. This contribution has been submitted to “Medical Decision Making”, which is ranked in category A in the VHB-JOURQUAL3 ranking [2]. It can be found in its entirety in Appendix C. Moreover, selected excerpts of the contribution have been presented at the 30th European Conference on Operational Research (EURO) conference in Dublin, Ireland, on June 25, 2019.

The operating room is the most critical resource and a major cost driver of a hospital. In fact, operating room expenditures account for around 40% of the hospital expenditures and an average operating room in the United States costs roughly 4,000 USD per hour. Consequently, it is of utmost importance to ensure an efficient management of the operating room environment. In order to schedule surgeries efficiently, accurate predictions of surgical durations are essential to avoid overestimating as well as underestimating both of which bear undesirable consequences. In today’s hospitals, predictions are prone to inaccuracy being based on either expert estimates or simple historic averaging. Within the last decades, a vast amount of research has been done. The author identifies 114 papers in this field and distinguishes between literature reviews, probability distribution fitting, expert prediction assessments, data mining and machine learning models, scheduling and sequencing, and others. In particular, machine learning based approaches seem promising and well suited for the prediction of surgical durations by learning the relationship between features and outputs from historic data.

This contribution proposes a new multi-objective approach based on neural networks. The model leverages a variety of patient-related, procedure-related, and operations-related features to predict multi-class labels for six different metrics with great practical relevance, i.e., operating room duration, incision suture duration, anesthetist duration, postoperative unit, postoperative length of stay, and discharge type. In fact, the model focuses not only on durations within the operating room, but also considers the influence of a scheduled surgical case on the further course of a patient's postoperative stay in the hospital as well as post-hospital treatments required both of which contributing to the efficiency of a hospital. Unlike other approaches, Schiele also considers features and labels that were derived from individual patient paths, e.g., the preoperative length of stay. The proposed model is generalizable and based on input data that is commonly available in most hospitals.

The proposed prediction model is based on neural networks and a large data set with $m \in \mathcal{M}$ historic samples $\left(\boldsymbol{x}^{(m)}, \boldsymbol{y}^{(m)}\right)$, i.e., surgical cases, where $\boldsymbol{x}_n, n \in \mathcal{N}$, describes the features and $\boldsymbol{y}_o$ the corresponding multi-class labels for $o \in \mathcal{O}$ outputs. Schiele evaluates the model with real-world data retrieved from Universitätsklinikum Augsburg, a 1,700-bed, maximum-care university hospital located in Southern Germany. In fact, surgery records as well as supporting unit records from 2010 to 2016 were retrieved from the hospital information system to derive more than $M = 150{,}000$ samples covering more than $P = 125{,}000$ distinct patients. After deliberation with the hospital management, the author uses patient-related features such as patient age, gender, patient type, emergency status, and admission type, procedure-related features such as International Classification of Procedures in Medicine code, operating room type, and surgical specialty, and operations-related features such as the number of previous surgeries, the preoperative length of stay in the hospital, the origin unit before being transferred to the operating room, the weekday, and the time of the day. He distinguishes between continuous and categorical features and presents their representations for the training samples. Categorical variables are converted into integers by using one-hot encoding. A multi-class label $\boldsymbol{y}_o^{(m)}$ is defined for each output $o \in \mathcal{O}$ with values $y_{o,c}^{(m)} \in [0,1]$ for $c \in \mathcal{C}_o$ classes where quantiles are used to categorize the surgical durations. The model has been implemented with `scikit-learn` in Python [40] and computations were done on a dedicated simulation node equipped with 56 physical Intel(R) Xeon(R) Platinum 8176 cores with enabled hyperthreading.

In order to evaluate the performance of the model, unweighted average precision (UAP), unweighted average recall (UAR), and unweighted average F-measure (UAF) are used. The author runs various computations with different hyperparameters and presents results for the models achieving the best performance. In particular, an UAP= 59.7%, an UAR= 58.7%, and an UAF= 59.1% are achieved on a neural network model with topology [10:10] for the prediction of the operating room duration and an UAP= 97.2%, an UAR= 97.4%, and an UAF= 97.2% are achieved on a model with topology [100:50:25] for the prediction of the postoperative unit. Moreover, confusion matrices are presented to illustrate the performance for each individual class of the multi-class output.

In summary, the proposed model seems to be of great practical value and could support clinicians in surgery scheduling. An aggregated version of the unweighted average performance metrics can be used to identify the model constellations with the best predictive power. The work opens up five possible directions for further research, i.e., additional features, customized performance metrics, comparision with other models, incorporation into an optimization framework, and application in the field.

2.4 Automated Classification of Airborne Pollen using Neural Networks

Demonstrating that AI-enabled decision support in health care extends beyond the scope of operating rooms, Schiele *et al.* [46] explore the automated monitoring of airborne pollen with the aim of improving personalized allergy management. The work at hand can be categorized into the dimension 'Smart devices and medication (P3)' belonging to the area 'Patient' of the HoF framework (see Section 2.1). They propose a new pollen classification model based on deep neural networks, present numerical results, and compare the performance with a state-of-the-art algorithm. This contribution has been published in "41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)", which is not ranked in the VHB-JOURQUAL3 ranking [2]. It can be found in its entirety in Appendix D. Moreover, selected excerpts of the contribution have been presented at the EMBC in Berlin, Germany, on July 26, 2019.

Clinical evidence reveals a general increase in both the incidence and prevalence of respiratory diseases over the last half century [7] which presumably is related to a parallel increase in the amount of airborne allergenic pollen [54]. Not only patients suffer from allergies, but also economic effects are noticeable, e.g., costs for visits to allergy specialists, lost working hours, and reduced working efficiency of patients. With prophylaxis being the first line of allergy management, accurate information on the occurrence and abundance of airborne pollen is needed in order to diminish exposure. Up to date, the volumetric Hirst-type sampler [8] is the biomonitoring gold-standard despite its fully manual operation from collection to chemical analysis and microscopic classification of different types of pollen [22]. Consequently, relevant risk alerts are announced to the public usually with a delay of at least 7 to 10 days.

Schiele *et al.* present a new automated pollen classification approach based on machine learning. The authors describe a 4-step automated pollen monitoring process, i.e., automated collection of air samples on a probe (step 1), detection of objects within each probe (step 2), classification of objects as either 'no pollen' or as a specific pollen type (step 3), and online publication of up-to-date information about the pollen concentration (step 4). For the collection of samples in step 1, airborne allergenic pollen have been monitored in Augsburg, Germany, in 2015 and 2016 using a novel automated pollen measuring device. The Bio-Aerosol Analyzer BAA500[1] continuously samples ambient air using a 3-stage virtual impactor such that pollen is deposited on a sticky surface and digitized with a camera. For the detection in step 2, the authors rely on the internal detection algorithm of the BAA500. The detected pollen samples were then manually labeled by aerobiology experts based on typical morphological features. In the study at hand, the 15 most abundant and allergenic pollen types worldwide were considered [54], i.e., *Alnus*, *Betula*, *Carpinus*, *Corylus*, *Fagus*, *Fraxinus*, *Plantago*, Poaceae, *Populus*, *Quercus*, *Salix*, *Taxus*, *Tilia*, *Ulmus*, and Urticaceae.

For the classification in step 3, the authors propose a model that is based on convolutional neural networks. Consisting of a cascade of so-called convolutional layers and one or several fully-connected neural layers, convolutional neural networks are a state-of-the-art method in image classification tasks such as visual object recognition, optical character recognition, or image based medical diagnostics [18]. In a pre-processing step, each pollen image is embedded in the center of a black background frame and the data set is split into three sets: the training set containing 60 % of all samples of each class, the validation set with 20 %, and the test set with 20 %. Raw pixel values of the embedded images are used to train the network consisting of three convolutional layers, each followed by a maximum-pooling step, a fully-connected layer with dropout, and an output layer. The described model has been implemented with the `Tensorflow` library in Python and the training was performed on a dedicated GPU cluster equipped with 16 Nvidia Titan X (Pascal) cards.

Using the aforementioned performance metrics (see Section 2.3), the model achieves an UAP of 83.0%, an UAR of 77.1%, and an UAF of 79.1% on the test set. A look at the confusion matrix reveals that *Taxus* and *Fagus* are predicted particularly well with a recall of 96 % and 92 %, respectively, and *Quercus* and *Fraxinus* show a

[1] `https://www.hund.de/en/instruments/pollen-monitor/`

rather bad performance with 35 % and 43 %, respectively. The authors also compare their results with the BAA500 internal classification algorithm which is based on a mathematical model calculating features such as area, perimeter, eccentricity, and roundness for each pollen sample. Overall, their model is able to predict less classes, but shows superior performance than the BAA500 model.

In summary, the proposed approach seems promising to advance towards an accurate, real-time, automatic dissemination of allergenic pollen information. The deep neural network based model is able to classify different types of airborne pollen with high accuracy and without requiring the objects to be round. The work opens up several possibilities for further research such as expansion to more pollen classes, development of an object detection algorithm (step 2), and expansion to automated monitoring of other air particles, e. g., fungal spores. The availability of real-time and accurate pollen information will contribute significantly to improve personalized allergy management.

3 Discussion of the contributions

This thesis is dedicated to AI-enabled decision support in health care and comprises the four aforementioned contributions of which each one fills a research gap in existing literature. The section at hand discusses how the contributions provide answers to the following research questions.

1. Can ideas from Industry 4.0 help mitigate the imminent challenges in health care?

2. How can hospitals navigate the variety of use cases on the way to the hospital of the future?

3. What role do big data and data analytics play in hospitals?

4. Which aspects should be considered for operating room management?

5. How can machine learning contribute to operating room scheduling at the tactical level?

6. How can machine learning contribute to surgery scheduling at the operational level?

7. Can machine learning contribute to other health care areas besides hospitals?

In the following, each research question is discussed in a separate section.

3.1 Can ideas from Industry 4.0 help mitigate the imminent challenges in health care?

The overall health care system and in particular hospitals around the globe face various social, organizational, medical, and financial challenges and must constantly adapt to a changing environment. In particular, the increasing cost pressure, the complex regulatory environment, and the acute shortage of staff contribute to this challenging setting and make it difficult to focus on value-adding tasks. However, instead of cutting costs and reducing value-added care, it is more promising to use concepts similar to Industry 4.0 in order to mitigate those challenges. Digitization, integration, automation, and personalization using cyber-physical systems and data analytics will help to advance the health care system. This has implications for the four traditional health care sectors, i.e., providers, payers, medical technology companies, and biopharma companies, as well as for patients and technology companies.

In the hospital of the future, those innovations will lead to more integration, less paper-based work, less bureaucracy, efficient processes, more automation, data-driven solutions, transparency, sustainability, and after all to more time for value-adding tasks, better medicine, higher satisfaction of staff, and better health of patients. However, it is also important to consider the differences between the smart factory and the hospital of the future. In a hospital, the core process is not production but diagnosis and therapy and the core are not products but humans. In fact, human-human interactions are essential for the recovery process of patients and will remain a valuable component in the future. Taking these differences into account, the concept of Industry 4.0 can well be transferred to the hospital setting serving as assistance for support processes such as cleaning and catering as well as core processes such as diagnostics and surgery. The concept is studied in the first contribution in this thesis [45] (see Section 2.1).

3.2 How can hospitals navigate the variety of use cases on the way to the hospital of the future?

Mainly driven by big data and analytical power, health care is on the cusp of a new era paving the way for the hospital of the future. Not few hospitals have already started to discuss and implement first use cases of this new concept, e.g., electronic health records or surgery robots, and have high ambitions for the upcoming years. Even though these individual lighthouses might be innovative and beneficial, most hospitals still lack a holistic strategy that guides them on the way to become a hospital of the future.

The first contribution in this thesis [45] investigates this problem (see Section 2.1). Based on scientific literature and an extensive survey among hospital practitioners, Schiele *et al.* derive a target picture that makes the hospital of the future in 2040 more tangible and facilitates strategic discussions. In order to support hospitals to navigate the variety of use cases, they propose a structured *HoF framework* that orchestrates the individual use cases in 32 *HoF dimensions* along seven *HoF areas.* Moreover, they present an evaluation matrix for prioritization and a four-step approach including a radar chart as supporting tool for self-assessment and benchmarking.

3.3 What role do big data and data analytics play in hospitals?

Data is essential for nearly all processes in a hospital such as diagnosis, treatment, prevention, hospital operations, and documentation. Hospitals have no shortage of data given the almost indefinite variety of sources such as electronic health records, lab results, magnetic resonance imaging, x-ray scans, medication plans, monitoring records, medical letters, clinical trials, insurance data, and billing information. In the near future, this will be further augmented by the integration of consumer wearables such as fitness bracelets, scales, glucometers, and smartphone applications. In order to leverage the full potential, it is essential to aggregate and manage all data in a centralized hospital information system. The availability of structured data pave the way for advanced methods such as big data analytics, machine learning, pattern recognition, simulation, and optimization offering a previously unimaginable variety of applications, e.g., advanced diagnostics, personalized medicine, post-care monitoring, predictive maintenance, detection of anomalies, paperless documentation, scheduling of resources, operational improvements, as well as prediction of bed occupancy levels, surgical durations, and clinical outcomes. The increasing number of scientific contributions that successfully apply machine learning in health care indicates that the approach is promising. The first contribution in this thesis [45] explores how big data and data analytics can advance hospitals and present numerous vivid applications (see Section 2.1).

Although levels of detail, systems, and tools vary significantly between hospitals, most keep track of the conducted surgeries and the patient movements:

Surgery records. A data set covering the conducted surgeries that often contains patient identifiers, patient types, emergency status, medical specialties, and timestamps for surgery-related events.

Supporting units records. A data set covering admission, transfers, and discharge during the hospital stay that often contains patient identifiers, room numbers, and timestamps for movement-related events.

Understanding individual patient paths through the hospital is beneficial to improve decision making at the tactical and operational level. Given an unique patient identifier, it is possible to interlink the two aforementioned data sources and reconstruct the respective paths for each patient. In fact, the second [47] and the third [44] contribution in this thesis leverage the reconstructed patient paths to improve decision making at the tactical and operational level, respectively (see Sections 2.2 and 2.3).

3.4 Which aspects should be considered for operating room management?

The operating room represents the core of a hospital. In the past six decades, extensive research has been carried out to optimize operating room management [10, 26, 42]. Based on the identified challenges and pitfalls, four critical aspects can be distilled. Figure 3.1 depicts the resulting *IUCT classification scheme* that summarizes all four aspects that should be considered for operating room management.

Integrity	Uncertainty
Complexity	Tangibility

Figure 3.1: IUCT classification scheme: four aspects should be considered for operating room management.

Integrity. Operating room decisions directly affect the connected up- and downstream units of a hospital such as ward and ICU. Particularly the ICU is crucial from a financial [27], medical [5], and organizational perspective [34] and hence, should to be included in the decision making process to improve the overall efficiency.

Uncertainty. Uncertainty is inherent to health services. Patient arrivals, length of stay as well as number, frequency, type, and outcome of interventions are not known in advance, but directly affect demand for resources and workload. In order to make decisions on a realistic basis, uncertainty can not be ignored.

Complexity. Real-world hospital settings are rather complex being shaped by a variety in many dimensions such as patient types, medical specialties, supporting units, and patient paths. Hence, conventional models struggle to explicitly imitate the real-world complexity in all details. Until today, there is no model that reflects all possible patient paths through the hospital without simplifications, e.g., Fügener et al. [23] neglect preoperative stays, multiple surgeries, and transfers from ward to the ICU.

Tangibility. Hospitals are risk-adverse. As patient safety is of utmost importance and capacity is expensive, any modifications to crucial processes require robust reasoning to avoid potentially lethal disruptions. In order to achieve lasting improvements in real-world hospitals, it is mandatory to build upon real data and to ensure that the hospital stakeholders understand the resulting benefits.

The second [47] and the third [44] contribution in this thesis propose data-driven models for the operating room management at the tactical and operational level, respectively, that both satisfy the IUCT classification scheme: considering operating rooms as well as supporting units, handling uncertainty and complexity reflected in seven years of underlying real data, and closely aligned with key stakeholders of the reference hospital Universitätsklinikum Augsburg (see Sections 2.2 and 2.3).

3.5 How can machine learning contribute to operating room scheduling at the tactical level?

In operating room scheduling at the tactical level, a MSS is developed to assign a hospital's operating room capacity to the different medical specialties. This decision does not only affect the operating rooms, but also has implications on the resulting bed occupancy levels in downstream units such as the ICU. The ICU is one of a hospital's most expensive resources and often represents an important bottleneck bearing the risk of blocked operating rooms and inferior patient treatment, i.e., lower probability of ICU admission, higher discharge rates, and increased danger of re-admission. Consequently, also the ICU should be considered when scheduling operating rooms (see IUCT scheme in Section 3.4). To do so, a supporting tool is required that predicts the impact of a MSS on downstream units.

Within the last decade, machine learning has gained momentum and helped to advance applications in health care, particularly for diagnosis such as classification of skin cancer [18]. Moreover, it is well-suited for forecasting and prediction problems in several industries, e.g., surgical durations [35]. Consequently, it seems beneficial and worth approaching to develop the supporting tool based on machine learning. The second contribution in this thesis [47] investigates the integrated operating room scheduling problem and proposes a neural network based supporting tool that predicts the resulting ICU bed occupancy levels for a given MSS (see Section 2.2). While traditional model struggle to reflect the hospital's real-world complexity and uncertainty, machine learning is well suited for this problem being able to learn the relationship between given MSS and resulting bed occupancy level directly from historical data. The approach is able to reflect more supporting units, patient types, and patient paths than previous work and the achieved predicting results outperform a state-of-the-art model. The proposed model serves as valuable supporting tool for hospital managers and can be incorporated in an optimization model.

3.6 How can machine learning contribute to surgery scheduling at the operational level?

In surgery scheduling at the operational level, clinicians consider several aspects such as estimated utilization of operating rooms, under- and overtime of staff, and the impact on downstream units. For an efficient operating room management, accurate predictions of surgical durations are essential since both overestimating as well as underestimating surgical durations bear undesirable consequences, i.e., idle time, lost revenues, waiting times, and cancellations. However, deriving accurate predictions is a complex endeavor due to inherent uncertainty and diversity. In most hospitals, the predictions are based on either expert estimates or simple averaging methods making them prone to inaccuracy. Besides surgical durations, also accurate predictions of operational implications such as postoperative unit, postoperative length of stay, and discharge type are important, but are rarely considered in today's hospitals.

Within the last decades, a vast amount of research has been done in the field of surgical scheduling [10], however, the problem of predicting surgical durations comprises only a minor fraction of the literature and just recently started to get more attention. Machine learning is well suited for the prediction of surgical durations by learning the relationship between features and outputs from historical data. The third contribution [44] investigates the problem at hand (see Section 2.3). Schiele presents a new multi-objective approach based on neural networks that is able to predict various perioperative durations as well as operational implications with great practical relevance. The model leverages a variety of patient-related, procedure-related, and operations-related factors retrieved from 7 years of real-world data covering more than 150,000 surgeries at Universitätsklinikum Augsburg. Convincing numerical results with high precision and recall values indicate that the approach is valid and could prove useful to support clinicians in their decision making.

3.7 Can machine learning contribute to other health care areas besides hospitals?

As shown in the second [47] and the third [44] contribution in this thesis, machine learning is well suited to improve decision making at a hospital's tactical and operational level, respectively (see Sections 2.2 and 2.3). Moreover, numerous studies and applications demonstrate the potential of machine learning in hospitals, e.g., to classify skin cancer [18], to predict pneumonia [28], to assess perioperative cardiac risks [30], to identify patients at high risk of postinduction hypotension [33], to support robot-assisted surgeries [52], to predict surgical durations [20] and other clinical events [17].

However, the potential of AI-enabled decision support in health care extends far beyond hospitals. The first contribution [45] in this thesis describes applications for each of the four traditional health care sectors, i.e., providers, payers, medical technology companies, and biopharma companies, as well as for patients and technology companies (see Sections 2.1). For example, machine learning supports biopharma companies to discover promising drugs by leveraging clinical data and partnering with biobanks [3, 53]. Providers use machine learning to automate claims handling and detect fraud, waste, and abuse. Medical technology companies improve the efficiency of their operations by identifying bottlenecks, preventing stock-outs on critical products, and using predictive maintenance. Also, technology companies become increasingly attracted by the new opportunities offered by machine learning in health care, e.g., to discover drugs (Atomwise), to connect patients with general practitioners (Babylon Health), and to analyze medical imaging (Zebra) [3].

Demonstrating the potential of AI outside of hospitals, the fourth contribution in this thesis [46] explores how machine learning can contribute to personalized allergy management (see Section 2.4). Pollen allergies are considered as a global epidemic nowadays, but today's alerts on high-risk allergenic pollen exposure are still based on conventional monitoring methods that are laborious and delayed by at least 7 to 10 days. Schiele *et. al* propose a fully automated approach based on convolutional neural networks to classify airborne pollen. Airborne allergenic pollen have been monitored in Augsburg, Germany, using a novel automatic Bio-Aerosol Analyzer (BAA 500, Hund GmbH). The proposed model achieves convincing results, i.e., an UAP of 83.0% and an UAR of 77.1% across 15 classes of pollen taxa. Automatic, real-time information on concentrations of airborne allergenic pollen will significantly contribute to the implementation of timely, personalized management of allergies in the future.

4 Conclusion

This thesis explores several aspects of AI-enabled decision support in health care. It motivates the need for further research on the intersection between health care and machine learning. The main part of this thesis consists of four contributions to the literature. A summary of each of the contributions is provided and the contributions are applied to answer seven open research questions introduced in this thesis.

The first contribution [45] introduces a vision for the hospital of the future in 2040 (see Section 2.1). Based on scientific literature and an extensive survey, a high-level target picture and a conceptual framework structured in 32 dimensions along seven areas is proposed in order to facilitate strategic discussions and navigate the multitude of use cases in a structured manner. A tangible four-step approach underlines the practical relevance and enables self assessment. The second contribution [47] studies the implications of MSSs on the downstream units at the tactical level of a hospital (see Section 2.2). This is the first prediction model for the integrated operating room scheduling problem that is based on neural networks. A case study underlines its effectiveness and presents prediction results that outperform state-of-the-art by 43%. The proposed model serves as a valuable supporting tool in regular management discussions and can be incorporated as an objective function in an optimization model. In the case study, the expected maximum ICU bed demand was reduced by 8.9%. The third contribution [44] examines the consequences of surgery scheduling at the operational level of a hospital (see Section 2.3). This is the first multi-objective classification model based on neural networks that is able to predict multi-class labels for six different metrics with great practical relevance. A case study demonstrates convincing prediction results, i.e., UAFs of at least 55.4% up to 97.2%. Finally, the fourth contribution [46] explores automated monitoring of airborne pollen with the aim of improving personalized allergy management (see Section 2.4). It introduces a 4-step automated pollen monitoring process and proposes a new pollen classification model based on deep neural networks. In a case study with 15 pollen types, the model achieves convincing prediction results, i.e., an UAF of 79.1%, showing superior performance than the state-of-the-art model.

The thesis poses several research questions which have not been answered by existing literature. The first contribution [45] shows how the concept of Industry 4.0 can help to mitigate the imminent challenges in health care (see Section 3.1). Digitization, integration, automation, and personalization using cyber-physical systems and data analytics have implications for the four traditional health care sectors as well as for patients and tech companies. In hospitals, the concept serves as assistance for support processes such as cleaning and catering as well as core processes such as diagnostics and surgery. In order to support hospitals to navigate the variety of use cases, the first contribution [45] proposes a structured HoF framework that orchestrates the individual use cases in 32 HoF dimensions along seven HoF areas (see Section 3.2). Big data and data analytics play an essential role in hospitals (see Section 3.3). The first contribution [45] explores how big data and data analytics can advance hospitals and present numerous vivid applications. The second [47] and third [44] contribution leverage two data sources of a hospital and reconstruct patient paths in order to improve decision making. A new classification scheme is presented summarizing four aspects that should be considered for operating room management (see Section 3.4). Two contributions [47, 44] address the operating room management and propose data-driven models that both satisfy the classification scheme. Machine learning is well suited to improve decision making at a hospital's tactical level (see Section 3.5). The second contribution in this thesis [47] investigates the integrated operating room scheduling problem and proposes a neural network based supporting tool that predicts the resulting ICU bed occupancy levels for a given MSS. At the operational level, machine learning can contribute to surgery scheduling by predicting implications of surgeries (see Section 3.6). The third contribution [44] presents a new multi-objective approach based on neural networks that is able to predict various perioperative durations as well as operational implications with great practical relevance. AI-enabled decision support has great potential for hospitals, but also for other health care areas (see Section 3.7). Some applications are described in the first contribution [45]. The fourth contribution [46] proposes a fully automated approach based on convolutional neural networks to classify airborne pollen.

This thesis opens up several opportunities for further research. While all contributions present relevant applications that support decision making in health care, it is worthwhile to apply and test them in practice. Field studies with selected hospitals on a global scale will be beneficial to further underline the relevance and improve the predictive performance. We intend to set up a global benchmarking database for the HoF framework that provides learnings from best-of-class hospitals. While the proposed neural network based models achieve convincing results, it is worthwhile to study the integration of additional features, evaluation with customized performance metrics, and comparison with alternative machine learning models. Ultimately, we believe that a joint consideration of tactical and operational decision levels would be beneficial. In particular, to predict the ICU occupancy level, one might consider patient-specific features or reflect the operating room occupancy on an hourly basis. Customized loss functions could be beneficial to penalize high congestion periods more severely. Metaheuristic or greedy random search could help to identify even better topologies for the neural network. Analogous to the second contribution, the model presented in the third contribution could also be incorporated into an optimization framework. The model in the fourth contribution could be expanded to other air particles and complemented with an object detection algorithm.

Facing many challenges, the health care system is on the cusp of a new era. Now is the right time for a data-driven health care industry. AI-enabled decision support has the potential to improve health care processes and even save lives. It is our intention to encourage further research and practical studies in this area to unfold the full potential and create lasting impact.

Bibliography

[1] CB Insights: Top Healthcare AI Trends To Watch. https://www.cbinsights.com/, 2019.

[2] Verband der Hochschullehrer für Betriebswirtschaft e.V.: Complete list of the journals in VHB-JOURQUAL3 in alphabetical order. https://vhbonline.org/en/service/jourqual/vhb-jourqual-3/complete-list-of-the-journals/, 2019.

[3] A. Aboshiha, R. Gallagher, and L. Gargan. Chasing Value as AI Transforms Health Care. *Boston Consulting Group*, 2019.

[4] M. Awad and R. Khanna. Machine Learning. In M. Awad and R. Khanna, editors, *Efficient Learning Machines: Theories, Concepts, and Applications for Engineers and System Designers*. Apress, Berkeley, CA, 2015. ISBN 978-1-4302-5990-9. doi: 10.1007/978-1-4302-5990-9_1.

[5] D. R. Baker, P. J. Pronovost, L. L. Morlock, R. G. Geocadin, and C. G. Holzmueller. Patient flow variability and unplanned readmissions to an intensive care unit. *Critical Care Medicine*, 37(11):2882–2887, 2009. ISSN 1530-0293. doi: 10.1097/ccm.0b013e3181b01caf.

[6] D. M. Berwick and A. D. Hackbarth. Eliminating waste in US health care. *Journal of the American Medical Association*, 307(14):1513–1516, 2012.

[7] J. Bunne, H. Moberg, L. Hedman, M. Andersson, A. Bjerg, B. Lundbäck, and E. Rönmark. Increase in Allergic Sensitization in Schoolchildren: Two Cohorts Compared 10 Years Apart. *The Journal of Allergy and Clinical Immunology: In Practice*, 5(2):457–463.e1, 2017. ISSN 2213-2198. doi: 10.1016/j.jaip.2016.09.025.

[8] J. T. M. Buters, C. Antunes, A. Galveias, K. C. Bergmann, M. Thibaudon, C. Galán, C. Schmidt-Weber, and J. Oteros. Pollen and spore monitoring in the world. *Clinical and Translational Allergy*, 8, 2018. ISSN 2045-7022. doi: 10.1186/s13601-018-0197-8.

[9] D. Bychkov, N. Linder, R. Turkki, S. Nordling, P. E. Kovanen, C. Verrill, M. Walliander, M. Lundin, C. Haglund, and J. Lundin. Deep learning based tissue analysis predicts outcome in colorectal cancer. *Scientific Reports*, 8(1): 3395, 2018. ISSN 2045-2322. doi: 10.1038/s41598-018-21758-3.

[10] B. Cardoen, E. Demeulemeester, and J. Beliën. Operating room planning and scheduling: A literature review. *European Journal of Operational Research*, 201 (3):921–932, 2010. ISSN 0377-2217. doi: 10.1016/j.ejor.2009.04.011.

[11] M. Carter, E. W. Hans, and R. Kolisch. Health care operations management. *OR Spectrum*, 34(2):315–317, 2012. ISSN 1436-6304. doi: 10.1007/s00291-012-0288-1.

[12] Centers for Medicare & Medicaid Services. National Health Expenditures 2016, 2016.

[13] M. Chen, Y. Hao, K. Hwang, L. Wang, and L. Wang. Disease Prediction by Machine Learning Over Big Data From Healthcare Communities. *IEEE Access*, 5:8869–8879, 2017. ISSN 2169-3536. doi: 10.1109/ACCESS.2017.2694446.

[14] P.-H. C. Chen, K. Gadepalli, R. MacDonald, Y. Liu, S. Kadowaki, K. Nagpal, T. Kohlberger, J. Dean, G. S. Corrado, J. D. Hipp, C. H. Mermel, and M. C. Stumpe. An augmented reality microscope with real-time artificial intelligence

integration for cancer diagnosis. *Nature Medicine*, 25(9):1453–1457, 2019. ISSN 1546-170X. doi: 10.1038/s41591-019-0539-7.

[15] M. M. Churpek, T. C. Yuen, C. Winslow, D. O. Meltzer, M. W. Kattan, and D. P. Edelson. Multicenter Comparison of Machine Learning Methods and Conventional Regression for Predicting Clinical Deterioration on the Wards. *Critical Care Medicine*, 44(2):368–374, 2016. ISSN 1530-0293. doi: 10.1097/CCM.0000000000001571.

[16] S. Daley. Surgical robots, new medicines and better care: 32 examples of AI in healthcare. https://builtin.com/artificial-intelligence/artificial-intelligence-healthcare, 2019.

[17] C. Esteban, O. Staeck, Y. Yang, and V. Tresp. Predicting Clinical Events by Combining Static and Dynamic Information Using Recurrent Neural Networks. *arXiv:1602.02685 [cs]*, 2016.

[18] A. Esteva, B. Kuprel, R. A. Novoa, J. Ko, S. M. Swetter, H. M. Blau, and S. Thrun. Dermatologist-level classification of skin cancer with deep neural networks. *Nature*, 542(7639):115–118, 2017. ISSN 1476-4687. doi: 10.1038/nature21056.

[19] Eurostat. Healthcare expenditure statistics, 2018.

[20] M. Fairley, D. Scheinker, and M. L. Brandeau. Improving the efficiency of the operating room environment with an optimization and machine learning model. *Health Care Management Science*, 22(4):756–767, 2019. ISSN 1572-9389. doi: 10.1007/s10729-018-9457-3.

[21] J. D. Fauw, J. R. Ledsam, B. Romera-Paredes, S. Nikolov, N. Tomasev, S. Blackwell, H. Askham, X. Glorot, B. O'Donoghue, D. Visentin, G. van den Driessche, B. Lakshminarayanan, C. Meyer, F. Mackinder, S. Bouton, K. Ayoub, R. Chopra, D. King, A. Karthikesalingam, C. O. Hughes, R. Raine, J. Hughes, D. A. Sim, C. Egan, A. Tufail, H. Montgomery, D. Hassabis, G. Rees, T. Back, P. T. Khaw,

M. Suleyman, J. Cornebise, P. A. Keane, and O. Ronneberger. Clinically applicable deep learning for diagnosis and referral in retinal disease. *Nature Medicine*, 24(9):1342–1350, 2018. ISSN 1546-170X. doi: 10.1038/s41591-018-0107-6.

[22] B. A. Federation. *Airborne Pollens and Spores: A Guide to Trapping and Counting*. National Pollen and Hayfever Bureau, Rotherdram, United Kingdom, 1995.

[23] A. Fügener, E. W. Hans, R. Kolisch, N. Kortbeek, and P. T. Vanberkel. Master surgery scheduling with consideration of multiple downstream units. *European Journal of Operational Research*, 239(1):227–236, 2014. ISSN 0377-2217. doi: 10.1016/j.ejor.2014.05.009.

[24] A. Fügener, G. M. Edenharter, P. Kiefer, U. Mayr, J. Schiele, F. Steiner, R. Kolisch, and M. Blobner. Improving Intensive Care Unit and Ward Utilization by Adapting Master Surgery Schedules. *A & A Case Reports*, 6(6):172–180, 2016. ISSN 2325-7237. doi: 10.1213/XAA.0000000000000247.

[25] I. Goodfellow, Y. Bengio, and A. Courville. *Deep Learning*. The MIT Press, 2016. ISBN 978-0-262-03561-3.

[26] F. Guerriero and R. Guido. Operational research in the management of the operating theatre: A survey. *Health Care Management Science*, 14(1):89–114, 2011. ISSN 1386-9620. doi: 10.1007/s10729-010-9143-6.

[27] N. A. Halpern and S. M. Pastores. Critical care medicine in the United States 2000-2005: An analysis of bed numbers, occupancy rates, payer mix, and costs. *Critical Care Medicine*, 38(1):65–71, 2010. ISSN 1530-0293. doi: 10.1097/CCM.0b013e3181b090d0.

[28] P. S. Heckerling, B. S. Gerber, T. G. Tape, and R. S. Wigton. Prediction of community-acquired pneumonia using artificial neural networks. *Medical Decision Making*, 23(2):112–121, 2003. ISSN 0272-989X. doi: 10.1177/0272989X03251247.

[29] A. Khazaee, A. Ebrahimzadeh, and A. Babajani-Feremi. Identifying patients with Alzheimer's disease using resting-state fMRI and graph theory. *Clinical Neurophysiology*, 126(11):2132–2141, 2015. ISSN 1872-8952. doi: 10.1016/j.clinph.2015.02.060.

[30] P. Lapuerta, G. J. L'Italien, S. Paul, R. C. Hendel, J. A. Leppo, L. A. Fleisher, M. C. Cohen, K. A. Eagle, and R. P. Giugliano. Neural network assessment of perioperative cardiac risk in vascular surgery patients. *Medical Decision Making*, 18(1):70–75, 1998. ISSN 0272-989X. doi: 10.1177/0272989X9801800114.

[31] Y. LeCun, Y. Bengio, and G. Hinton. Deep learning. *Nature*, 521(7553):436–444, 2015. ISSN 1476-4687. doi: 10.1038/nature14539.

[32] J. Li, Z.-Z. Chen, L. Huang, M. Fang, B. Li, X. Fu, H. Wang, and Q. Zhao. Automatic Classification of Fetal Heart Rate Based on Convolutional Neural Network. *IEEE Internet of Things Journal*, 6(2):1394–1401, 2019. ISSN 2372-2541. doi: 10.1109/JIOT.2018.2845128.

[33] C.-S. Lin, C.-C. Chang, J.-S. Chiu, Y.-W. Lee, J.-A. Lin, M. S. Mok, H.-W. Chiu, and Y.-C. Li. Application of an artificial neural network to predict postinduction hypotension during general anesthesia. *Medical Decision Making*, 31(2):308–314, 2011. ISSN 1552-681X. doi: 10.1177/0272989X10379648.

[34] N. Litvak, M. Rijsbergen, R. Boucherie, and M. Houdenhoven. Managing the overflow of intensive care patients. *European Journal of Operational Research*, 185:998–1010, 2008. doi: 10.1016/j.ejor.2006.08.021.

[35] N. Master, Z. Zhou, D. Miller, D. Scheinker, N. Bambos, and P. Glynn. Improving predictions of pediatric surgical durations with supervised learning. *International Journal of Data Science and Analytics*, 4(1):35–52, 2017. ISSN 2364-4168. doi: 10.1007/s41060-017-0055-0.

[36] K. McManus, E. K. Mallory, R. L. Goldfeder, W. A. Haynes, and J. D. Tatum. Mining Twitter Data to Improve Detection of Schizophrenia. *AMIA Joint*

Summits on Translational Science proceedings, 2015:122–126, 2015. ISSN 2153-4063.

[37] T. M. Mitchell. *Machine Learning*. McGraw-Hill Series in Computer Science. McGraw-Hill, New York, 1997. ISBN 978-0-07-042807-2.

[38] P. Natarajan, J. C. Frenzel, and D. H. Smaltz. *Demystifying Big Data and Machine Learning for Healthcare*. CRC Press, Taylor & Francis Group, 2017. ISBN 978-1-138-03263-7.

[39] OECD publishing. Health at a glance 2017: OECD indicators, 2017.

[40] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, and D. Cournapeau. Scikit-learn: Machine Learning in Python. *Machine Learning in Python*, 2011.

[41] E. Porter. A world of rising health care costs. *The New York Times*, 2013.

[42] M. Samudra, C. Van Riet, E. Demeulemeester, B. Cardoen, N. Vansteenkiste, and F. E. Rademakers. Scheduling operating rooms: Achievements, challenges and pitfalls. *Journal of Scheduling*, 19(5):493–525, 2016. ISSN 1099-1425. doi: 10.1007/s10951-016-0489-6.

[43] A. L. Samuel. Some Studies in Machine Learning Using the Game of Checkers. *IBM Journal of Research and Development*, 44(1.2):210–229, 1959.

[44] J. Schiele. Predicting Surgical Durations and Implications at the Operational Level. *Submitted*, 2019.

[45] J. Schiele and J. O. Brunner. A Framework for the Hospital of the Future. *Submitted*, 2019.

[46] J. Schiele, F. Rabe, M. Schmitt, M. Glaser, F. Häring, J. O. Brunner, B. Bauer, B. Schuller, C. Traidl-Hoffmann, and A. Damialis. Automated Classification of

Airborne Pollen using Neural Networks. In *41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2019. doi: 10.1109/EMBC.2019.8856910.

[47] J. Schiele, T. Koperna, and J. O. Brunner. Predicting ICU Bed Occupancy for Integrated Operating Room Scheduling via Neural Networks. *Naval Research Logistics*, 2020. doi: 10.1002/nav.21929.

[48] C. Thuemmler and C. Bai, editors. *Health 4.0: How Virtualization and Big Data Are Revolutionizing Healthcare*. Springer International Publishing, 2017. ISBN 978-3-319-47616-2.

[49] J. P. Tuwatananurak, S. Zadeh, X. Xu, J. A. Vacanti, W. R. Fulton, J. M. Ehrenfeld, and R. D. Urman. Machine Learning Can Improve Estimation of Surgical Case Duration: A Pilot Study. *Journal of Medical Systems*, 43(3):44, 2019. ISSN 1573-689X. doi: 10.1007/s10916-019-1160-5.

[50] A. Uyar, A. Bener, and H. N. Ciray. Predictive Modeling of Implantation Outcome in an In Vitro Fertilization Setting: An Application of Machine Learning Methods. *Medical Decision Making*, 35(6):714–725, 2015. ISSN 1552-681X. doi: 10.1177/0272989X14535984.

[51] J. M. van Oostrum, E. Bredenhoff, and E. W. Hans. Suitability and managerial implications of a Master Surgical Scheduling approach. *Annals of Operations Research*, 178(1):91–104, 2010. ISSN 1572-9338. doi: 10.1007/s10479-009-0619-z.

[52] M. Volkov, D. A. Hashimoto, G. Rosman, O. R. Meireles, and D. Rus. Machine learning and coresets for automated real-time video segmentation of laparoscopic and robot-assisted surgery. In *2017 IEEE International Conference on Robotics and Automation (ICRA)*, 2017. doi: 10.1109/ICRA.2017.7989093.

[53] I. Wallach, M. Dzamba, and A. Heifets. AtomNet: A Deep Convolutional Neural Network for Bioactivity Prediction in Structure-based Drug Discovery. *arXiv:1510.02855*, 2015.

[54] C. Ziello, T. H. Sparks, N. Estrella, J. Belmonte, K. C. Bergmann, E. Bucher, M. A. Brighetti, A. Damialis, M. Detandt, C. Galán, R. Gehrig, L. Grewling, A. M. Gutiérrez Bustillo, M. Hallsdóttir, M.-C. Kockhans-Bieda, C. De Linares, D. Myszkowska, A. Pàldy, A. Sánchez, M. Smith, M. Thibaudon, A. Travaglini, A. Uruska, R. M. Valencia-Barrera, D. Vokou, R. Wachter, L. A. de Weger, and A. Menzel. Changes to Airborne Pollen Counts across Europe. *PLoS ONE*, 7 (4), 2012. ISSN 1932-6203. doi: 10.1371/journal.pone.0034076.

Appendix A

A Framework for the Hospital of the Future

The following contribution [45] has been submitted to "Journal of Business Economics", which is ranked in category B in the VHB-JOURQUAL3 ranking [2]. The submitted version is reproduced below in its entirety.

A Framework for the Hospital of the Future

Julian Schiele[1,2], Jens O. Brunner[1,2,*]

[1] University Center of Health Sciences at Klinikum Augsburg (UNIKA-T)

Neusässer Straße 47, 86156 Augsburg, Germany

[2] Chair of Health Care Operations / Health Information Management

Faculty of Business and Economics, University of Augsburg

Universitätsstraße 16, 86159 Augsburg, Germany

[*] Corresponding author: jens.brunner@unikat.uni-augsburg.de

Submission date: December 30, 2019

A Framework for the Hospital of the Future

Abstract

Today's hospitals around the globe face social, organizational, medical, and financial challenges. The hospital of the future in 2040 will be characterized by digitization, integration, automation, and personalization in both support processes as well as core processes using cyber-physical systems and data analytics. There will be less bureaucracy, more efficient processes, higher staff satisfaction, more time spent for value-adding tasks and with the patient, better medicine, and - after all - better treatment of patients. The human component, however, will also remain essential in the future. Not few hospitals have already started to discuss and implement selected use cases such as electronic health records or surgery robots. Even though these individual lighthouses might be innovative and beneficial, most hospitals still lack a holistic strategy on the way to become a hospital of the future.

In this study, we present a conceptual framework for the hospital of the future structured in 32 dimensions along seven areas and based on four enablers as foundation. This *HoF framework* reflects both latest scientific findings as well as hands-on insights experienced in practice. We conducted an extensive survey with more than 265 hospital experts such as managers, physicians, nurses, and IT professionals. Hospital managers have high ambitions to enhance their hospital and expect significant benefits, but they are struggling to build momentum. Employees & skills as well as IT infrastructure & data security are critical enablers and there is need for action in nearly all dimensions of the framework. This study serves as valuable guidance for hospital managers to advance their hospital to the next level and for researchers to categorize their work. Eventually, we provide a four-step approach to help them getting started.

Keywords: Hospital of the future; hospital 4.0; smart hospital; vision; digital health; framework

Submission: December 30, 2019

1 Motivation

The overall health care system and in particular hospitals around the globe face various social, organizational, medical, and financial challenges and must constantly adapt to a changing environment (see Figure 1). Driven by an increasing life expectancy worldwide [69] as well as lower fertility rates [93], demographic change is a major driver of increasing health expenditures per capita [26]. Furthermore, it also imposes qualitative challenges such as multimorbidity to hospitals. Among others, urbanization and the shift to outpatient treatment require the hospital system to adapt accordingly, e.g., increasing home care delivery. Also global developments such as the climate change or global epidemics demand their tributes. Not only the society overall is changing, but also do patients as individuals: their health awareness and expectations are increasing and they demand more autonomy and participation in the treatment process, e.g., flexible access to care, regular updates, and online planning tools. In addition to those social challenges, also

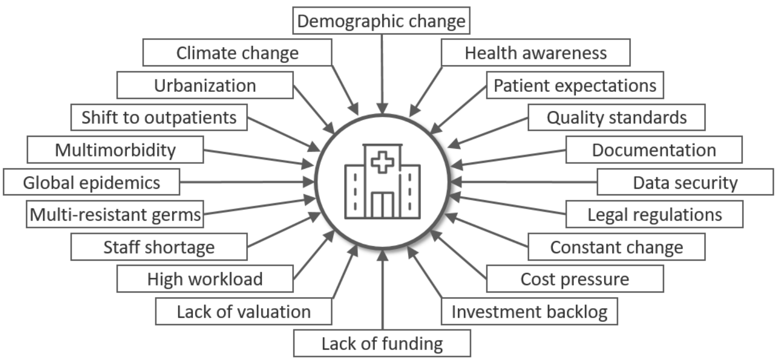

Figure 1: Today's hospitals face social, organizational, medical, and financial challenges.

organizational requirements become more demanding and complex. Next to their regular tasks, physicians and staff are asked to fulfill time-consuming documentation and comply with ever increasing legal regulations imposed by authorities. The regulatory environment of hospitals is confusing, complicated, and changing. In particular, data security is an highly sensitive topic that needs to be treated with utmost care. Prominent incidences such as 2006's hacker attack to blackmail German Lukaskrankenhaus [91] or 2019's global leak of more than 16 millions of picture archiving communication system (PACS) data

[39] (also see [92]) vividly demonstrate the high risk and the far-reaching consequences. The challenging situation is further aggravated by an increasing cost pressure from both public and private sponsors. In most countries around the globe, the health care industry accounts for a large share of expenditures and is still growing. For example, in the United States, 3.3 trillion USD or nearly 18% of its GDP were spent on health care in 2016 reflecting an annual growth of 4.3% compared to 2015 [19]. Similarly, the OECD countries spent on average 9% of their GDP [70] and Germany 11% [33]. A closer look reveals hospitals as main driver accounting for 32% in the United States and nearly 40% in the OECD [19, 70]. With rising costs, hospitals are increasingly attracting attention from sponsors in both the governmental and the private sector demanding more cost effectiveness while ensuring the same level of service quality. “Pressures to make operating margins will continue to be at the forefront of most hospital and health system leaders' minds” [66] - particularly, since it seems that nothing “will stop public spending on health care from rising” [74]. As a consequence, most hospitals suffer from an investment backlog and lack of funding. Commonly, the response to the cost pressure is to cut costs such as payment levels and benefit structures. This directly leads to one of the most critical challenges: the acute shortage of staff in hospitals. In German hospitals it is particularly challenging to find, recruit, and retain qualified nursing staff and the situation has further intensified with the introduction of the German regulation for the threshold of nursing staff in 2019[1]. However, there is not only a shortage of nurses, but the survey results confirm also a lack of qualified physicians, administration staff, and IT experts. Given the changing environment, adequate qualification and continuous training becomes even more important. Moreover, most personnel suffers from unpleasant working conditions, high workload, lack of valuation, and difficulties to reconcile work with personal time, i.e., on weekends and during night shifts. In this setting, it is very challenging to perform value-adding tasks with the necessary diligence and spend sufficient time with each individual patient.

[1] https://www.bundesgesundheitsministerium.de/personaluntergrenzen

The purpose of this paper is to develop a vision for the hospital in 2040 with the intention to advance hospitals and take some pressure off the system. To sharpen the presented vision, we have considered both latest findings published in scientific journals as well as valuable input provided by more than 265 practitioners in response to our survey. Our contribution is the conceptual *HoF framework* for the hospital of the future structured in 32 dimensions along seven areas and based on four enablers as foundation. Moreover, we underpin the vision with results of the extensive survey representing the perspective of hospital managers, physicians, nursery, and IT leaders. We believe that this study provides a valuable guidance for practitioners to navigate the multitude of use cases in a structured manner.

The remainder of this paper is organized in five sections. In Section 2, we explain the concept of *Health Care 4.0*, provide an overview of previous literature, distinguish between six stakeholders, and derive a definition for the hospital of the future. The target picture for the hospital of the future is detailed and depicted in Section 3. In Section 4, we present the *HoF framework* for the hospital of the future and describe each of the 32 dimensions in detail. Moreover, we also explain and evaluate the four critical enablers. Section 5 proposes a four-step approach which is intended to support practitioners who want to apply the framework and enhance their hospitals to the next level. Finally, we conclude our findings and discuss managerial insights in Section 6.

2 From Health Care 4.0 to the hospital of the future

Inspired by the fourth industrial revolution that led to the smart factory [108, 60, 56], similar concepts will also help to advance the health care system. Historically, *Industry 1.0* refers to mechanical production facilities in the late 18th century such as steam powered machines and the weaving loom. The second industrial revolution introduced electrical energy, the assembly line, and mass production. In *Industry 3.0*, computers and electronics paved the way for automation of machines. *Industry 4.0* or smart factory refers to a factory in which machines are augmented with sensors, everything is connected by the so-called internet of things (IoT), decisions are decentralized, processes are visualized and automated, and cloud computing as well as artificial intelligence (AI) are realized. Similarly, these technical, social, and organizational innovations can also make a valuable contribution to overcoming the challenges in the health care system. This concept is called *Health Care 4.0*. In the following, we will have a short look at previous literature addressing *Health Care 4.0*. Afterwards, we will focus on hospitals and develop a concept for the hospital of the future.

In the past years, many articles have been published on the concept of *Health Care 4.0* which is often referred to as smart health or digital health. Instead of cutting costs and reducing value-added care, this concept focuses on reducing waste and increasing efficiency. According to Berwick *et al.*, at least 20 % of total health care expenditures could be eliminated by addressing overtreatment, failures in coordination and execution of care processes, inefficient pricing, administrative complexity, fraud and abuse [15]. Thuemmler and Bai [95] define *Health Care 4.0* as “a strategic concept for the health domain derived from the *Industry 4.0* concept”. They argue that the patient flow and patient pathways would be the classical models of value chains within the health care industry, which is not any different to other industries. Moreover, they name virtualization in order to enable personalization as aim of *Health Care 4.0*. This should be achieved through the usage of cyber-physical systems, cloud computing, IoT, and mobile communication networks. They also highlight that in contrast to *Industry 4.0*, more attention must be paid to safety, security, and resilience in health infrastructures. Mohanty *et al.* [64] describe the concept as a “combination of various entities, including traditional health care, smart biosensors, wearable devices, ICT, and smart ambulance systems”. Particularly great attention was given to various applications of machine learning (or in broader terms: AI) in health care. Dua *et al.* [28] composed an extensive collection of articles on machine learning in health care informatics and present a variety of applications such as screening of arrhythmia from electrocardiography, regulation of glucose level for diabetic patients, detection of neuropsychiatric disorders such as Alzheimer's disease, and clinical decision making. Also Natarajan *et al.* [66] present a valuable collection of big data and machine learning based applications in health care, i.e., precision medicine, financial reporting, and medical imaging. Chen *et al.* [20] state that combining machine learning systems with human practitioners will improve our collective health. The fact that big data and machine learning offer unprecedented analytical power, wearables become omnipresent, and investment in digital health has overtaken funding for biotechnology [88] indicate that we are on the cusp of a new era in health care. Hence, it is not surprising that - next to the scientific literature - also consultancies, institutes, and other cooperation study the concept of *Health Care 4.0*. Aboshiha *et al.* [2] describe that "health care players are using AI to address significant inefficiencies and open up powerful new opportunities". They list various applications "from the delivery of remote health care services to the early diagnosis of disease and the hunt for new life-saving medicines" such as heart monitors, smart glucose pumps, and diagnostic devices. Across the health care value chain, they identify seven areas for opportunities of AI in health care: remote prevention and care, diagnostics support, treatment pathways and support, drug discovery and development, operations, marketing and sales, and support functions. Hipp *et al.* [45] describe use cases, benefits, challenges, and recommendations for the digital transformation of the health care sector. They present a three-step roadmap from (1) today's health care *on loose strings* which is characterized by a lack of continuity, transparency, and digitization to (2) a *central healthcare platform* in which stakeholders are digitally connected to a central entity towards (3) the future *patient-centered health system* in which patients have control over their information and stakeholder interaction. Similarly, Clawson *et al.* [23] observe that health care is one of the least mature, most regulated, and least efficient industries despite being among the largest in most developed countries. Consequently, they postulate the end of the current system and ask both policymakers and responsible industry leaders for a transformation such that "competitive forces will promote innovations and development that improve health care value". Finally, the concept of *Health*

Care 4.0 is also covered in numerous recent blogs [25, 27] and newspaper articles [65].

As stated in the aforementioned literature, we are on the cusp of a new era in health care. This has implications for the four traditional health care sectors, i.e., providers, payers, medical technology companies, and biopharma companies, as well as for patients and technology companies. Figure 2 depicts these six key stakeholders in the health care industry.

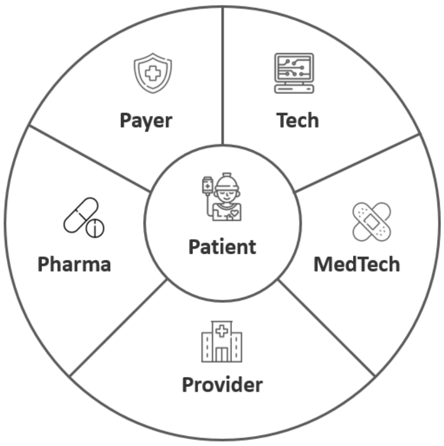

Figure 2: The six key stakeholders in the health care industry.

Patients are the core of the health care system and will see positive impact in many ways. Thanks to efficiency gains and technological as well as medical innovations, they are served with cheaper insurance, better diagnostics, more accurate monitoring, tailored pharmaceuticals, and an improved treatment. Additionally, they will participate more actively in health care processes and make use of remote health services from home.

Payer refers to private and statutory, for-profit and not-for-profit health insurance companies. Big data analytics and AI will yield major efficiencies in claims handling and detection of fraud, waste, and abuse. Also population health management (PHM) will play a more important role. The GKV 4.0 trend monitor [17] identifies and evaluates nine technology trends that are relevant for statutory health insurance companies, i.e., mobile, big data and advanced analytics, robotics, cloud computing and connectivity, AI, affective computing, IoT, virtual reality (VR) and augmented reality (AR), and blockchain.

Biopharma companies have already started to implement use cases of *Health Care 4.0*. For example, they improve the efficiency of drug discovery by leveraging clinical data and partnering with biobanks [2]. Moreover, precision medicine and genome research offer new business potentials.

Medical technology companies will improve the efficiency of their operations, i.e., identifying bottlenecks and preventing stock-outs on critical products. Moreover, they will develop new products and solutions such as intelligent monitoring and diagnostic devices. Hosseini *et al.* [47] provide an overview of data-driven business models and illustrate selected medtech companies that have already started.

Technology companies are not yet considered part of the traditional key segments, but increasingly become attracted by the new opportunities in health care. An increasing number of smaller companies and start-ups offer innovative solutions, i.e., to discover drugs (Atomwise), to connect patients with general practitioners (GPs) (Babylon Health), and to analyze medical imaging (Zebra) [2]. Given their consumer-friendliness and digital sophistication, particularly technology giants such as Alphabet, IBM, Apple, Amazon, and Alibaba are well suited and already started to invest significantly in health care. Saxena *et al.* [81] show how they might disrupt each of the industry's traditional segments. Their algorithms will become more integrated in core processes of biopharma and payers, they will increasingly handle clinical decision support and diagnostic tools for providers, and they will offer products and solutions in the domain of medical technology companies.

Provider refers to hospitals, GPs, and transitional care providers such as rehabs and hospices. They will profit from increased efficiency, automation of time-consuming support processes such as billing, better clinical decision making and allocation of resources such as nursery rosters, and more efficient drug utilization. Better diagnostic and treatment yield an improved outcome and consequently, also result in fewer complications and readmissions. In the following of this study, we focus on hospitals.

Although literature dedicated to the hospital of the future is rare, we found some previous work published by researchers, consultancies, institutes, and in newspapers. Afferni *et al.* [3] present a framework for the *Hospital 4.0* and discuss methodological and technological innovations. They name methods and policies, organization, and technology as main pillars and demand a transformation from today's *poly-functional center* where the patient is considered as object to an integrated, *patient-centered approach* where the patient is an active user of an integrated health service. In this model, the patient interacts only with dedicated client care services as single contact point while different expertises are provided by clinical care services. Even further in the future, they envision a *patient-centric approach* where smart patients are actively participating and collaborating within an ecosystem of integrated health care services. Ribera *et al.* [80] studied the hospital of the future to understand potential changes impacting European hospitals. Based on surveys and interviews with hospital decision makers from Karolinska University Hospital in Stockholm and Hospital Clínic of Barcelona, they develop a conceptual framework with five main dimensions: context, strategy and leadership, resources and capabilities, processes, and results. Furthermore, they extract 14 key messages to summarize the identified challenges and characteristics. Eichhorst *et al.* [30] analyze recent trends and conclude that small, independent full service hospitals will disappear. Instead, the health care system will be composed of a dynamic, integrated network structure in which various providers coexist and cooperate, i.e., excellence centers for complicated interventions, day clinics for routine tasks, outsourcing of clinical services such as in radiology, remote treatment at home, gyms, and external care solutions. Fraunhofer institute's Wibbeling *et al.* [101, 102, 103] published several position papers on topics such as *Hospital 4.0*, smart devices in hospitals, and human centered digital hospitals. Gimpel *et al.* [40] study digitization of logistic processes in hospitals such as bed management and warehousing. In a survey on digitization and integration in German hospitals [52], 87 % of the participants agree that the advantages of digitization outweigh the associated difficulties. The most critical challenges are lack of funding (61 %), lack of IT capacities (54 %), and heterogeneous IT structures (48 %). The newspaper article by Schwinn [83] provides some vivid insights into the everyday life of a connected hospital using the example of German Universitätsklinikum Hamburg-Eppendorf.

3 Sharpening the vision

Based on the aforementioned literature, the vision for the hospital of the future has now become more tangible. The hospital of the future in 2040 will be characterized by digitization, integration, automation, and personalization in both support processes as well as core processes using cyber-physical systems and data analytics. There will be more integration within the hospital and with other stakeholders, less paper-based work and less bureaucracy, efficient and optimized processes, more automation and data-driven solutions, transparency and sustainability, more time spent for value-adding tasks and with the patient, higher quality and safety, and after all better medicine and better treatment leading to higher satisfaction of staff and patients. However, it is also important to consider the differences between the smart factory and the hospital of the future. While production constitutes the core process in a factory, it is the diagnosis and therapy process in a hospital. While we see mostly centralized decision structures in factories, most decisions in hospitals are already made decentralized by physicians and nurses on the ground. While products are core in a factory, it is the patients in hospitals. While human-human interactions can be minimized in factories, they are essential for the recovery process of patients and will remain a valuable component in the future. This humanizing dimension is beneficial in multiple ways: patients find it easier to express their feelings, physicians are able to show compassion, humor is helpful to lessen fear and increase pain tolerance [89]. Taking these differences into account, the concept can be transferred to the hospital setting serving as assistance for support processes such as cleaning and catering as well as core processes such as diagnostics and surgery. Not few hospitals have already started to discuss and implement selected use cases such as electronic health records (EHR) or surgery robots. Even though these individual lighthouses might be innovative and beneficial, most hospitals still lack a holistic strategy on the way to become a hospital of the future.

It is the purpose of this study to develop such a structured framework for the hospital of the future in 2040. Our work is based on the aforementioned literature to account for latest scientific, theoretical, and state-of-the-art findings, as well as on a survey among hospital practitioners to complement the vision with their individual experiences from a practical perspective. The detailed survey has been conducted with more than 265 participants from public (65 % of participants), church and charity (25 %), and private (10 %) hospitals with focus on Germany (95 %). The participating hospitals differ in terms of level of care (59 % offering maximal care, 6 % primary care), size (31 % with more than 1,500 beds, 27 % less than 250 beds), and hospital type, i.e., covering big university hospitals (41 %) such as Charité in Berlin, UK Heidelberg, UK Augsburg, UK Köln, Universitätsmedizin Göttingen, and UK Hamburg-Eppendorf as well as hospital chains (2 %) such as Schön Kliniken, Asklepios, and Helios as well as regular hospitals (50 %) and others (7 %). A balanced picture is ensured thanks to the even distribution of the participants' professions, i.e., physicians and nursery (36 %), management, administration, and consulting (26 %), and IT experts (38 %). There were slightly more male participants (67 %) and ages range from less than 30 years (12 %) to 31-40 years (25 %) to 41-50 years (27 %) and older (37 %).

The survey reveals significant differences in the status of individual hospitals, but there is a lot of catching up to do everywhere. Many hospitals still rely on paper-based documents and fax machines while others already experiment with care robots and 3D printing. Asked for their key challenges and pain points, topics related to staff shortage were named most often (26 % of all given answers[2]) highlighting the difficulties associated with acquisition and retention of nursery, physicians, and IT experts. In German hospitals it is particularly difficult to find, hire, and retain qualified nursing staff and the situation has even intensified with the introduction of the German regulation for the

[2]Free text answers, up to three answers per participant

threshold of nursing staff in 2019[3]. Also, financial challenges are prevailing (11 %) including lack of funding, cost pressure, decreasing revenues, and backlog of investments. Many participants mention topics related to the working conditions (9 %). They are unsatisfied with their workload, have difficulties to reconcile private life with work, complain about weekend and night shifts, unbalanced schedules, and unpaid overtime. Topics related to documentation and bureaucracy were perceived as major challenge as well (5 %). Participants mention the increasing number of legal requirements, time-consuming and repetitive task, and repeated documentation in multiple systems. Many participants mention the hospital organization as significant challenge (5 %) including inefficient workflows, discontinuous processes with media interruptions, and in particular strict hierarchies that hinder innovative solutions in line with the saying "eminence instead of evidence". Others state concerns with the IT infrastructure and data security (5 %). They have difficulties with outdated systems and products and are afraid of data leaks and cyber attacks. On average, only 15 % of the processes in a hospital are automated, while 52 % are supported by simple IT tools such as Microsoft Excel, and 33 % are performed manually without any support. Most hospitals have already started to develop, detail, and implement selected use cases. Most common use cases are electronic health records (mentioned 69 times[4]), data management in hospital information systems (33), digital monitoring of vital patient parameters (25), and surgery robots such as Da Vinci (17). Some participants also mentioned staff rostering (11) and operating room (OR) scheduling using tools such as Agfa ORBIS (9), automated imaging in radiology (6), unit dose systems for pharmaceuticals (2), and additive manufacturing in areas such as oral surgery (2).

The survey results clearly confirm that hospitals have high ambitions for the future and expect significant benefits (93 % of participants) from the concept of the hospital of the future. Most hospitals (85 %) plan a timely implementation, thereof the majority (62 %) already within the next 2 years. Participants expressed their hope for a better medicine, increased patient safety, elimination of human errors, more efficient processes,

[3]https://www.bundesgesundheitsministerium.de/personaluntergrenzen
[4]Free text answers, up to five answers per participant

better cooperation and communication, and less bureaucracy. For each of seven key benefits (see Figure 3), the participants were asked to evaluate the expected potential to be addressed by the *HoF concept* on a scale from 0 (low potential) to 4 (high potential). Moreover, they were asked to evaluate the relative importance of each benefit on a scale from 0 (low importance) to 4 (high importance). Figure 3 depicts the resulting evaluation matrix. Most benefits find themselves in the upper right corner significating that the selected benefits are relevant and that the hospital of the future concept has high potential to achieve those benefits.

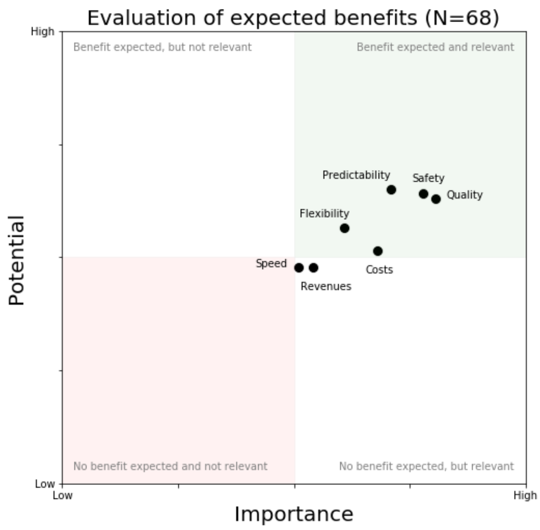

Figure 3: Evaluation by importance and potential for expected benefits.

However, most hospitals struggle to build momentum and do not feel well prepared (27 % completely or almost unprepared, 43 % only a little prepared). Participants highlight that they need additional and qualified staff, a shift of mindset of the staff, i.e., acceptance, motivation, courage, and support, high quality training to master the transformation, sufficient funding, as well as a clear vision, strategy, roadmap, and governance structure. Common obstacles are lack of acceptance and responsibility, required time and efforts, missing support of the management, restricting legal regulations, and lack of funding. Some participants also expressed their concern of blindly relying on machines (also see [62]), lack of social contacts, inhuman medicine, and dismissals.

In preliminary studies, we talked to industry experts and developed initial hypotheses. Furthermore, taking into account the extensive literature research and the detailed results of the survey, we are able to illustrate how we envision the hospital in 2040. Figure 4 depicts the target picture for the hospital of the future covering all essential areas. Overall, we distinguish between seven key areas. From area 1 (Patient) to area 7 (Infrastructure), each area is marked with a unique icon in Figure 4 and explained in the following.

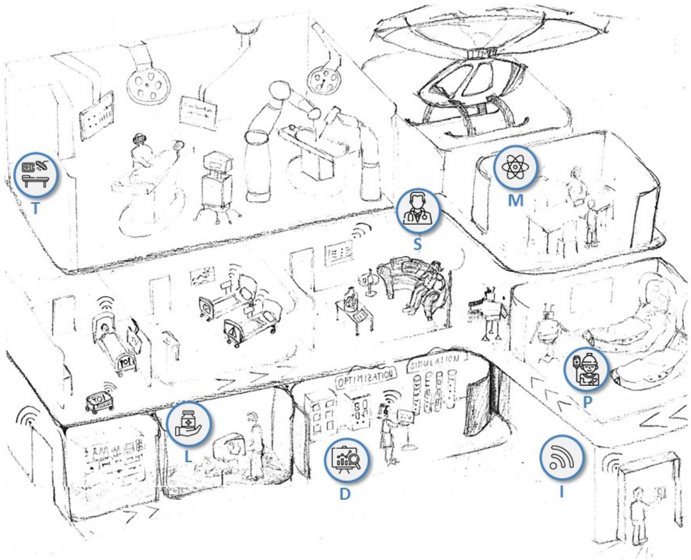

Figure 4: Target picture for the hospital of the future (sketch by Almut Rummel).

Patient (P) As described in Section 2, the individual patient is core in future *patient-centered models* [3]. Patients participate actively in various health care processes and have access to their digital files at all times. From the perspective of a patient, the experience already starts at home in form of automated monitoring and smart pharmaceuticals. Within the hospital, they are transported decentralized and autonomously to all stations that are needed for their recovery process.

Staff (S) In contrast to today's hospitals, physicians, nurses, and other staff are more satisfied in the future and their workload is more balanced thanks to smart staff scheduling. Mobile solutions, visualization and assistant systems make everyday work easier allowing to spend more time for value-adding tasks and with the patient.

Treatment and intervention (T) Diagnosis, intervention, and care are considered core processes. In the future, they are supported by robotics, planning tools, and smart AI solutions. Also, telemedicine and precision medicine play a vital role.

Logistics and supply (L) Supporting processes such as warehousing, cleaning, laundry, and catering are more efficient in the future. Connected devices, collaborative robots, automated replenishment, and autonomous transportation are incorporated in order to automate and optimize processes.

Management and organization (M) The hospital of the future looks different from a managerial and organizational point of view. In particular, the concept of value-based health care and the integration of various players along the value chain yield major improvements.

Data and control (D) Data is essential for nearly all processes in the hospital. A centralized hospital information system forms the foundation for paperless documentation, big data analytics, simulation, and optimization.

Infrastructure (I) The *HoF concept* has also implications for the design and construction phase of hospitals. Buildings, equipment, and systems are modular, integrated, standardized, and sustainable.

In Section 4, we propose the structured *HoF framework* for the hospital of the future along those seven areas and discuss each *HoF area* in more detail.

4 The conceptual HoF framework

We have learned that not few hospitals have already started to discuss and implement selected use cases, but still lack a holistic concept that guide them on the way towards the hospital of the future. In this section, we present the structured *HoF framework* that orchestrates the individual use cases in 32 *HoF dimensions* along the seven *HoF areas* that were identified in Figure 4. In the framework in Figure 5, the seven *HoF areas* are depicted as pieces of pie and illustrated by dedicated symbols. The *HoF dimensions* are depicted as white balls and marked with unique identifiers that correspond to the detailed description hereafter. Finally, we also distilled the four most critical enablers that are indispensable prerequisites to become a hospital of the future. They are depicted in the rectangular boxes below the pie chart.

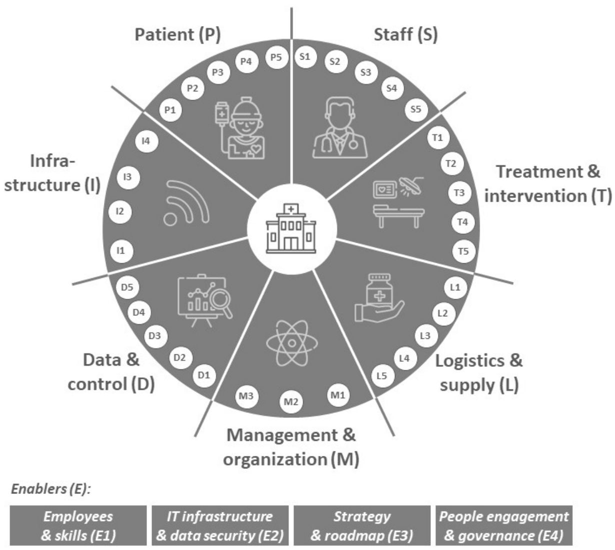

Figure 5: The *HoF framework* for the hospital of the future.

Hereafter, the key areas, the associated dimensions, and the enablers that constitute the framework for the hospital of the future are described in more detail.

4.1 Patient (P)

This area aggregates the five dimensions that are directly related to the patient. In the hospital of the future, patients make use of EHRs (P1), automated monitoring and preventive care (P2), as well as smart devices and medication (P3). Moreover, they participate actively (P4) and follow a decentralized patient flow (P5). The numbers in brackets for each dimension refer to the unique identifier as depicted in Figure 5.

Digital twin of the patient (P1). In the hospital of the future, all patient related information is stored digitally and accessible from mobile devices. This so-called *digital twin* of the patient goes much further than today's electronic medical records (EMRs), EHRs, or patient data management systems (PDMSs). In fact, it aggregates every piece of data at one location such as the entire medical history, insurance data, contact data, prescriptions, allergies, implants, detailed data of past surgeries, past lab results, computer tomography (CT) and x-ray scans, past treatments received by other health providers, real-time medication data and vital parameters, real-time location information, history of interaction with virtual agents and human staff. Both patients as well as medical staff are provided access via mobiles, tables, and at dedicated service-points in the hospital. Moreover, they are provided with learning modules tailored to their needs. Patients have full control over their data and have the option to contribute some of their data for medical studies. Both patients and staff benefit from this dimension. Tedious calls to GPs are a thing of the past, since patients stay automatically informed at all times about their current status, the outcome of diagnostics, and suggested treatment plans. Similarly, staff has access to required patient information regardless of location and time and are supported by automated warnings and recommendations based on big data analytics. Overall, this results in less bureaucracy, more transparency, higher efficiency, and improved patient outcomes. EHRs are studied in the literature [22, 61], and also implemented in practice. The hospital in Mühldorf am Inn, Germany, offers an app for patients [67], the UK Hamburg-Eppendorf have replaced paper-based documents with EHRs, and the United States government stimulated the usage of EHR with USD 30 billion in the HITECH Act.

Automated monitoring and preventive care (P2). In the future, continuous monitoring and tracking of patient parameters are essential on three levels: preventive, in-hospital, and post-care. On all levels, the gathered data is automatically analyzed in real-time, provides valuable insights, and triggers alarms or warning alerts in case of anomalies. Preventively, patients continuously monitor their state of health, e.g, vital signs, glucose level, and weight, using mobile technologies and smart devices such as weareables, e-skins, and smart home. Preventive monitoring which is also called quantified health can help to recognize health risks such as heart diseases and obesity before they become a major issue. Also, it helps to manage chronic health conditions such as chronic obstructive pulmonary, hypertension, diabetes, and high cholesterol by recommending personalized behavior plans. In the hospital, the state of health is continuously and automatically monitored by smart, patient-friendly devices such as smart beds and complemented with data from even more equipment such as electrocardiograms. In addition, detailed health checks are recorded as part of the admission process, e.g., as emergency triage, and before being released from the hospital. Post-care monitoring makes it possible for patients to leave the hospital much earlier after surgery than today. Even without physical contact, physicians can interact with patients, recurring pain can be treated, the risk of complications can be managed, and the chances of readmission can be predicted. On all three levels, we see positive impact such as reduced manual tasks, improved patient care, increased patient safety, and saved costs, i.e., an estimated USD 200 billion annually in the United States [2]. Already today, data from consumer wearables, scales, and glucometers are integrated into electronic medical records through smartphone applications such as Apple's HealthKit allowing to track changes and identify warning signs. As another example for preventive care, IBM and Malteser partnered to develop an inhome elderly care monitoring solution [12] using sensors for water, fire, video, and a bed mat with sensors which could be enriched by intelligent fall sensor systems to prevent falls, daily video calls, and remote trainings in the future. Agnihothri *et al.* [4] provide an approach to quantify the benefit of using mobile health technology to manage chronic conditions. An example for in-hospital monitoring is an early warning system to reduce postoperative surgical-site and sepsis infections based on patient's vital signs during the operation [38, 42]. Other use cases include smart eyeglasses to monitor blood flow and color [75], smart textiles and on-body sensors [64], non-contact measurement of respiratory and heart rate in hospital beds [73], monitoring of dementia people safety and health via mobile applications [106], early-warning system for depression based on social media images [79], and detection of schizophrenia based on Twitter interaction [63].

Smart devices and medication (P3). In the hospital of the future, pharmaceuticals and medical devices are smart, connected, and personalized in order to improve patient treatment and well-being. Bender *et al.* [95] define smart pharmaceuticals as "an electronic package, delivery system, or pill that offers one or more examples of *intelligent added value* ". As such smart added value, devices check the regular intake of drugs as prescribed, e.g., by using smart dispensers, send reminders, and are connected in order to analyze and document data. Moreover, medication plans can be tailored to individual daily needs. Smart devices and medication contribute to a higher quality of care, increased patient safety, and higher patient satisfaction. For example, the smart glucose pumps which are planted under skin provides great support to diabetes patients by automatically controlling glucose levels and injecting insuline if needed. Another example would be the treatment of patients with mental health by leveraging AI technology to offer daily chat conversations, mood tracking, and curated videos. Yousaf *et al.* [106] provide an overview of mobile apps for activities of daily living based cognitive training. Medication errors can be prevented by closed-loop medication systems, e.g., by scanning a patient's wristband for validation before drug intake as realized in German Agaplesion Diakonieklinikum Rotenburg. Other examples are tailored behaviour plans using automated monitoring of airborne pollen [82], smart walkers [6], portable spirometers [78], computer based training games [86], and IoT-enabled medicine boxes [90].

Patient participation (P4). As discussed in Section 2, smart patients are actively participating and collaborating within an ecosystem of integrated health care services [3]. In a self-administered health care system [95], patients stay well informed at all times, choose between several options and providers, and autonomously use self care applications such as digital self sign stations on arrival at the hospital or a patient app to receive the outcome after diagnostics. Moreover, they might also choose to participate in clinical studies by sharing selected data, e.g., submitting ophthalmology test results. Of course, patients who are unconscious or unable to make decisions are not forced to participate, but are cared for by smart systems and experienced staff. For the patient participation, human-machine interaction plays an important role. Advanced multi-directional interactions between patients and devices, health processes, and systems are enabled using user-friendly mobile apps, online patient interfaces, virtual chat bots, voice recognition, face recognition, and radio-frequency identification (RFID) chips. Patients book appointments in one-click digital booking systems, communicate with virtual agents from home to clarify first questions, check-in at the hospital using face recognition, adjust machines to their individual needs using RFID wristbands, and collect digital prescriptions. Hence, health services become more accessible, personalized, and user-oriented [87]. Recent research has also shown improved outcomes and lower readmission rates for patients that actively participate in the treatment process [50]. In a study by Jain *et al.* [48], patients seeking birth control were actively involved by online questionnaires leading to shorter and better visits. In Denmark, patients are provided online access to a digital platform which can be used to schedule appointments, retrieve digital prescriptions, or submit x-ray scans to the GP. Other examples are conversational robots that prepare patients for a treatment such as an MRT scan, robots that explain test results, mobile apps that provide diagnostic results such as skin cancer self exam, direct-to-consumer tests that can be done at home, e.g., genetic testing provider 23AndMe.

Decentralized, autonomous patient flow (P5). This dimension addresses the flow management and treatment pathways within a hospital. In the future, patients are guided by digital signage systems or transported by autonomous transport systems, e.g., in laser-guided beds or wheelchairs. Furthermore, a patient's treatment path from admission to discharge is steered dynamically and decentralized. In particular, a patient-specific device knows all stations that need to be visited, communicates in real-time with other devices and stations using RFID technology, and dynamically computes the optimal route under consideration of underlying constraints, e.g., MRI needs to be visited before surgery. The individual treatment requirements are updated regularly based on the outcome of previous stations and in line with the concept of value-based health care (VBHC) (see M2). The treatment path is not limited to the boundaries of the hospital, but does also cover stations outside of the hospital such as rehab facilities and GPs. This results in increased efficiency, lower waiting times, and higher flexibility. Decentralized, autonomous patient flow is inspired by recent development in other industries. Starting as a simple RFID chip in a diary processing plant, the stored specifications such as flavor and packaging could be communicated in each production process directly to the respective machines so that right cup size is selected and the yoghurt with the desired flavor is filled. In Audi's plant in Heilbronn, Germany, the Audi R8 is not moved on a traditional conveyor through the assembly area, but transported on a dynamic automated guided vehicle (AGV) guided by a laser scanner [56]. Asamoah *et al.* [8] showed that RFID-based information visibility for scheduling of lab services in a hospital result in lower wait times and better resource utilization.

4.2 Staff (S)

This area aggregates all dimensions that are directly related to the staff. In the hospital of the future, staff attends immersive trainings (S2), use smart visualization and assistant systems (S3), and mobile solutions (S4). Moreover, shortage of staff is reduced and staff satisfaction increased thanks to smart staff scheduling (S1) as well as dedicated HR initiatives (S5). The numbers in brackets refer to the unique identifier for the dimension as depicted in Figure 5.

Smart staff scheduling (S1). In the hospital of the future, personnel shift scheduling is done automated by software instead of highly-qualified personnel. The system computes a schedule under consideration of qualifications, time and location, contract regulations, fairness, training and development, and individual preferences that can be entered by the staff. Furthermore, big data analytics and machine learning are used to generate valuable insights and further increase quality and efficiency. This leads to better schedules, higher staff satisfaction, and lower planning overhead costs. An example for automated staff scheduling is the smart tool PLANFOX by XITASO Healthcare GmbH.

Immersive training (S2). The staff in the hospital of the future is better prepared for their tasks thanks to interactive and virtual training experiences. E-learning platforms and shared knowledge databases help them to stay up-to-date, quickly acquire additional qualifications, and learn from best practices of other hospitals. Realistic environments are created using VR, AR, and simulations such that new tasks can be trained before patient contact, e.g., planning and execution of complex surgeries to avoid complications. Some of the positive impacts are a reduced ramp-up time of new staff, higher quality of care, and higher patient safety. In Austria, the Niederösterreichischen Landeskliniken-Holding has launched a medical knowledge database called *eRef* for education and training purposes [57]. Elliman *et al.* [31] provide an overview of current developments for VR simulation applied to the education for nurses and Yousaf *et al.* [106] present studies addressing dementia education.

Visualization and assistant systems (S3). The staff is assisted by smart and integrated systems which show relevant information. This improves the staff satisfaction, provides support during complex tasks, and enhances the process quality. For instance, large electronic boards in the nursery room show real-time information about patients in the wards, their health status, and pending tasks. A smart glass connected to an expert system can support during repair and maintenance of medical equipment. While conducting surgeries, smart glasses can support surgeons, e.g., by indicating where to cut and where selected structures are located. The visualization and assistant systems can also be accessed from remote, e.g., to ask an expert for another opinion during surgery. Other examples are access control using face recognition, digital signage on rooms and beds, voice authentication using to verify identity.

Mobile solutions (S4). Mobile solutions support the staff in the hospital of the future and offer access to all required information regardless of the location. Employees can access EHR for individual patients, directly communicate with other providers, visualize performance metrics, and receive real-time updates. Moreover, some tasks that do not require physical presence might also be performed from home. This leads to higher staff satisfaction, higher efficiency, higher flexibility, higher transparency, and reduced response times. Some hospitals provide already mobile workstation for nurses. At the Lukaskrankenhaus Neuss, the project *Visite 3.0* has been rolled out to all nursing stations providing access to patient information, medication, prescriptions, digital forms, x-rag imaging, and lab results on tablets and mobile apps.

Staff acquisition and retention (S5). A professional, dedicated HR organization becomes a major component of the hospital of the future. To attract more potential employees, innovative and tailored recruiting campaigns are launched leveraging big data insights. Employee-friendly working models are developed that give employees more freedom and flexibility, e.g., flex-time, job rotation, and home office. Overtime, night shifts, and weekend work are balanced, monitored, and compensated adequately. Each employee is supported by a career advisor in order to discuss individual development paths, show appreciation for good work, remove hurdles, and resolve potential conflicts. Other possible tools to increase staff satisfaction include balanced staff schedules, anonymous evaluation, digital innovations, online learning modules, continuous staff feedback, team-building measures, bonus and benefits. It is the overarching goal to make working in a hospital more pleasant, prevent staff shortage, and increase the motivation of employees.

4.3 Treatment and intervention (T)

This area aggregates all dimensions that are directly related to the core processes in a hospital. Diagnostics are supported by machine learning methods (T2). Surgeries are scheduled automatically (T1) and conducted by collaborative surgery robots (T3). Telemedicine (T4) and precision medicine (T5) offer new possibilities for treatment and care. The numbers in brackets refer to the unique identifier for the dimension as depicted in Figure 5.

OR steering and resources management (T1). In the future, scarce and valuable resources such as OR capacity are allocated centrally by an IT-supported planning system using optimization methods and big data analytics. Under consideration of underlying constraints, preferences, and historic data, surgery schedules are optimized and dynamically refined in case of unexpected changes. Besides the OR, also the impact on downstream units such as intensive care units (ICUs), post-anesthesia care unit (PACU), and wards is considered. The system tracks progress in real-time and provides medical decision support based on previous surgeries. Overall, we see a higher utilization of scarce resources, better bed management, less waiting time, and reduced planning overhead costs. In the literature, several approaches have been proposed in order to develop tactical master surgery schedules [36], predict the duration of surgeries [97], predict the inpatient mortality [76], or estimate the required PACU time for different surgical procedures [34].

Advanced diagnostics (T2). Machine learning applications in diagnostics support radiologists in the future to identify conditions such as heart disease, skin cancer, or injuries earlier and more accurately. Medical imaging is one of the most effective and widespread applications of data science in health care. Images such as x-ray scans, MRIs, and mammographies are processed by machine learning algorithms to identify patterns and detect anomalies such as tumors and artery stenosis. Classification and recommendations are provided using big data analysis and comparison with benchmarks. Furthermore, also genome-based diagnostics is an helpful tool to optimize the accuracy of the diagnosis. Hospitals of the future collect and analyze all relevant data in one central competence center for diagnostics to leverage the full potential. Thanks to the advanced diagnostics, patient outcome is improved and workload of radiologists is reduced. Current state-of-the-art models show a convincing performance and sometimes even outperform human experts, e.g., diagnosis of irregular heart rhythms [77] and classification of malignant lesions [32]. Additional examples are analysis of urine [5], identification of neurodegenerative diseases such as Alzheimer [51], diagnosis of retinopathy [35], and prediction of chances to develop breast cancer [105].

Advanced robots in core processes (T3). Core processes in the future hospital such as intervention and care are performed or supported by advanced robots. In particular, collaborative robots (cobots) that are equipped with accurate sensors, high definition cameras, softer materials, and safety systems are well suited to work hand in hand with humans. Care robots can carry and lift patients, interact with patients [11], communicate lab results, and prepare patients for treatments such as MRI scans. Although they cannot replace human carers since they are unable to replicate the humanizing dimension, they support human nursers in their daily work. Surgery robots are characterized by a high accuracy, no trembling, high repeatability, and low error rate and hence improve patient safety and outcome particularly for complex interventions, e.g., removal of prostate tumor and eye surgeries. Nanobots are tiny robots that are inserted in the bloodstream and steered remotely. Since advanced robots support the staff with heavy and repetitive tasks, we also see higher staff satisfaction. For example, the Da Vinci surgery robot is already used by several hospitals, e.g., in urology and thoracic surgery [16]. Augsburg-based company German Bionic develops and produces exoskeletons that can be used to lift patients.

Telemedicine and remote care (T4). Advanced communication and visualization technologies enable real-time consultation and treatment of patients without physical presence of physicians and nurses. For some cases, human staff is not required at all and the patient communicates directly with virtual chat bots, e.g., to schedule an appointment or to conduct an initial consultation and collect required information before the face-to-face consultation with an physician. We distinguish between three major application areas, i.e., convenient in-home treatment, access to remote areas, and in the hospital. On demand in-home treatment via video calls is an convenient option for initial consultations and simple questions, chronic patients that need regular checks such as diabetics [59], and postcare after being released from the hospital, e.g., families with premature birth. In particular, preventive care and elderly care benefit significantly from telemedicine (also see P2). In addition to video appointments, also test samples can be collected at home and submitted to labs using drones, e.g., blood samples. Shin *et al.* [85] propose a mobile nucleic acid testing for sexually-transmitted infections such as Chlamydia. Telemedicine is very useful for remote areas with limited access to health services [64] or patients on the high seas. In the rural areas of Vorpommern-Greifswald in Germany, staff in the ambulance can consult a remote emergency doctor using mobile communications to transmit video signals and vital signs in real-time. In the hospital, ward rounds can be completed digitally without the physical presence of nurses. Also, surgeons can involve specialized experts during a treatment and ask them for their opinion on difficult cases [72] or let them conduct a surgery from remote by actively controlling the surgery robot.

Precision medicine (T5). Correlating individual patient characteristics with big data allows scientists to deliver precision medicine that is tailored to an individual patient's needs. For example, symptoms are be analyzed to narrow down a patient's diagnosis instead of applying the same set of tests to all patients. Furthermore, genomes are sequenced and analyzed to deliver more precise prescriptions and personalized treatments. In oncology, the commonly used chemotherapy that affects all cells are replaced by personalized therapy and medication that is tailored to the cells affected by the tumor. For this purpose, data on individual patient characteristics, national trends, and other data sources need to be collected, stored, and processed in a centralized manner, e.g., a national patient database. Recognizing the fact that *one size fits all* does not hold true for medication and care, precision medicine will have a significant impact on the delivery of health care in the future. Already in 2015, Obama stated: "You can match a blood transfusion to a blood type — that was an important discovery. What if matching a cancer cure to our genetic code was just as easy, just as standard? What if figuring out the right dose of medicine was as simple as taking our temperature?" [5]. Leveraging the potential of big data for precision medicine ensures the delivery of better patient care, increased patient safety, and reduced health care costs [76, 71, 49, 99]. Within the last 10 years, the costs and efforts to sequence a human genome have fallen significantly from USD 10 million and 10 months to USD 100 and 1 hour. Even today, we have initiatives such as

[5] https://obamawhitehouse.archives.gov/precision-medicine

the National Institutes of Health's 1000 Genome Project and companies like 23AndMe.

4.4 Logistics and supply (L)

This area aggregates all dimensions that are directly related to support processes in the hospital such as logistics, supply chain, catering, and cleaning. The warehouse in the hospital of the future is equipped with smart systems (L1) that enable automated replenishment (L2) and autonomous transportation (L4). Some support processes are conducted or supported by advanced robots (L3) and 3D printers (L5). The numbers in brackets refer to the unique identifier for the dimension as depicted in Figure 5.

Warehouse smartification (L1). Devices, resources, and products in the hospital of the future are integrated to cyber-physical systems using sensors and intelligent tools to gather information and trigger actions. For example, rack laser barriers and RFID gates are used in the warehouse to automatically scan and register in- and outcoming goods. Employees are supported by smart systems such as pick by vision and wearables for commissioning, encoding, and packaging tasks. This leads to efficient and transparent processes and reduce the logistic efforts. Some application areas are unit dose systems in the hospital pharmacy and sterile supply management, e.g., at UK Augsburg.

Automated replenishment (L2). Given the availability of cyber-physical systems (see L1), automated replenishment of materials and automated ordering can be realized in the hospital of the future. Sensors such as scale-triggered bins, cameras, or E-buttons detect if materials reach the critical stock level and automatically trigger the order for new materials. At the time of delivery, AGVs could pick up the new materials from the truck and deliver them within the hospital to the location where the critical stock level has been detected (see L4). Automated replenishment ensures process reliability, enables real-time stock control and optimization, and reduced repetitive manual tasks. Examples are boxes of bandages in the OR or mobile medication carts.

Advanced robots in support processes (L3). Support processes in the future hospital such as in warehousing, catering, and cleaning are performed or supported by advanced robots. Collaborative robots that are equipped with accurate sensors, high definition cameras, softer materials, and safety systems are used in cases where robots and humans directly work together. The mobile cobots are interconnected and can be easily programmed for new tasks. Possible application areas are unloading and commissioning of supply in the warehouse, bin-picking of drugs enabled by advanced cameras, and cleaning of floors and windows. Advanced robots automate monotonous, repetitive tasks and improve process efficiency, quality, and staff satisfaction. For example, some hospitals already use unit dose systems in the hospital pharmacy to sort different types of pharmaceuticals, commission them in small bags tailored to the patient's medication plan, and label the bag with name, room number, and time of usage. This use case could be combined with AGVs for automated delivery (L4) and smart patient wristbands for validation (P3).

Autonomous transportation (L4). In the future, transportation of materials, equipment, and persons are automated by usage of autonomous vehicles. Indoor AGVs are used to transport materials within the hospital, e.g., delivery of drugs and surgical instruments from the warehouse to the OR, management of empty beds, delivery of clean bed linen and meals. Pick-up from the AGVs can be realized with advanced robots (see L3). Autonomous wheelchairs or beds are used to transport patients within the hospital (see P5), e.g., from the check-in desk to the ward room. Autonomous ambulances and drones are used as connection with other health care facilities and areas outside of the hospital, e.g., transport the emergency doctor to the location of a car accident, collect transplants from another hospital, and deliver corpse to the morgue. Autonomous transportation automates monotonous tasks, reduce costs, and ensure higher speed and flexibility. In remote areas, saving time can even save lives. AGVs for hospital logistics and automated bed mover have been studied in the literature [9, 43]. The German ADAC currently tests the usage of manned volocopter for air rescue [6]. In Ghana, the government and

[6] `https://presse.adac.de/meldungen/adac-stiftung/luftrettung/einsatz-von-volocopter.`

US-based Zipline currently establish the world's largest medical drone network to deliver blood suppliers, vaccines, and life-saving pharmaceuticals such as snake venom antiserum to remote areas [18].

Additive manufacturing (L5). Additive manufacturing technologies such as rapid prototyping, selective laser sintering, laser beam melting, and binder jetting are already established in various industries, e.g., to manufacture complex parts in aerospace. Also in clinical practice, it is used for selected applications and becomes more relevant in the future [96, 21]. 3D printers using non-organic materials manufacture sterile supply items, personalized prostheses, and complex parts for surgery. Moreover, tissue engineering and bio printing offer entire new possibilities for organic parts, e.g., cardiac valves, hand bones, and maxillofacial surgery. Using additive manufacturing, hospitals are able to produce more complex parts with higher flexibility, personalize items for individual patients, and save costs and time. For example, the UK Eppendorf-Hamburg already operates a 3D printer to create models of hearts, bones, and lungs for training purposes. Scientists have manufactured a 3D-printed, patient-specific heart using a human-tissue-based hydrogel as ink which reduces the risk of rejection for the transplantation [68].

4.5 Management and organization (M)

This area aggregates all dimensions that are directly related to overarching managerial and organizational topics. The hospital of the future is continuously improved (M1), managed according to the principles of value-based health care (M2), and integrated with other stakeholders in the health care system (M3). The numbers in brackets refer to the unique identifier for the dimension as depicted in Figure 5.

Continuous improvements (M1). The guiding principle of continuous improvement is incorporated in the hospital of the future. Various technologies and methods such as sensors, feedback stations, and KPI monitoring are implemented with the goal to continuously improve processes and perform more value-adding activities. As a result,

`html`

we get more efficient processes, higher satisfaction of staff and patients, and reduced costs. Some possible applications are feedback stations for patients, definition and monitoring of KPIs for key processes, regular status updates for the hospital management, and monthly ideation competitions in which employees actively contribute to a better working environment and share their ideas (also see S5).

Value-based, patient-centered health care (M2).

While competition in today's health care industry is often based on wrong incentives, e.g., many and complex procedures, increasingly more providers start to distinguish themselves by delivering superior clinical outcomes. Outcomes-based competition has "the potential to improve the value delivered by the entire health system" since it focuses on "what really matters to patients and what ought to be the raison d'être of any health system: delivering high-quality care in a cost-efficient fashion" [23] what is referred to as value-based health care. According to Fung *et al.* [37], publicly release of performance data can stimulate the quality improvement activity in hospitals. Clawson *et al.* [23] distinguish between three levels with increasing complexity and responsibility, i.e., using standardized outcome metrics to improve clinical practice, linking reimbursement to outcomes, and managing the health outcomes for an entire population. Standardized and risk-adjusted outcome metrics are developed by organizations such as the International Society of Arthroplasty Registries and the International Consortium for Health Outcomes Measurement constituting a solid foundation for data analytics, e.g., for the Australian researchers that identified problems with DePuy's articular surface replacement (ASR) metal-on-metal implant in 2009 [24]. Reimbursement is already linked to outcome in some cases such as GlaxoSmithKline testing experimental pharmaceuticals for chronic obstructive pulmonary disease with the NHS in the UK and the Stockholm county council establishing OrthoChoice as reimbusement system for hip and knee arthroplasty in Sweden. According to Clawson *et al.* [23], some US-based single-provider integrated-delivery institutions such as Kaiser Permanente, Intermountain Healthcare, and the Geisinger Health System come closest to manage the entire population health since they are payer and provider, prioritize preventive care, and use only treatments with proven value. German prostate-cancer center Martini-Klinik continuously improves its performance by analyzing health outcomes data resulting in superior outcomes - which are significantly better than German average - and consequently to the highest patient volumes for radical prostatectomies worldwide.

Value chain integration (M3). The future health care system is composed of a dynamic, integrated network structure in which various providers coexist and cooperate, i.e., excellence centers for complicated interventions, day clinics for routine tasks, outsourcing of clinical services such as in radiology, remote treatment at home, gyms and external care solutions [30]. As part of this system, hospital departments are integrated within the hospital and with other stakeholders outside of the hospital such as GPs, labs, rehab facilities, suppliers, payers, morgues, and smart homes. They exchange information, coordinate health services, collaborate closely, discuss opinions, engage and evaluate each other. Traditionally separated departments within a hospital such as medical technology, IT, and building technology merge to one integrated center with shared data and clear responsibilities. While today's imaging data and lab results are stored and analyzed in various departments such as radiology, pathology, and clinical labs, in the future, one centralized diagnostic competence center is established leveraging the full potential of all data. Partnerships, strategic alliances, and close collaborations enable different providers to organize in flexible and distributed network structures, to specialize on core competences, and provide a holistic delivery of health services for the population (see M2). For example, while today's emergency rooms (ERs) are filled with non-emergency patients with minor issues, in the future, all patients are admitted at a central check-in area and referred to suitable providers such as GPs, day clinics, specialized clinics, or ERs. The value chain integration ensures better data availability, higher efficiency, reduced response time, higher patient safety, and better quality of care. For example, in some countries such as Finland, Canada, and Australia, various health providers share a common HIS. In Sweden, outcomes for individual types of intervention are analyzed and patients shifted to the hospital with higher survival rates, e.g., from Karolinska hospital to Danderyd hospital. Other applications include German IVENA system for interdisciplinary communication, Agfa's platform called EngageSuite to connect different stakeholders, or the partnership between Kaiser Permanente and Fresenius Medical Care to deliver high-quality care for renal-failure patients.

4.6 Data and control (D)

This area aggregates all dimensions that are directly related to usage of data which are key in the future. In the hospital of the future, data is gathered by smart sensors and processed on centralized hospital information systems (D1) allowing to analyze big data (D4), simulate and optimize processes (D5), track the location of devices and people (D3), and realize a paperless documentation. The numbers in brackets refer to the unique identifier for the dimension as depicted in Figure 5.

Centralized HIS and data warehouse (D1). Data in the hospital of the future is managed in a centralized hospital information system, also referred to as digital twin, and accessible via user-friendly interfaces. The centralized HIS combines systems such as radiology information system (RIS) and PACS and is connected to cloud-based software services and a storage solution for long-term archiving. The integration of additional information and services such as real-time tracking data create a digital representation of physical and non-physical elements in the hospital, e.g., buildings, rooms, patients, organs, and contracts. According to Kuhn [55], digital twins contain relevant information for existing and not-yet-existing objects and have the purpose to facilitate information exchange, realize virtual planning, and simulate properties of functional and physical nature. The system is not limited to individual hospitals, but share with various providers and other stakeholders in the health system. Structured digitization of all paper-based documents and gathering of additional data, e.g., by installing sensors, are prerequisites for the HIS. The highest priority is given to data security (see E2) to avoid cyber attacks and leakage of personal information. For example, advanced user right management, differential privacy [104], and knowledge-constrained access control can contribute to preserve privacy for patients' information [107]. The centralized HIS enables big data analysis, facilitate mobile accessibility, and create transparency. For example, the Charité in Germany has rolled out centralized data management for kidney diseases, in some countries such as Finland, Canada, and Australia, various health providers share a common HIS, and in the United States, the EHR usage has been increased significantly by posing penalties in line with the HITECH act.

Smart, paperless documentation (D2). As confirmed in our survey, most of a hospital's processes are not yet automated, a lot of documentation is still done on paper, and paperwork is very time-consuming, e.g., physicians spend roughly one-third of their time on paperwork. In the future, all documentation in the hospital is processed paperless and automated, e.g., appointments, doctor's letter, diagnostic findings, lab results, accounting, tax, care documentation, prescriptions, medication plans, internal status reports, and reports for payers and authorities. Digital documentation creates processes without media interruption and lays the foundation for further processing such as data analytics (D4). Virtual agents and chatbots can be applied to answer simple service queries and narrow down more complex questions. This results in less bureaucracy, less time spent on monotonous paperwork, and higher efficiency of processes. Moreover, it eliminates human errors such as unclean handwriting and lost notes [100]. Already today, medication plans are submitted in digital form to the hospital pharmacy, paper-based documents are digitized using natural language processing (NLP), physician notes are encoded to make them searchable, and audio systems are used to record information.

Real-time tracking (D3). In the hospital of the future, rooms, mobile devices, and even persons are equipped with technologies such as RFID and GPS to localize and track them in real-time. This allows for instance to find an available ultrasound device, localize the closest cardiologist, and get an overview of the location of nearby ambulances. Abkari *et al.* [1] study real-time locating techniques in a hospital environment. Real-time locating and tracking services reduce manual work, create transparency, and serve as foundation for big data analytics, simulation, and optimization. Possible applications are automated time recording of staff, real-time monitoring of patient flow, improved bed management, and route planning for ambulances and meal delivery. For example, ambulances are equipped with location-tracking devices to control traffic lights [44], patients are located using smart wearables or cameras to guide them through the hospital [54], and ORs are monitored to trigger cleaning services as soon as a patient leaves the OR.

Big data and analytics (D4). Big data and analytics seem to be made for health care given the fact that about one third of the world's stored data is gathered in this industry, e.g., electronic medical records, clinical trials, drug administration data, operations logs, insurance billing, regulatory compliance, and more recently also genomics data and real-time information from monitoring devices such as cameras, location-trackers, and wearables. Analyzing large amounts of data provides a insights and a deeper understanding of a hospital such that bottlenecks can be mitigated and processes optimized. We distinguish between the four stages of descriptive, diagnostic, predictive, and prescriptive analytics. Advanced methods such as machine learning, pattern recognition, and in particular deep learning [41, 58] are well suited to do so offering an almost indefinite variety of possible applications: predictive maintenance to prevent break-downs of machines, automated adjustment of lighting and HVAC to save energy, prediction of occupancy levels and readmission rates to make processes more efficient and provide better health care, and detection of anomalies in historical data to prevent fraud, wast, and abuse. Besides operational improvements, big data analytics is also essential for many other applications such as advanced diagnostics (T2), precision medicine (see T5), and preventive care (see P2). Big data and analytics show a variety of positive impacts such as reduced break-downs, improved stability, and tailored medicine and care. For example, IBM and Mayo Clinic processed 4.4 million patient records and obtained better diagnostic results, a Texas hospital analyzed data to identify high-risk cardiac patients resulting in a reduced readmission rate from 26 % to 21 %[7], and Bermejo *et al.* [14] developed a model to prevent early readmission of COPD patients.

Simulation and optimization (D5). Based on the structured data in the centralized HIS, processes in the hospital are modeled, simulated, and optimized in order to provide decision support for the hospital management. For example, the processes in an operating room can be simulated to identify bottlenecks, optimize the scheduling, and evaluate the impact of modifications. Moreover, the entire patient flow through the hospital can be simulated and updated with real-time data. Also, methods from the field of process mining can be applied [98]. Simulation and optimization make processes more efficient and save both time and costs. For example, extensive work has been done in the fields of case mix planning [46], ICU management [10], and various applications using Data Envelopment Analysis (DEA) [53].

4.7 Infrastructure (I)

This area aggregates all dimensions that are directly related to hospital buildings and equipment. Hospitals of the future are modular, integrated buildings (I1) equipped with standardized systems (I3) and running sustainable operations (I2) adapted to the care needed in the future, e.g., palliative care. The numbers in brackets refer to the unique identifier for the dimension as depicted in Figure 5.

Modular, integrated buildings (I1). Agile planning methods and modular designs are the leading principles when building the hospital of the future. Architects and construction engineers construct buildings that are flexible, resilient, and robust and ensure to stay compatible with a variety of systems that will be used in 20 years. The building consists of several "simple and slim" modules with an interchangeable design to allow for dynamic modifications, expansions, and repurposing in response to future developments. Shorter distances and travel times are enabled by clustering areas that belong together and limiting the height of buildings to 4-5 floors. Above all, future buildings are designed around patient needs instead of today's specialist departments, e.g., a dedicated center offering all health services for respiratory diseases, and promote the well-being of patients, e.g., a pleasant hotel feeling in admission area and wards. Examples are hybrid ORs and fully variable rooms that can be used for various levels of care intensity.

Sustainability (I2). Driven by increased environmental awareness, ecologically sustainable operations and efficient usage of energy, materials, and resources play a major role for the hospital of the future. Smart systems monitor, analyze, and control energy consumption in the building using sensors and big data analytics, e.g., HVAC (see I4). Where applicable, disposable and plastic items such as gloves are replaced with reusable items made from eco-friendly materials. Enhancements such as energy-efficient buildings, LED lighting, and improved waste management not only improve the hospital's image and increase satisfaction, but also save the planet's valuable resources and reduce costs.

Standardization and interoperability (I3). Systems, equipment, and interfaces in the hospital of the future are standardized and interoperable with each other. Hospitals avoid isolated solutions, compare different suppliers, and promote open-source software. Open application programming interfaces (APIs), standardized technology, and accessible standards create a new ecosystem [95]. The compatibility between different systems provides equal access to all suppliers. Lower entry barriers increase competition among the suppliers, e.g., innovative start-ups enter the market, giving hospitals the freedom to purchase each individual item from the respective best supplier and integrate all items by *plug-and-play* to a working system. For example, up-to-date patient entertainment can be connected as add-on to a smart bed independent of the supplier.

Hospital smartification (I4). Devices, resources, and products in the hospital of the future are integrated to cyber-physical systems using sensors and intelligent tools to gather information and trigger actions. The so-called smartification of buildings and equipment facilitates the realization of various use cases. For example, doors are equipped with access control systems using RFID or face recognition. Smart systems monitor, analyze, and control energy consumption in a building using sensors and big data analytics, e.g., HVAC (see I2). In a ward room, the patients can control windows, blinds, heating, light, and nurse call via voice recognition systems such as Amazon's Alexa. Smart beds provide the staff with useful information such as care and cleaning history. Visitors are supported by smart signage systems that guide them through the hospital. This leads to efficient and transparent processes, enable real-time monitoring, and increase satisfaction of staff and patients. For example, Wibbeling *et al.* [101] describe smart beds and the German start-up hospimatix upgrades ward rooms with sensors and control systems.

4.8 Evaluation of HoF dimensions

The proposed framework organizes the full range of technological, social, and organizational innovations that constitute the hospital of the future along 32 dimensions. We have asked the survey participants to evaluate those dimensions by importance and maturity level (on a scale from low (0) to high (4), respectively). Figure 6 depicts the resulting evaluation matrix where the colors indicate the areas and the numbers refer to the unique identifier of the dimensions as described above. The location of a dimension in the evalu-

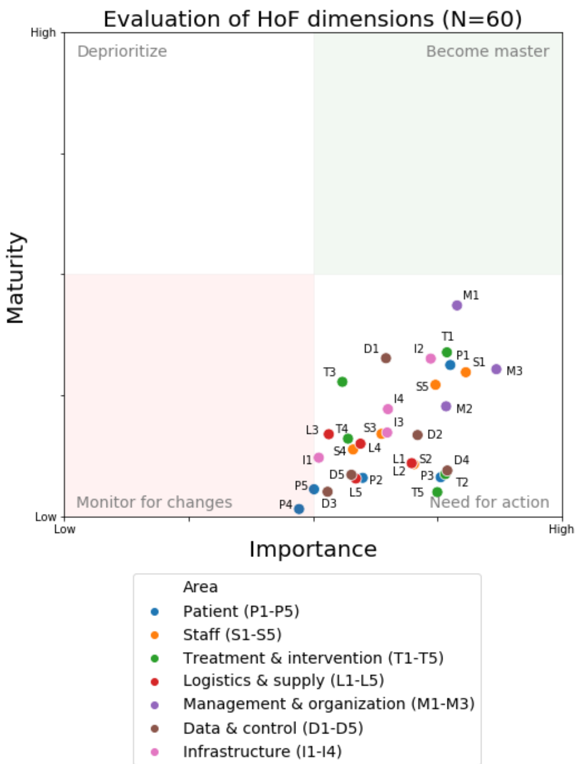

Figure 6: Evaluation of *HoF dimensions.*

ation matrix reflects the respective maturity and importance. In particular, the left side of the matrix indicates lower importance while the right side indicates higher importance. Analogously, the bottom half indicates lower maturity while the top half indicates higher maturity. Consequently, the following recommendations are obtained for the quadrants of the evaluation matrix:

1. Deprioritize Dimensions in the top left quadrant should be deprioritized to be able to focus the efforts on more relevant topics instead.

2. Monitor for changes Dimensions in the bottom left quadrant should be monitored closely to be able to react in case they become more relevant in the future.

3. Need for action Dimensions in the bottom right quadrant should be advanced with great efforts in order to build up adequate capabilities.

4. Become master Dimensions in the top right quadrant should be further strengthened and professionalized to become best-in-class.

The survey results depicted in Figure 6 reveal that almost all dimensions show need for action. Although the participants evaluated their importance as rather high, the hospitals' capabilities are rather limited. Value chain integration (M3), OR steering and resources management (T1), and smart staff scheduling (S1) were evaluated as most relevant. Patient participation (P4) and decentralized, autonomous patient flow (P5) were evaluated as least relevant. Continuous improvements (M1) show the highest maturity, while patient participation (P4) shows the lowest.

4.9 Enablers (E)

In order for the vision to become reality, hospitals need to address several prerequisites. As confirmed in our survey, most hospitals struggle to build momentum and do not feel well prepared (27 % completely or almost unprepared, 20 % only a little prepared). Furthermore, participants highlight that they need additional and qualified staff, a shift of mindset of the staff, i.e., acceptance, motivation, courage, and support, high quality training to master the transformation, sufficient funding, as well as a clear vision, strategy, roadmap, and governance structure. Consequently, we have distilled four enablers that are most critical for the implementation of use cases in any of the framework's 32 dimensions, i.e., employees & skills (E1), IT infrastructure & data security (E2), strategy & roadmap (E3), and people engagement & governance (E4). On the journey towards the hospital of the future, hospitals need to ensure first and foremost that all of those four enablers are addressed.

Employees & skills (E1) As confirmed in our survey, shortage of staff is one of the most critical challenges in today's hospitals. For a successful transition to the hospital of the future, hospitals need to ensure sufficient numbers of employees such as physicians, nurses, IT experts, and administrative staff. Only with sufficient, qualified, and motivated staff, there will be capacity to develop, implement, and professionalize the applications. Moreover, many innovations in the hospital of the future are driven by technology. A number of jobs in the future require technical competencies such as IT and medical technology, but also quality management, logistics, supply chain, and health care operations management. Given the rapidly changing environment, also social competencies such as learning capacity and problem solving become more important. Consequently, it is essential that employees stay up-to-date and human factors represent no limitation [94]. Hiring new employees is a challenging task given the limited financial resources and the so-called *fight for talents*, i.e., only 3 % of data scientists in the US work in health care. Besides hiring new employees, hospitals also need to develop an approach to training and qualifying existing staff in order to retain them in the long term.

IT infrastructure & data security (E2) IT infrastructure & data security are major challenges in most hospitals. Since many applications in the hospital of the future rely on data, a solid and reliable IT infrastructure is crucial for success. In particular, up-to-date hardware and software needs to be provided along the entire process from data gathering, i.e, sensors, cameras, and wearables, over connectivity, i.e., ubiquitous WLAN, to central data storage and analytics capabilities. Data needs to be organized, governance rules established, and compatibility ensured. Cloud services offer the possibility to outsource storage and computations to virtual machines. Data security is a major concern if it comes to health care data. In particular, patient-related information such as imagery, prescriptions, genomic data, or EHRs is very sensitive and often a prime target for cyber attacks. According to a recent study [13], almost two third of all German hospitals have already become victims of cyber attacks. Prominent incidences such as 2006's hacker attack to blackmail German Lukaskrankenhaus with ransomware [91] or 2019's global leak of more than 16 millions of PACS data [39] (also see [92]) vividly demonstrate the high risk and the far-reaching consequences. Consequently, data security in hospitals needs to be treated with utmost caution and ensured by a variety of measures such as advanced encryption, modern firewalls, separated network structures, a rigorous update policy, an external data security officer, a dedicated security operations center, and a set of policies, rules, and standards, e.g., an effective risk access control model [84]. Network structures and data exchanges should be continuously monitored with AI-powered anomaly detection algorithms (see D4). The new field of data ethics becomes more important, e.g., patients sovereignty over their data needs to be ensured. Patients at Stanford Medical School can decide whether to contribute their data to a research database which is stored on a cloud run by the US-based start-up Oasis Lab and secured with smart contract-software and blockchain technology. When training machine learning models and running big data analytics, data scientists should consider federated and privacy-preserving approaches such as differential privacy [104] providing a mathematical guarantee that it is impossible to infer from any outcome to any individual's data [29], i.e., Germany-based apheris AI.

Strategy & roadmap (E3) The strategy for implementing the hospital of the future must be anchored as a key element in the overall hospital strategy. Most hospitals lack a clear strategic vision and a holistic target picture which is essential to manage the transformation process in a structured manner and to engage the staff (see E4). Only a bold vision and strong commitment of the hospital management enable cultural change which is essential to change the system [3]. The target picture as well as the structured framework presented in this study are intended to serve as guidance for the development of a strategic vision. Section 5 provides some useful recommendations for the usage of the framework and the development of a roadmap.

People engagement & governance (E4) As confirmed in our survey, lack of acceptance and responsibility, missing motivation and courage, required time and efforts, and missing support of the management are among the most common obstacles. Some participants also expressed their concern of blindly relying on machines (also see [62]), lack of social contacts, inhuman medicine, and dismissals. Changing the culture of staff is essential for a successful transformation to the hospital of the future [3, 94]. Providing high quality training for the employees is not only important to close the gaps to the required capabilities (see E1), but also to motivate the employees, take away concerns and fears, and master the transformation and change process. While the entire hospital organization transforms from today's specialist departments to a more *patient-centered structure* (see I1), also for the realization of the vision, organizational structures and processes need to be put in place, e.g., a dedicated team that ensures the overall progress, clear responsibilities for each dimension, regular meetings and status updates, and tailored benefit structures. In particular, the management must lead with a clear vision (see E3), show courage, and shift from an authoritarian to a more advisory leadership style.

The evaluation matrix in Figure 7a depicts the relative importance and maturity for each of the four discussed key enablers (on a scale from low (0) to high (4), respectively). All four enablers were evaluated as important, but lack maturity. IT infrastructure & data security (E2) was rated most important while - next to employees & skills (E1) - also showing the highest level of maturity. For each of those two enablers, a more detailed evaluation is presented in Figure 7b and Figure 7c, respectively. In employees & skills in Figure 7b, we distinguish between technical (depicted in blue) and social (depicted in orange) competencies. Employees with IT skills were evaluated as most important,

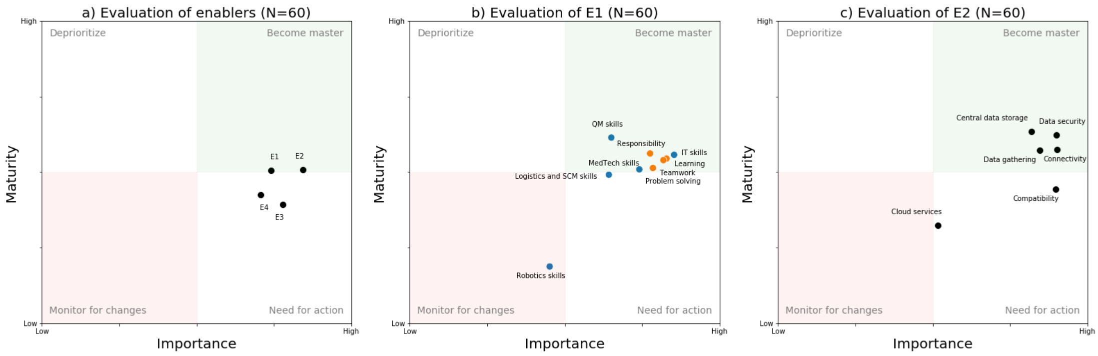

Figure 7: Evaluation by importance and maturity for (a) the four key enablers and deep-dives for (b) employees & skills as well as (c) IT infrastructure & data security.

quality management skills show the highest maturity, and robotics skills were evaluated with rather low importance and maturity. According to the survey, the gaps between required skills and existing capabilities are planned to be closed mostly via training of employees (48 % of participants) followed by hiring of new employees (23 %). Narrowed down to individual skills, hiring particularly applies for IT skills (43 %) and training is more suited for social competencies (on average 56 %), i.e., for learning capacity (71 %). Among the IT infrastructure & data security in Figure 7c, all topics - except for cloud services - were rated as highly important with data security & connectivity being the most important ones.

5 Deriving implications

In the previous section, we presented a conceptual framework for the hospital of the future. Given the pressing challenges and the fast-changing environment, hospitals need to move quickly to address all four key enablers, develop a long-term vision, and conduct many pilots. This will help hospitals to familiarize with the new concept, initiate a cultural change, attract qualified employees, build up strong partnerships, and gain first hands-on experiences.

For practitioners who wish to use our framework as guidance, we propose a four-step approach to get started. First, managers should involve key persons in their hospital, conduct interviews, and collect all relevant information to assess the status quo. The purpose of these initial investigations is to identify current pain points and key challenges and learn about already existing use cases, pilots, and planned projects. The experiences of all relevant groups such as management, physicians, nursery, and IT should be represented. Second, the current maturity level for each of the 32 *HoF* dimensions should be assessed. In order to do so, a dedicated questionnaire is processed in discussion rounds with the hospital management. Each dimension is divided into several aspects whose status needs to be assessed with one of the following scores: 0 (not started), 1 (piloted), 2 (applied), 3 (routine), or 4 (mastered). Filling out the entire questionnaire will result in an average maturity score between 0 and 4 for each *HoF* dimension. The overall maturity level of a hospital can be illustrated by indicating the score for each dimension in a radar chart as depicted in Figure 8. Augmenting the maturity level in the radar chart by a line

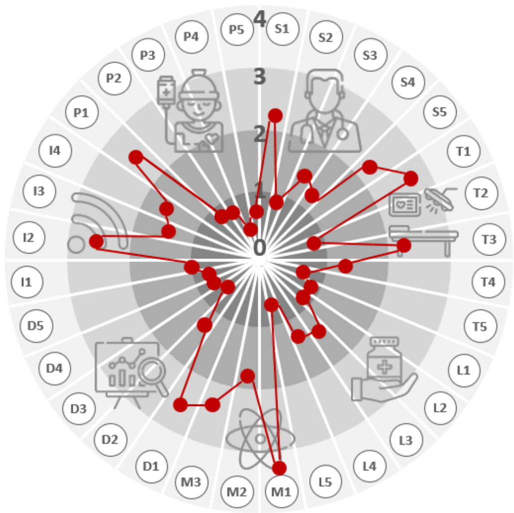

Figure 8: Radar chart to assess maturity of a hospital.

for the target level will help to identify areas for development and draw a target picture similar to the one in Figure 4. Since the starting point as well as the aspirations differ significantly between hospitals, also a hospital's target level is individual and needs to be decided on in management discussions. It is important to note that there is no need to achieve the highest score in every dimension. In fact, the target level should rather be tailored to the needs and aspirations of the individual hospital. Third, the derived radar chart might also be used to compare the maturity level with other hospitals. Benchmarking creates transparency, identifies best-in-class hospitals, and offers valuable insights to learn from each other. The maturity level depicted in Figure 8 is based on the results from the survey and serves as first indication for an industry benchmark. Fourth, based on the maturity assessment and the defined target level, hospital managers will be able to derive a tangible implementation roadmap. It is the merit of using the framework that the timeline, tasks, and responsibilities can be broken down to areas, dimensions, and individual use cases. This structure might also help to set up a governance structure. To ensure a successful transformation and leverage the full potential, a center for health care operations and analytics with link to the IT department, top management, and the individual medical departments should be established.

Moreover, the framework enables researchers to categorize their work into one of the 32 dimensions, develop structured literature overviews, and uncover research gaps. The four-step approach will be beneficial in order to initiate and coordinate joint research projects with hospitals, identify new opportunities, and facilitate communications between researchers and practitioners.

6 Conclusion

Mainly driven by big data and analytical power, health care is on the cusp of a new era which has implications for providers, payers, medical technology companies, and biopharma companies, as well as for patients and technology companies. This comes at the right time for hospitals, since they face a variety of organizational, medical, and financial challenges such as lack of staff and lack of funding that – similar to the smart factory – can be mitigated by these innovations. The hospital of the future in 2040 will be characterized by digitization, integration, automation, and personalization in both support processes as well as core processes using cyber-physical systems and data analytics. Not few hospitals have already started to discuss and implement first use cases such as EHR or surgery robots, however, most hospitals still lack a holistic strategy on the way to become a hospital of the future. We found that hospital managers have high ambitions to enhance their hospital and expect significant benefits, but they are struggling to build momentum. To make the hospital of the future more tangible, we developed the conceptual *HoF framework* structured in 32 dimensions along seven areas and based on four enablers as foundation. The framework is based on latest scientific findings as well as an extensive survey representing the perspective of hospital managers, physicians, nursery, and IT leaders.

This study is intended as guidance for hospital managers to navigate the transformation in a structured manner. The high-level target picture facilitates strategic discussions, the *HoF framework* helps to structure the use cases, and each *HoF dimension* is described in detail and enhanced with examples from scientific literature and practice. Furthermore, we proposed a four-step approach and a radar chart as supporting tool for self-assessment and benchmarking. To make the most of the transformation to the new era, hospitals must prepare themselves, develop a clear strategy, and secure the right talents and data.

The presented *HoF* framework is based on scientific literature as well as insights obtained from our survey. Going forward, the framework needs to be tested and detailed in practice. In our future work, we will conduct several field studies with selected hospitals to assess their current maturity and derive a roadmap tailored to their needs. A global survey will be beneficial to underpin our findings beyond the scope on Germany. The experience gained from those field studies and the global survey will enable us to further refine the *HoF framework*. This would also allow us to quantify required investments, running costs, and expected upsides for the concept. We plan to set up an extensive benchmarking database which can be accessed online and provides insights per region, hospital type, and hospital size. Eventually, it is our intention to encourage further research and practical studies in this area to support hospitals on their transition towards the hospital of the future.

Acknowledgments

We want to thank Universitätsklinikum Augsburg's Dr. Thomas Koperna and Universitätsmedizin Göttingen's Anne-Katrin Kirchner for their valuable input and helpful discussions during the development of the framework. A special thank you to Patrick Pflaum for his thesis which paved the way for this study. We thank Almut Rummel for drawing the target picture (Figure 4). Furthermore, we are grateful for the support of Günther Gartner as well as Harald Mahr enabling us to conduct and promote the survey. This study would not have been possible without all the participants of the survey to whom we would like to express our sincere thanks.

References

[1] S. E. Abkari, A. Jilbab, and J. E. Mhamdi. Real time positioning over WSN and RFID network integration. In *2018 4th International Conference on Advanced Technologies for Signal and Image Processing (ATSIP)*, pages 1–5, Mar. 2018. doi: 10.1109/ATSIP.2018.8364443.

[2] A. Aboshiha, R. Gallagher, and L. Gargan. Chasing Value as AI Transforms Health Care. *Boston Consulting Group*, 2019.

[3] P. Afferni, M. Merone, and P. Soda. Hospital 4.0 and Its Innovation in Methodologies and Technologies. In *2018 IEEE 31st International Symposium on Computer-Based Medical Systems (CBMS)*, pages 333–338, June 2018. doi: 10.1109/CBMS.2018.00065.

[4] S. Agnihothri, L. Cui, M. Delasay, and B. Rajan. The value of mHealth for managing chronic conditions. *Health Care Management Science*, Oct. 2018. ISSN 1572-9389. doi: 10.1007/s10729-018-9458-2.

[5] A. I. Ahmed, H. Baz, and S. Lotfy. Urinalysis: The Automated Versus Manual Techniques; Is It Time To Change? *Clinical Laboratory*, 62(1-2):49–56, 2016. ISSN 1433-6510.

[6] J. Alves, E. Seabra, I. Caetano, J. Gonçalves, J. Serra, M. Martins, and C. P. Santos. Considerations and Mechanical Modifications on a Smart Walker. In *2016 International Conference on Autonomous Robot Systems and Competitions (ICARSC)*, pages 247–252, May 2016. doi: 10.1109/ICARSC.2016.30.

[7] R. Amarasingham, P. C. Patel, K. Toto, L. L. Nelson, T. S. Swanson, B. J. Moore, B. Xie, S. Zhang, K. S. Alvarez, Y. Ma, M. H. Drazner, U. Kollipara, and E. A. Halm. Allocating scarce resources in real-time to reduce heart failure readmissions:

A prospective, controlled study. *BMJ Quality & Safety*, 22(12):998–1005, Dec. 2013. ISSN 2044-5415, 2044-5423. doi: 10.1136/bmjqs-2013-001901.

[8] D. A. Asamoah, R. Sharda, H. N. Rude, and D. Doran. RFID-based information visibility for hospital operations: Exploring its positive effects using discrete event simulation. *Health Care Management Science*, 21(3):305–316, Sept. 2018. ISSN 1572-9389. doi: 10.1007/s10729-016-9386-y.

[9] J. Bačík, F. Ďurovský, M. Biroš, K. Kyslan, D. Perduková, and S. Padmanaban. Pathfinder–Development of Automated Guided Vehicle for Hospital Logistics. *IEEE Access*, 5:26892–26900, 2017. doi: 10.1109/ACCESS.2017.2767899.

[10] J. Bai, A. Fügener, J. Schoenfelder, and J. O. Brunner. Operations research in intensive care unit management: A literature review. *Health Care Management Science*, 21(1):1–24, Mar. 2018. ISSN 1572-9389. doi: 10.1007/s10729-016-9375-1.

[11] V. Banthia, Y. Maddahi, M. May, D. Blakley, Zixin Chang, A. Gbur, Chi Tu, and N. Sepehri. Development of a graphical user interface for a socially interactive robot: A case study evaluation. In *2016 IEEE 7th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON)*, pages 1–8, Oct. 2016. doi: 10.1109/IEMCON.2016.7746294.

[12] A. Bauer, O. Hüfner, and T. Ruemping. Why you need to know about IBM's Elderly Care solution. *IBM Insurance Industry Blog*, 2018.

[13] R. Berger. Roland Berger Krankenhausstudie 2017. 2017.

[14] P. Bermejo, J. A. Gámez, J. M. Puerta, M. A. Esquivias, and P. J. Tárraga. Construction of a Semi-Naive Model to Predict Early Readmission of COPD Patients by Using Quality Care Information. *2016 IEEE 16th International Conference on Data Mining Workshops (ICDMW)*, pages 233–240, 2016. doi: 10.1109/icdmw.2016.0040.

[15] D. M. Berwick and A. D. Hackbarth. Eliminating waste in US health care. *Journal of the American Medical Association*, 307(14):1513–1516, 2012.

[16] J. Bodner, H. Wykypiel, G. Wetscher, and T. Schmid. First experiences with the da Vinci™ operating robot in thoracic surgery. *European Journal of Cardio-Thoracic Surgery*, 25(5):844–851, May 2004. ISSN 1010-7940. doi: 10.1016/j.ejcts.2004.02.001.

[17] Boston Consulting Group. GKV 4.0 - Trend-Monitor, 2017.

[18] J. Bright. Drone delivery startup Zipline launches UAV medical program in Ghana. *TechCrunch*, Apr. 2019.

[19] Centers for Medicare & Medicaid Services. National Health Expenditures 2016, 2016.

[20] J. H. Chen and S. M. Asch. Machine Learning and Prediction in Medicine — Beyond the Peak of Inflated Expectations. *The New England journal of medicine*, 376(26):2507–2509, June 2017. ISSN 0028-4793. doi: 10.1056/NEJMp1702071.

[21] J. K. C. Chen and H. T. T. Do. Perspective of the 3D Printing Technology Applied on Medical Resource Integration and Service Innovation Business Model. In *2017 Portland International Conference on Management of Engineering and Technology (PICMET)*, pages 1–11, July 2017. doi: 10.23919/PICMET.2017.8125298.

[22] A. Chhatlani, A. Dadlani, M. Gidwani, M. Keswani, and P. Kanade. Portable Medical Records Using Internet of Things for Medical Devices. In *2016 8th International Conference on Computational Intelligence and Communication Networks (CICN)*, pages 446–451, Dec. 2016. doi: 10.1109/CICN.2016.93.

[23] J. Clawson, P. Lawyer, C. Schweizer, and S. Larsson. Competing on Outcomes: Winning Strategies for Value-Based Health Care. Technical report, BCG, 2014.

[24] D. Cohen. Out of joint: The story of the ASR. *BMJ*, 342(may13 2):d2905–d2905, May 2011. ISSN 1756-1833. doi: 10.1136/bmj.d2905.

[25] L. Davidson. Data Science in Healthcare: How It Improves Care, Mar. 2019.

[26] C. de Meijer, B. Wouterse, J. Polder, and M. Koopmanschap. The effect of population aging on health expenditure growth: A critical review. *European Journal of Ageing*, 10(4):353–361, Dec. 2013. ISSN 1613-9372. doi: 10.1007/s10433-013-0280-x.

[27] T. Degner. Healthcare AI & Machine Learning: 4.5 Things To Know, 2018.

[28] S. Dua, U. R. Acharya, and P. Dua, editors. *Machine Learning in Healthcare Informatics*. Intelligent Systems Reference Library. Springer-Verlag, Berlin Heidelberg, 2014. ISBN 978-3-642-40016-2.

[29] C. Dwork. The Promise of Differential Privacy: A Tutorial on Algorithmic Techniques. In *2011 IEEE 52nd Annual Symposium on Foundations of Computer Science*, pages 1–2, Palm Springs, CA, USA, Oct. 2011. IEEE. ISBN 978-0-7695-4571-4 978-1-4577-1843-4. doi: 10.1109/FOCS.2011.88.

[30] S. Eichhorst, J. Jenkins, and N. Stern. Das Krankenhaus der Zukunft. Technical report, McKinsey & Company, 2016.

[31] J. Elliman, M. Loizou, and F. Loizides. Virtual Reality Simulation Training for Student Nurse Education. In *2016 8th International Conference on Games and Virtual Worlds for Serious Applications (VS-GAMES)*, pages 1–2, Sept. 2016. doi: 10.1109/VS-GAMES.2016.7590377.

[32] A. Esteva, B. Kuprel, R. A. Novoa, J. Ko, S. M. Swetter, H. M. Blau, and S. Thrun. Dermatologist-level classification of skin cancer with deep neural networks. *Nature*, 542(7639):115–118, Feb. 2017. ISSN 1476-4687. doi: 10.1038/nature21056.

[33] Eurostat. Healthcare expenditure statistics, 2018.

[34] M. Fairley, D. Scheinker, and M. L. Brandeau. Improving the efficiency of the operating room environment with an optimization and machine learning model. *Health Care Management Science*, 22(4):756–767, Dec. 2019. ISSN 1572-9389. doi: 10.1007/s10729-018-9457-3.

[35] J. D. Fauw, J. R. Ledsam, B. Romera-Paredes, S. Nikolov, N. Tomasev, S. Blackwell, H. Askham, X. Glorot, B. O'Donoghue, D. Visentin, G. van den Driessche, B. Lakshminarayanan, C. Meyer, F. Mackinder, S. Bouton, K. Ayoub, R. Chopra, D. King, A. Karthikesalingam, C. O. Hughes, R. Raine, J. Hughes, D. A. Sim, C. Egan, A. Tufail, H. Montgomery, D. Hassabis, G. Rees, T. Back, P. T. Khaw, M. Suleyman, J. Cornebise, P. A. Keane, and O. Ronneberger. Clinically applicable deep learning for diagnosis and referral in retinal disease. *Nature Medicine*, 24(9): 1342–1350, Sept. 2018. ISSN 1546-170X. doi: 10.1038/s41591-018-0107-6.

[36] A. Fügener, G. M. Edenharter, P. Kiefer, U. Mayr, J. Schiele, F. Steiner, R. Kolisch, and M. Blobner. Improving Intensive Care Unit and Ward Utilization by Adapting Master Surgery Schedules. *A&A Practice*, 6(6):172, Mar. 2016. ISSN 2575-3126. doi: 10.1213/XAA.0000000000000247.

[37] C. H. Fung, Y.-W. Lim, S. Mattke, C. Damberg, and P. G. Shekelle. Systematic review: The evidence that publishing patient care performance data improves quality of care. *Annals of Internal Medicine*, 148(2):111–123, Jan. 2008. ISSN 1539-3704. doi: 10.7326/0003-4819-148-2-200801150-00006.

[38] A. Gbegnon, J. Monestina, and J. W. Cromwell. Machine Learning Algorithm for Accurate, Automated, Real-Time Prediction of Surgical Site Infections using EHR Data. *Journal of Surgical Research*, 186(2):527, Feb. 2014. ISSN 0022-4804, 1095-8673. doi: 10.1016/j.jss.2013.11.380.

[39] J. Gillum, J. Kao, and J. Larson. Millions of Americans' Medical Images and Data Are Available on the Internet. Anyone Can Take a

Peek. https://www.propublica.org/article/millions-of-americans-medical-images-and-data-are-available-on-the-internet, Sept. 2019.

[40] H. Gimpel and J. Übelhör. Logistik: Transparenz, Sicherheit und Effizienz. *Management & Krankenhaus*, 37, 2018.

[41] I. Goodfellow, Y. Bengio, and A. Courville. *Deep Learning*. The MIT Press, 2016. ISBN 978-0-262-03561-3.

[42] E. Gultepe, J. P. Green, H. Nguyen, J. Adams, T. Albertson, and I. Tagkopoulos. From vital signs to clinical outcomes for patients with sepsis: A machine learning basis for a clinical decision support system. *Journal of the American Medical Informatics Association*, 21(2):315–325, Mar. 2014. ISSN 1067-5027. doi: 10.1136/amiajnl-2013-001815.

[43] Z. Guo, X. Xiao, and H. Yu. Design and Evaluation of a Motorized Robotic Bed Mover With Omnidirectional Mobility for Patient Transportation. *IEEE Journal of Biomedical and Health Informatics*, 22(6):1775–1785, Nov. 2018. doi: 10.1109/JBHI.2018.2849344.

[44] S. Hebbar, P. Pattar, and V. Golla. A Mobile ZigBee Module in a Traffic Control System. *IEEE Potentials*, 35(1):19–23, Jan. 2016. doi: 10.1109/MPOT.2014.2341331.

[45] R. Hipp, C. Schlude, and N. Göller. Position paper - Healthcare of the future: The digital revolution of the healthcare sector. Technical report, Porsche Consulting, 2017.

[46] S. Hof, A. Fügener, J. Schoenfelder, and J. O. Brunner. Case mix planning in hospitals: A review and future agenda. *Health Care Management Science*, 20(2): 207–220, June 2017. ISSN 1572-9389. doi: 10.1007/s10729-015-9342-2.

[47] M. Hosseini. What will the future look like under Industry 4.0 and digital transformation in the healthcare space? Technical report, Roland Berger, 2015.

[48] T. Jain, E. B. Schwarz, and A. Mehrotra. A Study of Telecontraception. *New England Journal of Medicine*, 381(13):1287–1288, Sept. 2019. ISSN 0028-4793. doi: 10.1056/NEJMc1907545.

[49] J. L. Jameson and D. L. Longo. Precision medicine–personalized, problematic, and promising. *The New England Journal of Medicine*, 372(23):2229–2234, June 2015. ISSN 1533-4406. doi: 10.1056/NEJMsb1503104.

[50] D. Jha, A. Frye, and J. Schlimgen. Evaluating variables of patient experience and the correlation with design. *Patient Experience Journal*, 4(1):33–45, Apr. 2017. ISSN 2372-0247. doi: 10.35680/2372-0247.1176.

[51] A. Khazaee, A. Ebrahimzadeh, and A. Babajani-Feremi. Identifying patients with Alzheimer's disease using resting-state fMRI and graph theory. *Clinical Neurophysiology: Official Journal of the International Federation of Clinical Neurophysiology*, 126(11):2132–2141, Nov. 2015. ISSN 1872-8952. doi: 10.1016/j.clinph.2015.02.060.

[52] J.-C. Kischkewitz. Der weite Weg zum Krankenhaus 4.0. *Krankenhaus Technik Management*, 2016.

[53] S. Kohl, J. Schoenfelder, A. Fügener, and J. O. Brunner. The use of Data Envelopment Analysis (DEA) in healthcare with a focus on hospitals. *Health Care Management Science*, 22(2):245–286, June 2019. ISSN 1572-9389. doi: 10.1007/s10729-018-9436-8.

[54] W. Ku, Y. Liou, C. Yang, M. Lee, C. Lin, W. Lin, and T. Tsai. The study of the indoor positioning based on smart wearable technologies. In *2016 IEEE International Conference on Signal and Image Processing (ICSIP)*, pages 646–649, Aug. 2016. doi: 10.1109/SIPROCESS.2016.7888342.

[55] T. Kuhn. Digitaler Zwilling. *Informatik-Spektrum*, 40(5):440–444, Oct. 2017. ISSN 1432-122X. doi: 10.1007/s00287-017-1061-2.

[56] D. Küpper, K. Kuhlmann, S. Köcher, T. Dauner, and P. Burggräf. The factory of the future. Technical report, Boston Consulting Group, 2016.

[57] T. Laschitz. Neue Wege in der Aus- und Weiterbildung: Erfolgreiche Etablierung der eRef in der Niederösterreichischen Landeskliniken-Holding. https://www.kma-online.de/aktuelles/it-digital-health/detail/erfolgreiche-etablierung-der-eref-in-der-niederoesterreichischen-landeskliniken-holding-a-40257, 2019.

[58] Y. LeCun, Y. Bengio, and G. Hinton. Deep learning. *Nature*, 521(7553):436–444, May 2015. ISSN 1476-4687. doi: 10.1038/nature14539.

[59] D. Y. Lee, J. Park, D. Choi, H.-Y. Ahn, S.-W. Park, and C.-Y. Park. The effectiveness, reproducibility, and durability of tailored mobile coaching on diabetes management in policyholders: A randomized, controlled, open-label study. *Scientific Reports*, 8(1):3642, Feb. 2018. ISSN 2045-2322. doi: 10.1038/s41598-018-22034-0.

[60] J. Lee, B. Bagheri, and H.-A. Kao. A Cyber-Physical Systems architecture for Industry 4.0-based manufacturing systems. *Manufacturing Letters*, 3:18–23, Jan. 2015. ISSN 2213-8463. doi: 10.1016/j.mfglet.2014.12.001.

[61] K. J. Leonard. The Role of Patients in Designing Health Information Systems: The Case of Applying Simulation Techniques to Design an Electronic Patient Record (EPR) Interface. *Health Care Management Science*, 7(4):275–284, Nov. 2004. ISSN 1572-9389. doi: 10.1007/s10729-004-7536-0.

[62] L. McDonald, S. V. Ramagopalan, A. P. Cox, and M. Oguz. Unintended consequences of machine learning in medicine? *F1000Research*, 6:1707, 2017. ISSN 2046-1402. doi: 10.12688/f1000research.12693.1.

[63] K. McManus, E. K. Mallory, R. L. Goldfeder, W. A. Haynes, and J. D. Tatum. Mining Twitter Data to Improve Detection of Schizophrenia. *AMIA Joint Summits on Translational Science proceedings. AMIA Joint Summits on Translational Science*, 2015:122–126, 2015. ISSN 2153-4063.

[64] S. P. Mohanty, U. Choppali, and E. Kougianos. Everything you wanted to know about smart cities: The Internet of things is the backbone. *IEEE Consumer Electronics Magazine*, 5(3):60–70, July 2016. doi: 10.1109/MCE.2016.2556879.

[65] V. E. Morgan. Gesundheit 4.0. *New Business*, 2018.

[66] P. Natarajan, J. C. Frenzel, and D. H. Smaltz. *Demystifying Big Data and Machine Learning for Healthcare*. CRC Press, Taylor & Francis Group, 2017. ISBN 978-1-138-03263-7.

[67] N. Neumann. Digitalisierung mit Augenmaß. https://www.cerner.com/de/de/blog/digitalisierung-mit-augenmass, 2019.

[68] N. Noor, A. Shapira, R. Edri, I. Gal, L. Wertheim, and T. Dvir. 3D Printing of Personalized Thick and Perfusable Cardiac Patches and Hearts. *Advanced Science*, 6(11):1900344, 2019. ISSN 2198-3844. doi: 10.1002/advs.201900344.

[69] OECD publishing. Society at a glance 2011. OECD social indicators, 2011.

[70] OECD publishing. Health at a glance 2017: OECD indicators, 2017.

[71] R. B. Parikh, J. S. Schwartz, and A. S. Navathe. Beyond Genes and Molecules - A Precision Delivery Initiative for Precision Medicine. *The New England Journal of Medicine*, 376(17):1609–1612, Apr. 2017. ISSN 1533-4406. doi: 10.1056/NEJMp1613224.

[72] L. E. A. Pico, O. R. Cuenca, D. J. S. Alvarez, and P. A. P. Salgado. Knowledge management model for teleconsulting in telemedicine. *Studies in Health Technology and Informatics*, 137:130–140, 2008. ISSN 0926-9630.

[73] E. Pinheiro, O. Postolache, and P. Girão. Theory and Developments in an Unobtrusive Cardiovascular System Representation: Ballistocardiography. *The Open Biomedical Engineering Journal*, 4:201–216, Oct. 2010. ISSN 1874-1207. doi: 10.2174/1874120701004010201.

[74] E. Porter. A world of rising health care costs. *The New York Times*, 2013.

[75] G. Prouski, M. Jafari, and H. Zarrabi. Internet of Things in Eye Diseases, Introducing a New Smart Eyeglasses Designed for Probable Dangerous Pressure Changes in Human Eyes. In *2017 International Conference on Computer and Applications (ICCA)*, pages 364–368, Sept. 2017. doi: 10.1109/COMAPP.2017.8079762.

[76] A. Rajkomar, E. Oren, K. Chen, A. M. Dai, N. Hajaj, M. Hardt, P. J. Liu, X. Liu, J. Marcus, M. Sun, P. Sundberg, H. Yee, K. Zhang, Y. Zhang, G. Flores, G. E. Duggan, J. Irvine, Q. Le, K. Litsch, A. Mossin, J. Tansuwan, D. Wang, J. Wexler, J. Wilson, D. Ludwig, S. L. Volchenboum, K. Chou, M. Pearson, S. Madabushi, N. H. Shah, A. J. Butte, M. D. Howell, C. Cui, G. S. Corrado, and J. Dean. Scalable and accurate deep learning with electronic health records. *npj Digital Medicine*, 1 (1):1–10, May 2018. ISSN 2398-6352. doi: 10.1038/s41746-018-0029-1.

[77] P. Rajpurkar, A. Y. Hannun, M. Haghpanahi, C. Bourn, and A. Y. Ng. Cardiologist-Level Arrhythmia Detection with Convolutional Neural Networks. *arXiv:1707.01836 [cs]*, July 2017.

[78] C. Ramos Hernández, M. Núñez Fernández, A. Pallares Sanmartín, C. Mouronte Roibas, L. Cerdeira Domínguez, M. I. Botana Rial, N. Blanco Cid, and A. Fernández Villar. Validation of the portable Air-Smart Spirometer. *PloS One*, 13(2):e0192789, 2018. ISSN 1932-6203. doi: 10.1371/journal.pone.0192789.

[79] A. G. Reece and C. M. Danforth. Instagram photos reveal predictive markers of depression. *EPJ Data Science*, 6(1):15, Aug. 2017. ISSN 2193-1127. doi: 10.1140/epjds/s13688-017-0110-z.

[80] J. Ribera, G. Antoja, M. Rosenmöller, and P. Borrás. Hospital of the Future - A new role for leading hospitals in Europe. Technical report, IESE Center for Research in Healthcare Innovation Management, Accenture, 2016.

[81] S. Saxena, N. Holobinko, C. Liu, and A. Porwal. Preparing for a New Kind of Shakeup in Health Care. *Boston Consulting Group*, 2019.

[82] J. Schiele, F. Rabe, M. Schmitt, M. Glaser, F. Häring, J. O. Brunner, B. Bauer, B. Schuller, C. Traidl-Hoffmann, and A. Damialis. Automated Classification of Airborne Pollen using Neural Networks. In *41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 4474–4478, July 2019. doi: 10.1109/EMBC.2019.8856910.

[83] M. Schwinn. Das vernetzte Krankenhaus. *Süddeutsche Zeitung*, 2019.

[84] M. Shi, R. Jiang, X. Hu, and J. Shang. A privacy protection method for health care big data management based on risk access control. *Health Care Management Science*, July 2019. ISSN 1572-9389. doi: 10.1007/s10729-019-09490-4.

[85] D. J. Shin, P. Athamanolap, L. Chen, J. Hardick, M. Lewis, Y. H. Hsieh, R. E. Rothman, C. A. Gaydos, and T. H. Wang. Mobile nucleic acid amplification testing (mobiNAAT) for Chlamydia trachomatis screening in hospital emergency department settings. *Scientific Reports*, 7(1):1–10, July 2017. ISSN 2045-2322. doi: 10.1038/s41598-017-04781-8.

[86] B. G. Silverman, J. Holmes, S. Kimmel, C. Branas, D. Ivins, R. Weaver, and Y. Chen. Modeling Emotion and Behavior in Animated Personas to Facilitate Human Behavior Change: The Case of the HEART-SENSE Game. *Health Care Management Science*, 4(3):213–228, Sept. 2001. ISSN 1572-9389. doi: 10.1023/A:1011448916375.

[87] A. Solanas, C. Patsakis, M. Conti, I. S. Vlachos, V. Ramos, F. Falcone, O. Postolache, P. A. Perez-martinez, R. D. Pietro, D. N. Perrea, and A. Martinez-Balleste.

Smart health: A context-aware health paradigm within smart cities. *IEEE Communications Magazine*, 52(8):74–81, Aug. 2014. doi: 10.1109/MCOM.2014.6871673.

[88] P. Sonnier. *The Fourth Wave: Digital Health*. Independently published, Nov. 2017. ISBN 978-1-976791-55-0.

[89] A. S. Sprouse-Blum, G. Smith, D. Sugai, and F. D. Parsa. Understanding Endorphins and Their Importance in Pain Management. *Hawaii Medical Journal*, 69(3): 70–71, Mar. 2010. ISSN 0017-8594.

[90] M. Srinivas, P. Durgaprasadarao, and V. N. P. Raj. Intelligent medicine box for medication management using IoT. In *2018 2nd International Conference on Inventive Systems and Control (ICISC)*, pages 32–34, Jan. 2018. doi: 10.1109/ICISC.2018.8399097.

[91] S. Steffen. Hackers hold German hospital data hostage. https://www.dw.com/en/hackers-hold-german-hospital-data-hostage/a-19076030, Feb. 2016.

[92] M. Stites and O. S. Pianykh. How Secure Is Your Radiology Department? Mapping Digital Radiology Adoption and Security Worldwide. *American Journal of Roentgenology*, 206(4):797–804, Mar. 2016. ISSN 0361-803X. doi: 10.2214/AJR.15.15283.

[93] The World Bank Group. Fertility rate, total (births per woman) | Data. https://data.worldbank.org/, 2016.

[94] H. Thimbleby. Technology and the Future of Healthcare. *Journal of Public Health Research*, 2(3), Dec. 2013. ISSN 2279-9028. doi: 10.4081/jphr.2013.e28.

[95] C. Thuemmler and C. Bai, editors. *Health 4.0: How Virtualization and Big Data Are Revolutionizing Healthcare*. Springer International Publishing, 2017. ISBN 978-3-319-47616-2.

[96] J. Tuomi, K.-S. Paloheimo, J. Vehviläinen, R. Björkstrand, M. Salmi, E. Huotilainen, R. Kontio, S. Rouse, I. Gibson, and A. A. Mäkitie. A novel classification and online platform for planning and documentation of medical applications of additive manufacturing. *Surgical Innovation*, 21(6):553–559, Dec. 2014. ISSN 1553-3514. doi: 10.1177/1553350614524838.

[97] J. P. Tuwatananurak, S. Zadeh, X. Xu, J. A. Vacanti, W. R. Fulton, J. M. Ehrenfeld, and R. D. Urman. Machine Learning Can Improve Estimation of Surgical Case Duration: A Pilot Study. *Journal of Medical Systems*, 43(3):44, Jan. 2019. ISSN 1573-689X. doi: 10.1007/s10916-019-1160-5.

[98] W. M. P. van der Aalst. *Process Mining: Discovery, Conformance and Enhancement of Business Processes*. Springer, Berlin, Heidelberg, 2011. ISBN 978-3-642-19345-3. doi: 10.1007/978-3-642-19345-3_1.

[99] F. R. Vogenberg, C. Isaacson Barash, and M. Pursel. Personalized medicine: Part 1: Evolution and development into theranostics. *P & T: A Peer-Reviewed Journal for Formulary Management*, 35(10):560–576, Oct. 2010. ISSN 1052-1372.

[100] WHO. WHO calls for urgent action to reduce patient harm in healthcare. https://www.who.int/news-room/detail/13-09-2019-who-calls-for-urgent-action-to-reduce-patient-harm-in-healthcare, 2019.

[101] S. Wibbeling and M. Raida. Krankenhaus 4.0 - Digitalisierung im Krankenhaus unterstützt durch Smart Devices - Positionspapier. Technical report, Fraunhofer-Institut für Materialfluss und Logistik: Health Care Logistics, 2019.

[102] S. Wibbeling, M. Hintze, and W. Deiters. Krankenhaus 4.0 - Industrie 4.0 im Gesundheitswesen - Positionspapier. Technical report, Fraunhofer-Innovationszentrum für Logistik und IT, 2018.

[103] S. Wibbeling, S. Laciok, and M. Hintze. Humanzentriertes digitales Krankenhaus

- Positionspapier. Technical report, Fraunhofer-Institut für Materialfluss und Logistik: Health Care Logistics, 2018.

[104] A. Wood, M. Altman, A. Bembenek, M. Bun, M. Gaboardi, J. Honaker, K. Nissim, D. O'Brien, T. Steinke, and S. Vadhan. Differential Privacy: A Primer for a Non-Technical Audience. *Vanderbilt Journal of Entertainment & Technology Law, Berkman Klein Center Research Publication*, 21(17), 2018.

[105] A. Yala, C. Lehman, T. Schuster, T. Portnoi, and R. Barzilay. A Deep Learning Mammography-based Model for Improved Breast Cancer Risk Prediction. *Radiology*, 292(1):60–66, May 2019. ISSN 0033-8419. doi: 10.1148/radiol.2019182716.

[106] K. Yousaf, Z. Mehmood, I. A. Awan, T. Saba, R. Alharbey, T. Qadah, and M. A. Alrige. A comprehensive study of mobile-health based assistive technology for the healthcare of dementia and Alzheimer's disease (AD). *Health Care Management Science*, June 2019. ISSN 1572-9389. doi: 10.1007/s10729-019-09486-0.

[107] R. Zhang, D. Chen, and X. Shang. Privacy preserving for patients' information: A knowledge-constrained access control model for hospital information systems. In *2016 IEEE 14th International Conference on Industrial Informatics (INDIN)*, pages 921–926, July 2016. doi: 10.1109/INDIN.2016.7819293.

[108] D. Zuehlke. SmartFactory—Towards a factory-of-things. *Annual Reviews in Control*, 34(1):129–138, Apr. 2010. ISSN 1367-5788. doi: 10.1016/j.arcontrol.2010.02.008.

Appendix B

Predicting ICU Bed Occupancy for Integrated Operating Room Scheduling via Neural Networks

The contribution [47] has been accepted for publication in "Naval Research Logistics", which is ranked in category B in the VHB-JOURQUAL3 ranking [2].

J. Schiele, T. Koperna, and J. O. Brunner. Predicting ICU Bed Occupancy for Integrated Operating Room Scheduling via Neural Networks. *Naval Research Logistics*, 2020. doi: 10.1002/nav.21929

Appendix C

Predicting Surgical Durations and Implications at the Operational Level

The following contribution [44] has been submitted to "Medical Decision Making", which is ranked in category A in the VHB-JOURQUAL3 ranking [2]. The submitted version is reproduced below in its entirety.

Predicting Surgical Durations and Implications at the Operational Level[1]

Julian Schiele, M. Sc.

University Center of Health Sciences at Klinikum Augsburg (UNIKA-T)

Neusässer Straße 47, 86156 Augsburg, Germany

Chair of Health Care Operations / Health Information Management

Faculty of Business and Economics, University of Augsburg

Universitätsstraße 16, 86159 Augsburg, Germany

E-mail: julian.schiele@unikat.uni-augsburg.de

ORCID: 0000-0002-4118-2889

Number of words: 4,017

Presented: June 25, 2019

30th European Conference on Operational Research conference, Dublin, Ireland

Submission date: December 30, 2019

[1]Financial support for this study was provided entirely by a contract with the University of Augsburg, Bavaria, Germany. The funding agreement ensured the authors' independence in designing the study, interpreting the data, writing, and publishing the report.

Predicting Surgical Durations and Implications at the Operational Level

Abstract

The operating room is the most critical resource and a major cost driver of a hospital. In times of increased economic pressure, it is of utmost importance to ensure an efficient management of the operating room environment. When scheduling surgeries at the operational level, clinicians consider several aspects such as estimated utilization of operating rooms, under- and overtime of staff, and the impact on downstream units. To do so, they rely on accurate predictions of surgical durations and operational implications. However, the predictions in most hospitals are based on either expert estimates or simple averaging methods making them prone to inaccuracy. In this study, we present a new multi-objective approach based on artificial neural networks that is able to predict various perioperative durations as well as operational implications with great practical relevance. Our model leverages a variety of patient-related, procedure-related, and operations-related factors some of which are derived from individual patient paths. The required input data is commonly available in most hospitals. We evaluate the model with 7 years of real world data covering more than 150,000 surgeries at Universitätsklinikum Augsburg, a German tertiary care hospital having 1,700 beds and serving all surgical specialties. Developed in close collaboration with the operating room management, the proposed model will be of great practical value and could support clinicians in their decision making.

Keywords: Surgical duration; surgery scheduling; length of stay; operating room; efficiency; prediction; machine learning; artificial neural network

Submission: December 30, 2019

1 Introduction

Most countries around the globe spend a significant share of their gross domestic product on health care, particularly for hospitals. Figure 1 illustrates the United States' health expenditures in 2016 (3.3 trillion USD), of which hospitals account for nearly one third (1.1 trillion USD) [4]. Hence, "pressures to make operating margins will continue to be at the forefront of most hospital and health system leaders' minds" [31] - particularly, since it seems that nothing "will stop public spending on health care from rising" [34]. Being

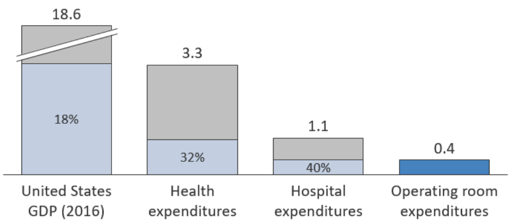

Figure 1: United States' health expenditures 2016 in trillion USD.

the core of a hospital, the operating room (OR) is one of the most expensive resources [9] causing around 40% of the hospital expenditures (see Figure 1) [5, 26]. While OR costs vary significantly between countries, hospitals, and procedures [28], an average OR hour in the United States accounts for roughly 4,000 USD [25, 36] and an average cancelled hour corresponds to roughly 1,500 USD in lost revenues [6, 7, 27]. Consequently, efficient utilization of the OR is of particular interest. A significant share of expenditures can be eliminated by addressing overtreatment, failures in coordination, administrative complexity, fraud and abuse [2] and by scheduling surgeries more efficiently [14].

Accurate predictions of surgical durations are essential for an efficient OR management. Both *overestimating* as well as *underestimating* surgical durations bear undesirable consequences. The former is associated with idle time and up to 60% higher costs of staff [38], low utilization of OR and other resources, lost revenues, and longer indirect waiting times of patients, i.e., the duration between time of request and appointment. The latter is associated with staff overtime, need for rescheduling and cancellation of subsequent surgeries (which affects $10 - 40\%$ of elective surgeries [32]), negative impact on downstream units, ramification on patient outcomes, and longer direct waiting times of patients, i.e., the duration between appointment and actual surgery. However, deriving accurate predictions for surgical durations is a complex endeavor for several reasons, particularly due to inherent *uncertainty* and *diversity* in processes, specialties, staff capabilities, and patient characteristics [8, 12]. Uncertainty is introduced by factors that are not known a priori, e.g., unexpected bleeding, whereas diversity refers to factors that are known a priori, e.g., patient age, emergency status, and type of procedure. In today's hospitals, predictions are either based on *expert estimates* or on *historical averaging* given electronic health records. The former is often biased by financial incentives, psychological pressure, and convenience [1, 32, 24], seizes numerous man-hours of valuable medical staff, and often results in inaccurate predictions, i.e., overestimation in 32% and underestimation in 42% of the time [21]. The latter is formed as arithmetic mean of historic procedure durations for an "average patient", neglects patient-specific factors such as age, sex, and emergency status, and yields only slightly better predictions [43]. A data-driven model that integrates the critical factors, predicts the surgical durations, and supports experts in their decision making could improve the accuracy [42, 9] and would be of great practical value.

Within the last decades, a vast amount of research has been done in the field of surgical scheduling [3], however, the problem of predicting surgical durations comprises only a minor fraction of the literature and just recently started to get more attention (see Figure 2). Most notably, Strum et al. [37] compare log-normal and normal distributions to model surgical durations, Tuwatananurak et al. [40] apply a proprietary machine learning algorithm, and Fairley et al. [11] present a machine learning approach for a pediatric hospital. Overall, we identified 114 papers in this field and distinguished between six areas according to their main focus as depicted in Figure 2, namely literature reviews (9 papers), probability distribution fitting (13), expert prediction assessments (6), data mining and machine learning models (68), scheduling and sequencing (12), and others.

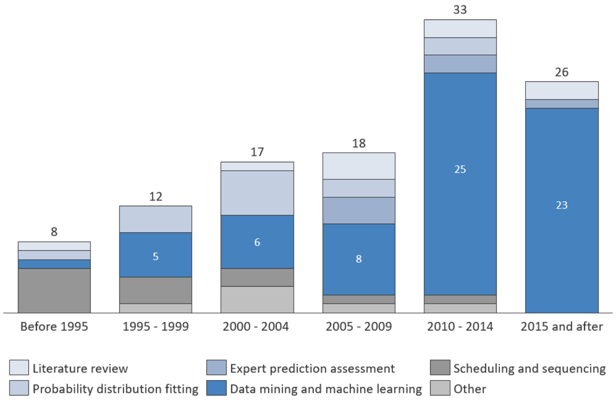

Figure 2: Recent literature addressing the prediction of surgical durations.

The individual characteristics of a surgical case not only determine the duration for which the operating room is occupied, but also have an influence on further perioperative durations, the further course of a patient's postoperative stay in the hospital, and even the subsequent, post-hospital treatments required. Ultimately, these are the metrics that contribute to the efficiency of a hospital. Given accurate predictions, these implications can well be considered during decision making at the operational level to improve the quality of surgery scheduling. It is the purpose of this study to leverage patient-related, procedure-related, and other operations-related factors in order to derive accurate predictions. Our contribution is a new multi-objective classification model based on artificial neural networks that is able to predict multi-class labels for six different metrics with great practical relevance, i.e., operating room duration, incision suture duration, anesthetist duration, postoperative unit, postoperative length of stay (LOS), and discharge type. Unlike other approaches, we also consider features that are derived from the individual patient path, e.g., the preoperative LOS. The model is generalizable and can also be applied to other hospitals since it is based on input data that is commonly available in most hospitals. We evaluate the model with real world data covering more than 150,000 surgeries from 125,000 patients within 2010 to 2016 and present convincing numerical results. We believe that the proposed model is of great practical value and could support clinicians in their decision making.

The paper at hand is organized in four sections. In this Section 1, we set up the context and stated the objectives of our work. Section 2 describes the underlying data of the reference hospital, explains the used methodology, and proposes our general multi-objective prediction model based on neural networks. In Section 3, we present the numerical results for the study. Eventually, we discuss our findings, limitations, and future work in Section 4.

2 Methods

In this section, we propose our prediction model to predict surgical durations and implications of a surgery at a hospital's operational level. The problem at hand corresponds to finding a computable function $\boldsymbol{x} \mapsto f(\boldsymbol{x}) = \boldsymbol{y}$ that maps an input vector $\boldsymbol{x}$, i.e., patient-related, procedure-related, and operations-related features, to an output vector $\boldsymbol{y}$, i.e., surgical durations and operational implications of the surgery. However, the relationship between inputs and outputs is not trivial and hence, it is difficult to develop an explicit model. Instead, we use machine learning to learn the relationship from a large data set with $m \in \mathcal{M} = \{1, \ldots, M\}$ historic samples $\left(\boldsymbol{x}^{(m)}, \boldsymbol{y}^{(m)}\right)$, i.e., surgical cases, where $\boldsymbol{x}_n, n \in \mathcal{N} = \{1, \ldots, N\}$, describes the features and $\boldsymbol{y}_o$ the corresponding multi-class labels for $o \in \mathcal{O} = \{1, \ldots, O\}$ outputs. In the following, we describe the underlying data source, definition of features and labels, pre-processing steps, training and evaluation of the machine learning based model, and implementation details.

The data for this study was retrieved from the hospital information system of our partner hospital, Universitätsklinikum Augsburg, a 1,700-bed, maximum-care university hospital located in Southern Germany. Our work is based on two sources of input data covering seven years from 2010 to 2016, namely *surgery records* containing timestamps and further details about surgical interventions and *supporting unit records* containing timestamps and further details about admission, transfers, and discharge of patients. Given the unique identifier for each patient, we are able to interlink the two data sources and reconstruct the path through the hospital for each individual patient. Based on our experience, these data sets are commonly available in most hospitals and can be retrieved directly from the hospital information system, e.g., Agfa Healthcare Orbis in the reference hospital. For this study, we exclude samples with missing information and restrict the analysis to the most common procedure types. In total, we consider M = 151,767 surgical cases covering P = 125,191 distinct patients. This comprehensive data set as well as our close cooperation with key stakeholders of the reference hospital provide a solid foundation for our approach. The study at hand has no external funding source.

After deliberation with the hospital management, we use *patient-related*, *procedure-related*, and *operations-related* parameters as input for the predictive model. As patient-related variables we consider patient age, gender, patient type, emergency status, and admission type. Procedure-related variables are International Classification of Procedures in Medicine (ICPM) code, OR type, and surgical specialty. Operations-related variables refer to the number of previous surgeries, the preoperative LOS in the hospital, the origin unit before being transferred to the OR, the weekday, and the time of the day. We distinguish between *continuous features* and *categorical features*. Table 1 depicts the continuous features ($n = 1, n = 2, n = 3$) including their means and standard deviations for three data sets. Note that the entire data set is divided into training set (60% of

Table 1: Continuous features considered in this study and their representations for the samples in the training set, validation set, and test set.

$n \in \mathcal{N}$	**Continuous feature**	**Training** (60%)	**Validation** (20%)	**Test** (20%)
1	Patient age, y	60.8 ± 21.6	60.8 ± 21.4	60.9 ± 21.5
2	Previous surgeries, count	0.5 ± 1.6	0.5 ± 1.7	0.5 ± 1.6
3	Preoperative LOS, min	3761.2 ± 9204.4	3721.8 ± 9348.0	3784.6 ± 9653.6

Note: Values are means ± standard deviations before oversampling and normalization. LOS = length of stay.

all samples), validation set (20%), and test set (20%) to avoid bias during training and evaluation of the model. There is no significant difference between the realizations within those three sets. In order to obtain values for features $n = 2$ and $n = 3$, we reconstruct each individual patient path and determine the location for each time during the hospital stay. Table 2 shows the categorical features considered in this study. We apply one-

Table 2: Categorical features considered in this study and their representations for the samples in the training set.

$n \in \mathcal{N}$	**Categorical feature**	**Representation**
4 – 5	Gender	Female (46.6), Male (53.4)
6 – 7	Patient type	Inpatient (80.4), Outpatient (19.6)
8 – 10	Emergency status	Elective (77.7), Urgent (16.5), Emergency (5.8)
11 – 16	Admission type	Regular (57.6), Work (2.5), Traffic/sport (0.0), War-disabled (0.0), Emergency (20.0), Others (19.9)
17 – 138	Procedures, ICPM	5-787 (3.0), 5-790 (2.7), 5-361.03 (2.2), 5-144.5a (2.4), etc.
138 – 144	OR type	Central (39.5), Roof (28.3), South (12.7), Uro/Endo (2.1), Sectio (1.8), Others (15.6)
145 – 157	Specialty	Trauma surg. (17.0), Ophthalmologic surg. (15.3), Urology (12.3), Cardio. surg. (10.0), Vascular surg. (9.3), General surg. (9.1), Gynecology (7.5), ENT (7.2), Dermatology (5.9), Neurosurg. (4.2), Pediatrics (2.2), Oral surg. (0.1), Radiotherapy (0.0)
158 – 162	Preoperative unit	Ward (72.7), Outside (20.0), ICU (3.4), ED (2.9), IMC (0.9)
163 – 169	Weekday	Mon (19.4), Tue (19.3), Wed (19.0), Thu (19.4), Fri (19.7), Sat (1.7), Sun (1.5)
170 – 173	Time of day	Morning (54.4), Afternoon (37.1), Evening (4.3), Night (4.2)

Note: Values in brackets are percentages before oversampling and normalization. There was no significant difference between training, validation, and test sets. ICPM = International Classification of Procedures in Medicine; OR = operating room; ENT = ear-nose-throat, ICU = intensive care unit; ED = emergency department; IMC = intermediate care unit.

hot encoding to convert the categorical variables into integers. For example, we convert the parameter 'Gender' with values 'Female' and 'Male' into two binary features called 'Female' and 'Male' with values $\{0, 1\}$, respectively. The values in brackets refer to the occurrence as percentage in the training set. In order to ensure a sufficient sample size for each procedure type, we restrict the analysis to the 122 most common procedure types. In fact, we include all full-digit ICPM codes that appeared at least 500 times, merge the remaining ICPM codes based on their first 5 digits, and also include all 5-digit ICPM codes with at least 500 samples. The features $n = 170$ to $n = 173$ are derived from the starting time of the respective surgery using 7am, 12pm, 5pm, and 8pm as thresholds. Having defined the input features, each sample shall now be complemented with the respective output values. The proposed model is able to predict three different surgical durations ($o = 1, o = 2, o = 3$) as well as three operational implications of the surgery ($o = 4, o = 5, o = 6$) as depicted in Table 3. As common for multi-class classification

Table 3: Labels considered in this study and their representations.

$o \in \mathcal{O}$	**Multi-class label**	**Representation**
1	OR duration	Q1 (≤ 27 min), Q2 (46), Q3 (72), Q4 (118), Q5 (> 118)
2	Incision suture dur.	Q1 (≤ 11 min), Q2 (24), Q3 (42), Q4 (79), Q5 (> 79)
3	Anesthetist dur.	Q1 (≤ 60 min), Q2 (86), Q3 (118), Q4 (178), Q5 (> 178)
4	Postoperative unit	Ward, IMC, ICU, Others
5	Postoperative LOS	Q1 (≤ 0 days), Q2 (1), Q3 (4), Q4 (8), Q5 (> 8)
6	Discharge type	Treatment completed, Completed and postoperative treatment planned, Transfer to other hospital, Discharge to rehab/care, Death, Others

Note: Values before oversampling. Values in brackets indicate the threshold for the quantiles, e.g., Q2 (46) refers to the range between 27 min and 46 min. OR = operating room; dur. = duration; IMC = intermediate care unit; ICU = intensive care unit; LOS = length of stay.

problems, we define a multi-class label $\boldsymbol{y}_o^{(m)}$ for each output $o \in \mathcal{O}$ with values $y_{o,c}^{(m)} \in [0, 1]$ for $c \in \mathcal{C}_o$ classes. For example, if the patient in sample $m = 1$ is transferred to the ward after surgery, the multi-class label for output $o = 4$ is obtained as $\boldsymbol{y}_{o=4}^{(1)} = \begin{bmatrix} 1 & 0 & 0 & 0 \end{bmatrix}^{\mathrm{T}}$. The OR duration is computed as difference between the arrival of a patient in the OR and the maximum of either the end of the procedure or the end of the anesthesia time. For the first output ($o = 1$), we then use the 0.2-, 0.4-, 0.6-, and 0.8-quantiles to define the five classes $c \in \mathcal{C}_{o=1}$ for the multi-class label, i.e., $y_{o=1,c} \in [0, 1], \forall c \in \{1, \ldots, 5\}$. The respective quantiles in minutes are shown in Table 3. The multi-class label for the second output ($o = 2$) is defined as the quantiles for the duration between incision and suture of the surgery. The anesthetist duration ($o = 3$) describes the duration between arrival and departure of the anesthetist. Note that surgical cases without anesthesia are neglected for the computations on $o = 3$ leading to a smaller effective sample size of

$N_3 = 114{,}911$. The postoperative downstream station to which a patient is transferred after surgery can be reconstructed from the individual patient path and is categorized as 'Ward', 'IMC' (intermediate care unit), 'ICU' (intensive care unit), or 'Others' resulting in the multi-class label with values $y_{o=4,c} \in [0,1], \forall c \in \{1,\ldots,4\}$ for the fourth output ($o = 4$). The classes for the fifth output ($o = 5$) are defined as quantiles for the duration between end of the surgery and discharge from the hospital. Depending on the type of discharge from the hospital, a patient is categorized into one of seven classes resulting in $y_{o=6,c} \in [0,1], \forall c \in \{1,\ldots,7\}$. The aforementioned features together with the respective multi-class labels form the samples that serve as foundation for our machine learning based model which we describe in the following.

The classification of each surgery case sample into one of the considered classes for each output is done by an *artificial neural network (ANN)* model. Within the last decade, machine learning based approaches, particularly ANNs, have gained momentum across many industries. They serve as powerful classifiers in health care, e.g., to detect skin cancer [10], to predict pneumonia [17], to classify airborne pollen [35], and for other computer-aided diagnosis applications [31, 39], and serve as medical decision support at the operational level of a hospital, e.g., to assess perioperative cardiac risks [20], identify patients at high risk of postinduction hypotension [23], and predict implantation outcome of individual embryos [41]. ANNs are well suited for the prediction of surgical durations by learning the relationship between features and outputs from historic data. A comprehensive overview on deep learning is provided in [15] and [22]. Figure 3 depicts a schematic representation of our proposed ANN model. The model consists of one *input layer* which is defined by the number of features, i.e., $N = 173$ for this study, several *output layers*, i.e., $O = 6$, each of which represents the respective multi-class label and is defined by the number of classes in this label, and finally a number of *hidden layers* $l \in \mathcal{L}_o$ consisting of several neurons. A *neuron* represents the basic building block of an ANN and it is characterized by three mathematical operations. In each neuron, the dot product between its input vector $\boldsymbol{x}$ and its internal weight vector $\boldsymbol{w}$ is computed and

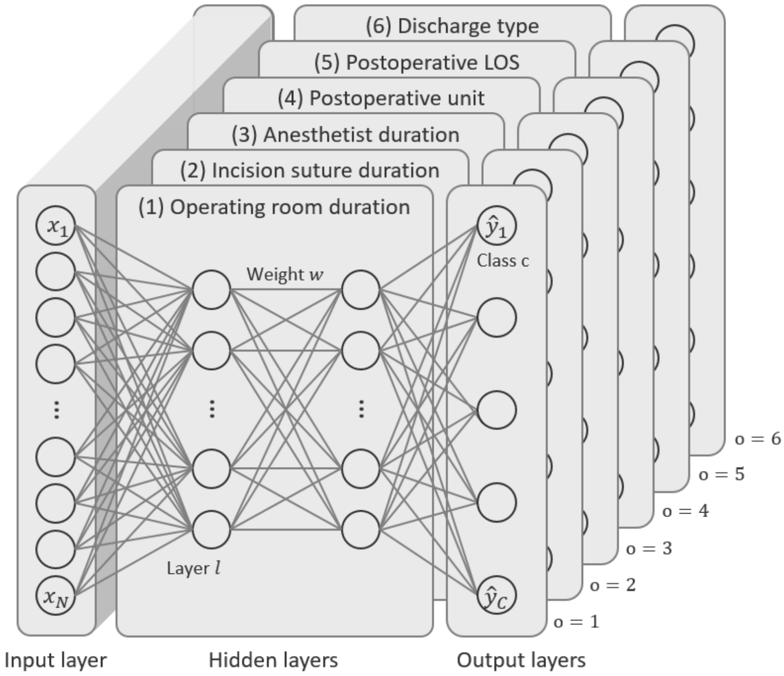

Figure 3: Schematic representation of artificial neural network model with $O = 6$ outputs.

a bias b is added (see Eq. 1). Then, an activation function $g(z)$ is applied to the term resulting in the prediction $\hat{y}$ (see Eq. 2).

$$z = \boldsymbol{w}\boldsymbol{x}^{\mathrm{T}} + b \tag{1}$$

$$\hat{y} = g(z) \tag{2}$$

Due to the lack of a well-established theoretic protocol to determine an ANN's topology [30], we experimented with different numbers of neurons in the hidden layers and tuned the hyperparameters on the validation set until we ended with a powerful model. Hence, the model computes a prediction $\hat{\boldsymbol{y}}_o$ for each multi-class label $\boldsymbol{y}_o, \forall o \in \mathcal{O}$, with the objective to minimize the loss function that describes the deviation between predicted and measured values. We used the Adam optimizer [19] with log loss, an initial learning rate of $\alpha = 10^{-4}$, and $E = 100{,}000$ epochs to train the multi-layer perceptron classifier iteratively [13, 16, 18]. As evaluation metric, we use the unweighted average of the class-specific F-measures (F1) being defined as the harmonic mean between recall and precision for each class. The class-specific precision is defined as the share of true positives, i.e., all samples of the respective class that were correctly labeled, and all predicted positives, i.e., including samples that were wrongly assigned to the respective class in Eq. (3).

$$\text{Precision} = \frac{\text{True positives}}{\text{True positives} + \text{False positives}} \tag{3}$$

Eq. (4) defines the class-specific recall as the share of true positives and all positives, i.e., including samples that were wrongly not assigned to the respective class.

$$\text{Recall} = \frac{\text{True positives}}{\text{True positives} + \text{False negatives}} \tag{4}$$

The proposed ANN model to classify surgical durations and operational implications was implemented with `scikit-learn` in Python [33]. For data processing, we use `Pandas`, `NumPy`, `SciPy`, `Matplotlib`, and `scikit-learn`. After having selected features and labels, we perform oversampling in case of unbalanced classes, encode the categorical features with `category_encoders`'s `OneHotEncoder`, and use `scikit-learn`'s `train_test_split` to split the resulting data set randomly into three subsets, i.e., M_{train} samples are assigned to the training set (60%), M_{val} samples to the validation set (20%), and M_{test} samples to the test set (20%). Afterwards, each feature is scaled and normalized individually with `scikit-learn`'s `MinMaxScaler` such that it is in the range between zero and one. When fitting the estimator, we only use the training set in order to prevent a spillover of information, i.e., the constants computed with the training set are also used to scale the validation and test set. Scores and confusion matrices are computed using `scikit-learn`'s `metrics`. Pre-processing and training of the ANN were performed on a dedicated simulation node equipped with 56 physical Intel(R) Xeon(R) Platinum 8176 cores with enabled hyperthreading.

3 Results

We use unweighted average precision (UAP), unweighted average recall (UAR), and unweighted average F-measure (UAF) to evaluate the performance of our classification model. Assuming that all classes of a multi-class label are equally important, we choose the unweighted scores as performance metrics. In total, 91,059 training samples are used to train the predictive model, 30,354 validation samples are used to select the best hyperparameters of the model, and 30,354 test samples are used to evaluate the performance of the model. Table 4 shows the selected hyperparameters as well as the resulting performance on the test set for each of the six outputs. Overall, we run various computations with different hyperparameters of the model. We use the tuple $r = (o, h) \in \mathcal{R} = \mathcal{O} \times \mathcal{H}$ as unique identifier for a particular run with set output $o \in \mathcal{O}$ and set ANN hyperparameters $h \in \mathcal{H}$. Table 4 shows the models achieving the best performance. For the prediction

Table 4: Performance of selected models on the test set.

$r \in \mathcal{R}$	**Output**	**ANN hyperparameters**	**UAP**	**UAR**	**UAF**
$(1, 1)$	OR duration	[10:10], Adam, ReLu	59.7	58.7	59.1
$(2, 1)$	Incision suture duration	[10:10], Adam, ReLu	58.4	57.5	57.5
$(3, 1)$	Anesthetist duration	[10:10], Adam, ReLu	55.8	55.6	55.4
$(4, 2)$	Postoperative unit	[100:50:25], Adam, ReLu	97.2	97.4	97.2
$(5, 1)$	Postoperative LOS	[10:10], Adam, ReLu	62.0	60.9	60.1
$(6, 2)$	Discharge type	[100:50:25], Adam, ReLu	84.3	84.8	84.2

Note: Performance values are percentages. $[x : x]$ describes the number of hidden layers and number of nodes per layer in the ANN topology. ANN = artificial neural network; UAP = unweighted average precision; UAR = unweighted average recall; UAF = unweighted average F-measure; OR = operating room.

of the OR duration ($o = 1$), we configure the ANN model with a topology of two hidden layers with 10 neurons each, ReLu activation, and the Adam optimizer ($h = 1$). We achieve an $\mathrm{UAP}_{(1,1)}$ of 59.7%, an $\mathrm{UAR}_{(1,1)}$ of 58.7%, and an $\mathrm{UAF}_{(1,1)}$ of 59.1% on the test set which significantly outperform a random classification model with respective values of 20.0% for five classes $C_1 = 5$. We use the subscript $r = (1, 1)$ to indicate that the performance metrics relate to the first run, i.e., prediction of the OR duration using the [10:10] ANN model. The three surgical durations ($o = 1, o = 2, o = 3$) show a similar performance with the anesthesist duration showing the lowest performance in all three metrics. The three surgical implications ($o = 4, o = 5, o = 6$) show even better performance with UAFs of at least 60.1% and up to 97.2%. The performance for the postoperative unit in $r = (4, 2)$ with an $\text{UAP}_{(4,2)}$ of 97.2%, an $\text{UAR}_{(4,2)}$ of 97.4%, and an $\text{UAF}_{(4,2)}$ of 97.2% is particularly good. Our model also achieves high accuracy for the prediction of the discharge type, i.e., $\text{UAP}_{(6,2)} = 84.3\%$, $\text{UAR}_{(6,2)} = 84.8\%$, and $\text{UAF}_{(6,2)} = 84.2\%$, despite having more classes, i.e., $C_6 = 6$ The computation time for the training of the ANN model varies significantly depending on output and ANN hyperparameters, i.e., from 460 min for $r = (3, 1)$ up to 131 hours for $r = (6, 2)$.

In order to study a model's performance for each individual class of the multi-class output, we use a normalized confusion matrix. In a confusion matrix, each row represents the share of instances in an true class and each column represents the instances in a predicted class. The values are normalized per true label, i.e., per row, such that the diagonal shows the class-specific recall. Figure 4 depicts the confusion matrices for the surgical durations ($o = 1, o = 2, o = 3$). In the confusion matrix for the OR duration

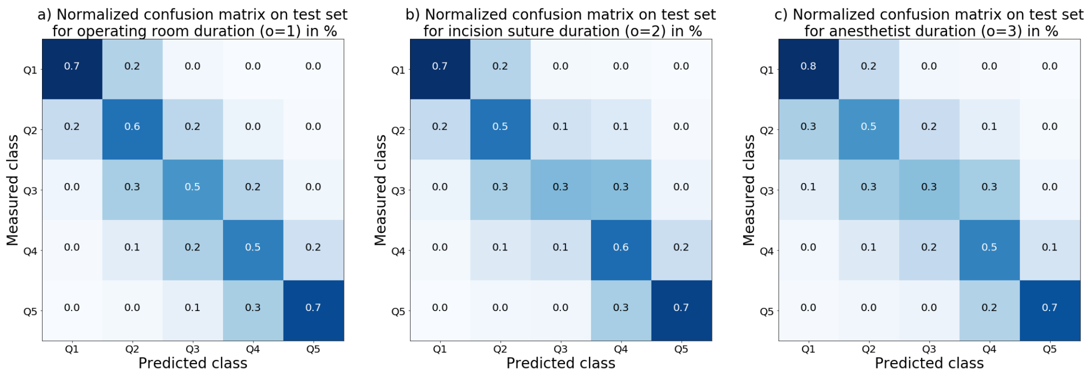

Figure 4: Normalized confusion matrices for our model showing predicted vs. measured (a) operating room duration, (b) incision suture duration, and (c) anesthetist duration. The percentage of possible instances is shown as color scale and noted in each square.

depicted in Figure 4a, samples in 'Q1' (≤ 27 min) are classified particularly well showing a recall of 74.5%. However, samples in 'Q3' show a slightly worse performance (45.2%)

since it is often confused with 'Q2' (25.7%) and 'Q4' (22.6%). Similarly, also for $o = 2$ in Figure 4b and for $o = 3$ in Figure 4c, the class 'Q1' (≤ 11 min and ≤ 60 min, respectively) shows best performance, i.e., 74.0% and 77.3%, and the class 'Q3' the worst, 33.5% and 32.8%, respectively. Figure 5 depicts the confusion matrices for the operational implications of a surgery ($o = 4, o = 5, o = 6$). The confusion matrix for the postoperative

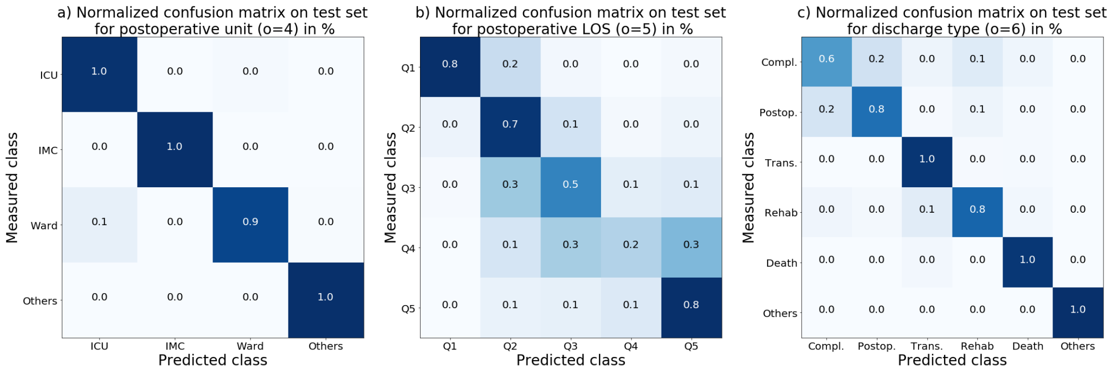

Figure 5: Normalized confusion matrices for our model showing predicted vs. measured (a) postoperative unit, (b) postoperative LOS, and (c) discharge type. The percentage of possible instances is shown as color scale and noted in each square. Note: LOS = length of stay; Compl. = treatment completed; Postop. = completed and postoperative treatment planned; Trans. = transfer to other hospital; Rehab = discharge to rehab/care.

unit depicted in Figure 5a confirms that all classes are predicted particularly well showing recalls from 91.5% ('Ward') up to 99.8% ('IMC'). Also, the postoperative LOS ($o = 5$) is predicted particularly well with a recall of up to 76.8% for class 'Q1'. However, class 'Q4' shows a rather bad performance of 24.1% being confused very often with 'Q3' (27.2%) and 'Q5' (34.8%). Most classes of discharge types ($o = 6$) are predicted with high accuracy, only the class 'Treatment completed' with a recall of (58.9%) is often confused with the class 'Completed and postoperative treatment planned' (21.9%).

4 Discussion

In the paper at hand, a new multi-objective classification model based on ANNs for the prediction of various surgical durations and operational implications was presented. We showed that patient-related, procedure-related, and operations-related parameters retrieved from the hospital information system can be used as features for an ANN model in order to predict multiple multi-class labels. In particular, also individual patient paths were reconstructed to incorporate features and labels that go beyond the time in the OR. Applied to an extensive data set from Universitätsklinikum Augsburg, the proposed model achieved convincing results, i.e., UAFs of at least 55.4% up to 97.2%. The prediction of the postoperative unit ($o = 4$) shows the highest UAF which might partly result from the fact that this output comprises only $C_4 = 4$ different classes, i.e., 'ICU', 'IMC', 'Ward', and 'Others'. The prediction of the anesthetist duration ($o = 3$) shows the lowest UAF which might be due to the smaller sample size of $N_3 = 114{,}911$ containing patients that undergo anesthesia as well as unreliable documentation. For all outputs, the model significantly outperforms a random classification model with values of 25.0% for four classes ($o = 4$), 20.0% for five classes ($o = 1, o = 2, o = 3, o = 5$), and 16.7% for six classes ($o = 6$), respectively. No comparison of our results with the literature was possible since we were unable to identify a similar multi-objective classification model for surgical durations and operational implications.

In order to evaluate and compare the predictive performance of the overall model, we propose to use aggregated versions of the performance metrics. Given the hyperparameters $h_o, \forall o \in \mathcal{O}$, for each output o, we use $\boldsymbol{h} = (h_1, \ldots, h_O)$ to describe the overall model. The weighted aggregated versions are defined as aggregated UAP (AUAP) in Eq. 5, aggregated UAR (AUAR) in Eq. 6, and aggregated UAF (AUAP) in Eq. 7

$$\mathrm{AUAP}_{\boldsymbol{h}} = \frac{\sum_{o \in \mathcal{O}} \kappa_o \mathrm{UAP}_{(o,h_o)}}{\sum_{o \in \mathcal{O}} \kappa_o} \tag{5}$$

$$\mathrm{AUAR}_{\boldsymbol{h}} = \frac{\sum_{o \in \mathcal{O}} \kappa_o \mathrm{UAR}_{(o,h_o)}}{\sum_{o \in \mathcal{O}} \kappa_o} \tag{6}$$

$$\mathrm{AUAF}_{\boldsymbol{h}} = \frac{\sum_{o \in \mathcal{O}} \kappa_o \mathrm{UAF}_{(o,h_o)}}{\sum_{o \in \mathcal{O}} \kappa_o} \tag{7}$$

where the relative importance of each output is adjusted with the weight $\kappa_o \in \mathcal{K}, \forall o \in \mathcal{O}$.

Those metrics might be helpful to find the best model for individual preferences in terms of output relevance. In case of equally important outputs, i.e., $\kappa_o = 1, \forall o \in \mathcal{O}$, the proposed model using the hyperparameters h_o for each output as depicted in Table 4 achieves an AUAP of 69.6%, an AUAR of 69.1%, and an AUAF of 68.9%.

The limitations of our study pave the way for future research. We see mainly five directions to enhance the proposed approach in the future. First, the integration of additional features might further improve the predictive power of the model. Medical history, previous diagnosis results, patient weight, patient height, and known risk factors are just a few of the examples for additional patient-related features. Similarly, room numbers and used medical equipment could be integrated as procedure-related features and season, workload, and occupancy levels in downstream units as operations-related features. Additionally, the incorporation of staff-related features such as name of surgeon, seniority of surgeon, size of surgical team, name of anesthetist, and number of conducted surgeries would be a valuable extension of the model, however, they might also interfere with data protection regulations. Second, the development of customized performance metrics might help to tailor the predicted outcome to individual preferences. For example, weighting the classes of each multi-class label would allow to increase the predictive power for some selected classes, e.g., for very long surgical durations or for high congestion periods. Master *et al.* [29] introduced a customized loss function that does not penalize the deviation below or above a certain predicted duration, e.g., predicted durations below 15 min or above 60 min. Third, future work should compare the aggregated performance metrics achieved by our model with the ones achieved by ANN models with different hyperparameters and also with other machine learning based models such as support vector machines. Fourth, the incorporation of the proposed prediction model in an optimization framework might help to derive optimal surgery schedules for various objectives, e.g., minimize the daily overtime or minimize the daily number of surgeries that generate ICU patients. One might define a weighted objective function based on the predictions $\hat{\boldsymbol{y}}_o, o \in \mathcal{O}$, and minimize it with metaheuristics, e.g., a genetic algorithm.

Fifth, the proposed model should be applied and tested by clinicians in practice accompanied by a detailed assessment of their experiences with the supporting tool. Also, the comparison of the predictive performance of our model with their manual predictions might generate valuable insights.

It is our intention to encourage the usage of historic data and machine learning based methods to derive more accurate predictions for the implications of surgeries at the operational level. Ultimately, this will improve the quality of surgery schedules, save valuable staff time, and enhance the efficiency in the OR and beyond. We hope that the proposed model proves useful to inform better decisions.

Acknowledgments

We want to thank Universitätsklinikum Augsburg's Dr. Thomas Koperna and University Augsburg's Prof. Dr. Jens O. Brunner for their valuable input and helpful discussions. A special thank you to Henrina Herold for her thesis which paved the way for this study. We also thank the Institute for Communications Engineering of the Technical University of Munich for using their computational resources while training the neural network model.

References

[1] A. E. Abouleish, F. Dexter, R. H. Epstein, D. A. Lubarsky, C. W. Whitten, and D. S. Prough. Labor Costs Incurred by Anesthesiology Groups Because of Operating Rooms Not Being Allocated and Cases Not Being Scheduled to Maximize Operating Room Efficiency. *Anesthesia & Analgesia*, pages 1109–1113, 2003. ISSN 0003-2999.

[2] D. M. Berwick and A. D. Hackbarth. Eliminating waste in US health care. *Journal of the American Medical Association*, 307(14):1513–1516, 2012.

[3] B. Cardoen, E. Demeulemeester, and J. Beliën. Operating room planning and scheduling: A literature review. *European Journal of Operational Research*, 201(3):921–932, Mar. 2010. ISSN 0377-2217. doi: 10.1016/j.ejor.2009.04.011.

[4] Centers for Medicare & Medicaid Services. National Health Expenditures 2016, 2016.

[5] B. Denton, J. Viapiano, and A. Vogl. Optimization of surgery sequencing and scheduling decisions under uncertainty. *Health Care Management Science*, 10(1):13–24, Feb. 2007. ISSN 1572-9389. doi: 10.1007/s10729-006-9005-4.

[6] F. Dexter, J. Blake, D. Penning, and D. Lubarsky. Calculating a Potential Increase in Hospital Margin for Elective Surgery by Changing Operating Room Time Allocations or Increasing Nursing Staffing to Permit Completion of More Cases: A Case Study. *Anesthesia & Analgesia*, 94:138–42, table of contents, Feb. 2002. doi: 10.1097/00000539-200201000-00026.

[7] F. Dexter, E. Marcon, R. Epstein, and J. Ledolter. Validation of Statistical Methods to Compare Cancellation Rates on Day of Surgery. *Anesthesia & Analgesia*, 101:465–73, table of contents, Sept. 2005. doi: 10.1213/01.ANE.0000154536.34258.A8.

[8] F. Dexter, E. U. Dexter, and J. Ledolter. Influence of procedure classification on process variability and parameter uncertainty of surgical case durations. *Anesthesia & Analgesia*, 110(4):1155–1163, Apr. 2010. ISSN 1526-7598. doi: 10.1213/ANE.0b013e3181d3e79d.

[9] M. J. C. Eijkemans, M. van Houdenhoven, T. Nguyen, E. Boersma, E. W. Steyerberg, and G. Kazemier. Predicting the unpredictable: A new prediction model for operating room times using individual characteristics and the surgeon's estimate. *Anesthesiology*, 112(1):41–49, Jan. 2010. ISSN 1528-1175. doi: 10.1097/ALN.0b013e3181c294c2.

[10] A. Esteva, B. Kuprel, R. A. Novoa, J. Ko, S. M. Swetter, H. M. Blau, and S. Thrun. Dermatologist-level classification of skin cancer with deep neural networks. *Nature*, 542(7639):115–118, Feb. 2017. ISSN 1476-4687. doi: 10.1038/nature21056.

[11] M. Fairley, D. Scheinker, and M. L. Brandeau. Improving the efficiency of the operating room environment with an optimization and machine learning model. *Health Care Management Science*, 22(4):756–767, Dec. 2019. ISSN 1572-9389. doi: 10.1007/s10729-018-9457-3.

[12] A. Fügener, S. Schiffels, and R. Kolisch. Overutilization and underutilization of operating rooms - insights from behavioral health care operations management. *Health Care Management Science*, 20 (1):115–128, Mar. 2017. ISSN 1386-9620. doi: 10.1007/s10729-015-9343-1.

[13] X. Glorot and Y. Bengio. Understanding the difficulty of training deep feedforward neural networks. In *Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics*, pages 249–256, Mar. 2010.

[14] C. Gomes, B. Almada-Lobo, J. Borges, and C. Soares. Integrating Data Mining and Optimization Techniques on Surgery Scheduling. In S. Zhou, S. Zhang, and G. Karypis, editors, *Advanced Data Mining and Applications*, Lecture Notes in Computer Science, pages 589–602, Berlin, Heidelberg, 2012. Springer. ISBN 978-3-642-35527-1. doi: 10.1007/978-3-642-35527-1_49.

[15] I. Goodfellow, Y. Bengio, and A. Courville. *Deep Learning*. The MIT Press, 2016. ISBN 978-0-262-03561-3.

[16] K. He, X. Zhang, S. Ren, and J. Sun. Delving Deep into Rectifiers: Surpassing Human-Level Performance on ImageNet Classification. *IEEE International Conference on Computer Vision (ICCV 2015)*, 1502, Feb. 2015. doi: 10.1109/ICCV.2015.123.

[17] P. S. Heckerling, B. S. Gerber, T. G. Tape, and R. S. Wigton. Prediction of community-acquired pneumonia using artificial neural networks. *Medical Decision Making*, 23(2):112–121, 2003 Mar-Apr. ISSN 0272-989X. doi: 10.1177/0272989X03251247.

[18] G. E. Hinton. Connectionist learning procedures. *Artificial Intelligence*, 40(1-3):185–234, Sept. 1989. ISSN 00043702. doi: 10.1016/0004-3702(89)90049-0.

[19] D. P. Kingma and J. Ba. Adam: A Method for Stochastic Optimization. *arXiv:1412.6980 [cs]*, Dec. 2014.

[20] P. Lapuerta, G. J. L'Italien, S. Paul, R. C. Hendel, J. A. Leppo, L. A. Fleisher, M. C. Cohen, K. A. Eagle, and R. P. Giugliano. Neural network assessment of perioperative cardiac risk in vascular surgery patients. *Medical Decision Making*, 18(1):70–75, 1998 Jan-Mar. ISSN 0272-989X. doi: 10.1177/0272989X9801800114.

[21] D. Laskin, A. Abubaker, and R. Strauss. Accuracy of Predicting the Duration of a Surgical Operation. *Journal of oral and maxillofacial surgery*, 71:446–7, Feb. 2013. doi: 10.1016/j.joms.2012.10.009.

[22] Y. LeCun, Y. Bengio, and G. Hinton. Deep learning. *Nature*, 521(7553):436–444, May 2015. ISSN 1476-4687. doi: 10.1038/nature14539.

[23] C.-S. Lin, C.-C. Chang, J.-S. Chiu, Y.-W. Lee, J.-A. Lin, M. S. Mok, H.-W. Chiu, and Y.-C. Li. Application of an artificial neural network to predict postinduction hypotension during general anesthesia. *Medical Decision Making*, 31(2):308–314, 2011 Mar-Apr. ISSN 1552-681X. doi: 10.1177/0272989X10379648.

[24] A. Macario. Truth in Scheduling: Is It Possible to Accurately Predict How Long a Surgical Case Will Last? *Anesthesia & Analgesia*, 108:681–5, Apr. 2009. doi: 10.1213/ane.0b013e318196a617.

[25] A. Macario. What does one minute of operating room time cost? *Journal of clinical anesthesia*, 22:233–6, June 2010. doi: 10.1016/j.jclinane.2010.02.003.

[26] A. Macario, T. Vitez, B. Dunn, and T. McDonald. Where Are the Costs in Perioperative Care?: Analysis of Hospital Costs and Charges for Inpatient Surgical Care. *Anesthesiology*, 83(6):1138–1144, Dec. 1995. ISSN 0003-3022.

[27] A. Macario, F. Dexter, and R. Traub. Hospital Profitability per Hour of Operating Room Time Can Vary Among Surgeons. *Anesthesia & Analgesia*, 93:669–75, Oct. 2001. doi: 10.1097/00000539-200109000-00028.

[28] N. Master, D. Scheinker, and N. Bambos. Predicting Pediatric Surgical Durations. *arXiv:1605.04574 [stat]*, May 2016.

[29] N. Master, Z. Zhou, D. Miller, D. Scheinker, N. Bambos, and P. Glynn. Improving predictions of pediatric surgical durations with supervised learning. *International Journal of Data Science and Analytics*, 4(1):35–52, Aug. 2017. ISSN 2364-4168. doi: 10.1007/s41060-017-0055-0.

[30] A. S. Miller, B. H. Blott, and T. K. hames. Review of neural network applications in medical imaging and signal processing. *Medical and Biological Engineering and Computing*, 30(5):449–464, Sept. 1992. ISSN 1741-0444. doi: 10.1007/BF02457822.

[31] P. Natarajan, J. C. Frenzel, and D. H. Smaltz. *Demystifying Big Data and Machine Learning for Healthcare*. CRC Press, Taylor & Francis Group, 2017. ISBN 978-1-138-03263-7.

[32] J. Pandit and A. Carey. Estimating the duration of common elective operations: Implications for operating list management. *Anaesthesia*, 61:768–76, Aug. 2006. doi: 10.1111/j.1365-2044.2006.04719.x.

[33] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, and D. Cournapeau. Scikit-learn: Machine Learning in Python. *Machine Learning in Python*, 2011.

[34] E. Porter. A world of rising health care costs. *The New York Times*, 2013.

[35] J. Schiele, F. Rabe, M. Schmitt, M. Glaser, F. Häring, J. O. Brunner, B. Bauer, B. Schuller, C. Traidl-Hoffmann, and A. Damialis. Automated Classification of Airborne Pollen using Neural Networks. In *41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, July 2019. doi: 10.1109/EMBC.2019.8856910.

[36] R. D. Shippert. A Study of Time-Dependent Operating Room Fees and How to save $100 000 by Using Time-Saving Products. *The American Journal of Cosmetic Surgery*, 22(1):25–34, Mar. 2005. ISSN 0748-8068. doi: 10.1177/074880680502200104.

[37] D. P. Strum, J. H. May, and L. G. Vargas. Modeling the uncertainty of surgical procedure times: Comparison of log-normal and normal models. *Anesthesiology*, 92(4):1160–1167, Apr. 2000. ISSN 0003-3022.

[38] K. Tankard, T. Acciavatti, J. Vacanti, M. Heydarpour, S. Beutler, H. Flanagan, and R. Urman. Contributors to Operating Room Underutilization and Implications for Hospital Administrators. *The Health Care Manager*, 37:1, Mar. 2018. doi: 10.1097/HCM.0000000000000214.

[39] C. Thuemmler and C. Bai, editors. *Health 4.0: How Virtualization and Big Data Are Revolutionizing Healthcare*. Springer International Publishing, 2017. ISBN 978-3-319-47616-2.

[40] J. P. Tuwatananurak, S. Zadeh, X. Xu, J. A. Vacanti, W. R. Fulton, J. M. Ehrenfeld, and R. D. Urman. Machine Learning Can Improve Estimation of Surgical Case Duration: A Pilot Study. *Journal of Medical Systems*, 43(3):44, Jan. 2019. ISSN 1573-689X. doi: 10.1007/s10916-019-1160-5.

[41] A. Uyar, A. Bener, and H. N. Ciray. Predictive Modeling of Implantation Outcome in an In Vitro Fertilization Setting: An Application of Machine Learning Methods. *Medical Decision Making*, 35(6):714–725, Aug. 2015. ISSN 1552-681X. doi: 10.1177/0272989X14535984.

[42] I. H. Wright, C. Kooperberg, B. A. Bonar, and G. Bashein. Statistical modeling to predict elective surgery time. Comparison with a computer scheduling system and surgeon-provided estimates. *Anesthesiology*, 85(6):1235–1245, Dec. 1996. ISSN 0003-3022.

[43] A. Wu, C.-C. Huang, M. Weaver, and R. Urman. Use of Historical Surgical Times to Predict Duration of Primary Total Knee Arthroplasty. *The Journal of Arthroplasty*, 31, May 2016. doi: 10.1016/j.arth.2016.05.038.

Appendix D

Automated Classification of Airborne Pollen using Neural Networks

The contribution [46] has been accepted for publication in “41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)”, which is not ranked in the VHB-JOURQUAL3 ranking [2]. It is accessible in IEEE Xplore and indexed by PubMed and Medline.

J. Schiele, F. Rabe, M. Schmitt, M. Glaser, F. Häring, J. O. Brunner, B. Bauer, B. Schuller, C. Traidl-Hoffmann, and A. Damialis. Automated Classification of Airborne Pollen using Neural Networks. In *41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2019. doi: 10.1109/EMBC.2019.8856910