



The exploitation of data to support decision-making in healthcare: a systematic literature review and future research directions

Luigi Jesus Basile¹  · Nunzia Carbonara² · Umberto Panniello² · Roberta Pellegrino²

Received: 5 May 2023 / Accepted: 13 November 2024
© The Author(s) 2024

Abstract

The development of new technologies and their continued adoption allow data to be collected, analysed and exploited for decision-making. Data can play an important role in the healthcare industry since it is a complex system where every decision is strongly affected by risk and uncertainty. Although the proliferation of data and the awareness of the importance of new technologies to support decision-making in presence of risk and uncertainty, there is a lack of understanding of the inter-relations between data, decision-making process and risk management in healthcare organizations and their role to deliver healthcare services. Pursued by this research gap, the objective of this study is to understand how data can optimize decisions confronted with risk and uncertainty in the main domains (structure, process, outcome) of healthcare organizations. Thus, we conducted a systematic literature review based on the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) methodology, by selecting and analysing peer-reviewed journal articles from three databases: Scopus, Web of Science and PubMed. The paper's findings suggest that although data are widely used to optimize the decisions in the healthcare organization domains in presence of risk and uncertainty, there are still many scientific and practice gaps that lead to the definition of a future research agenda.

Keywords Data-driven · Decision-making · Healthcare · Risk · Uncertainty · Artificial intelligence

Extended author information available on the last page of the article

1 Introduction

In recent years, digital transformation has led to an exponential increase in data creation (Kraus et al. 2021; OECD 2020). This surge in data offers significant opportunities to enhance decision-making (DM) processes across various sectors, including healthcare (Galaitis et al. 2021; Kraus et al. 2021a, b). The exploitation of data provides a promising opportunity to enable decision-makers to generate valuable insights and knowledge to prepare for and manage upcoming risks and uncertainty, thus, supporting the risk management (RM) process (Basile et al. 2024b; Dicuonzo et al. 2021). Healthcare organizations are undergoing a profound digital transformation, characterized by the accelerated adoption of digital technologies, evidenced by the increasing volume of research and the growing number of applications in this sector (Basile et al. 2024a; Dal Mas et al. 2023; Kraus et al. 2021a, b; Sechi et al. 2020). The healthcare decision-making is particularly challenging due to its inherently dynamic, uncertain nature and complexity (Brehmer 1992; Champion et al. 2019; Han et al. 2019). Healthcare organizations must navigate the interactions with various stakeholders, such as insurance companies, pharmaceutical firms, families, and patients, whose non-linear interactions can lead to unexpected and often undesirable outcomes (Lipsitz 2012).

In this context, decision-making in healthcare impacts different organizational levels, which have been categorized into three main domains: (i) structure, which is related to professional and organizational resources, (ii) process, considering the tasks performed for patients, (iii) outcome, concerning the outcomes of the care provided and the outcomes of the care provided (Donabedian 1988).

Various studies have been conducted to highlight the benefits of using data in other industries (Dicuonzo et al. 2021; Tantalaki et al. 2019), however, the use of data and more specifically their interrelations with DM and RM in healthcare organizations remain underexplored. Addressing these challenges requires leveraging new technologies and data to support DM and RM processes, considering the numerous variables, environments, and interactions involved in healthcare (Champion et al. 2019; Chen et al. 2012).

Recent studies have started to explore these concepts (Basile et al. 2024b; Dal Mas et al. 2023; Gomes et al. 2023; Kraus et al. 2021a, b) but there is still room for further research to provide a cohesive understanding of their interrelations. This research aims to contribute to this growing body of knowledge by systematically reviewing the literature to uncover the potential of data towards decision optimization in conditions of risk and uncertainty considering the three healthcare organization's domains. The guiding research question for this study is:

- How can data optimize decisions confronted with risk and uncertainty in the main domains (structure, process, outcome) of healthcare organizations?

This paper is structured as follows: Sect. 2 outlines the theoretical background, covering the existing literature on data for decision-making in healthcare, as well as risk and uncertainty in healthcare decisions. Section 3 describes the methodology used for the systematic literature review (SLR) and the development of the framework based

on the study's rationale. Section 4 presents and discusses the findings from the literature in relation to the proposed framework. Section 5 suggests directions for future research. Finally, Sect. 6 concludes the paper.

2 Theoretical background

2.1 Data for decision-making in healthcare

The exponential increase in data production (Al-Ali et al. 2017; Florence and Shyamala Kumari 2019; Provost and Fawcett 2013) and the proliferation of analytical techniques (Ahmed et al. 2018; Galetsi et al. 2020; Tantalaki et al. 2019) are driving various industries towards data-driven decision-making to achieve better outcomes and performance (OECD 2015). Consequently, decision-makers are acquiring the skills and resources necessary to gather, store, retrieve, and analyze data to support their decisions (Meyer et al. 2014). National Health Systems are also promoting the development of data-analytical skills among physicians in healthcare organizations (Howard et al. 2015).

Healthcare is widely recognized as one of the most promising industries for data-driven decision-making (Galetsi et al. 2020; Ngiam and Khor 2019; Weerasinghe et al. 2022). The use of data for decision-making in healthcare organizations has gained significant attention from both academics and practitioners. Numerous applications have been proposed to leverage data for supporting decisions in healthcare (P.-T. Chen et al. 2020; Weerasinghe et al. 2022). Decisions in healthcare can be made by various stakeholders across three main domains of the organization: structure, process, and outcome (Donabedian 1988).

At the structure level, decisions involve professional and organizational resources related to the delivery of healthcare services, such as material resources, human resources, and organizational activities. For example, data can be used to predict likely trajectories or to schedule activities based on the expected length of stay (Simsek et al. 2020, 2021). At the process level, decisions pertain to patient care and the activities patients engage in to receive services. Data can support physicians in making decisions, such as selecting the best treatment and care pathways (Sun et al. 2021). Finally, the outcome domain focuses on the desired results of the care provided by healthcare practitioners, such as patient satisfaction. In this domain, data can be utilized to improve patient health outcomes, such as identifying at-risk patients or diagnosing diseases at an early stage.

Despite the increasing adoption of data in healthcare, its transformation into knowledge to improve the outcomes of healthcare services seems to be still limited (OECD 2019). In this context, this paper aims at understanding how data can optimize decisions in the main domains (structure, process, outcome) of the healthcare organization.

2.2 Uncertainty and risks in decisions in healthcare

Coping with uncertainty and risks has always been a cornerstone for scholars and practitioners across various industries (Lipshitz and Strauss 1997). Although risk and uncertainty can arise in different contexts, they often stem from common sources such as the complexity of relationships between variables, the indeterminacy of outcomes, and the lack of information (Han et al. 2011, 2019). Healthcare organizations, in particular, operate in high-risk and uncertain environments (Babrow et al. 1998; Champion et al. 2019; Spiegelhalter 2008). According to Han et al. (2011), risks and uncertainties in healthcare can emerge from the organization of healthcare facilities, the processes of diagnosis and prognosis, and the psychological conditions of patients.

To mitigate clinical and non-clinical uncertainties and risks, systematic methods for risk analysis and assessment are essential (Cagliano et al. 2011; Ferroni et al. 2017; Odone et al. 2019; Tartaglia et al. 2012). The International Standard Organization (ISO) has established a standard for applying risk management to decrease the risks of adverse events (ISO 2018). According to ISO, the risk management process involves defining the scope, context, and criteria, and it consists of several phases: Risk Assessment (RA), Risk Treatment (RT), and Risk Monitoring (RM).

The RA phase includes risk identification, risk analysis, and risk evaluation. It involves identifying risks that may affect an organization's objectives, analyzing these risks to understand their nature and characteristics, and evaluating them. The RT phase focuses on selecting and implementing optimal strategies to address identified risks. RM involves continuously monitoring risks to ensure the effectiveness of the adopted strategies and determining if new actions are necessary to mitigate risks.

Integrating data into healthcare decision-making can significantly enhance the ability to manage risk and uncertainty and improve the performance of medical systems (Aceto et al. 2020; Kraus et al. 2021a, b). Despite the increasing availability of data, its implementation as a tool for supporting risk management in healthcare remains underdeveloped (Basile et al. 2024b; Dicuonzo et al. 2021). Although there is growing research interest in understanding how to deal with risk and uncertainty in healthcare, much remains to be clarified (Han et al. 2019). Even though the growing interest of research into understanding how to deal with risk and uncertainty in healthcare, there is still much to be clarified (Han et al. 2019). Therefore, this paper aims to explore how data can help healthcare organizations optimize decisions confronted with risk and uncertainty.

3 Methodology

In this paper, we carry out a systematic literature review (SLR) to understand the potential of data to support decisions affected by risk and uncertainty in the healthcare organization setting. The systematic review methodology is based on a scientific, transparent and replicable protocol that aims to reduce biases related to the selective reporting of outcomes and arbitrariness in decision-making when extracting and using data from primary research (Moher et al. 2009).

3.1 Review process

In the current study, the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) have been adopted as the protocol for the systematic review (Moher et al. 2009). PRISMA provides a roadmap to report a systematic review in a transparent, objective, and explicit way, and it develops the review in four main stages: Identification, Screening, Eligibility and Inclusion. The whole review process is shown in Fig. 1.

As the first step (i.e., identification), in order to ensure the comprehensiveness of the research the authors selected three databases to retrieve the scientific papers: Scopus, Web of Science and PubMed. Then, keywords to conduct the research on the databases have been selected. Specifically, the final set of keywords is based on the three main pillars of this research, namely: data, the healthcare industry, and the

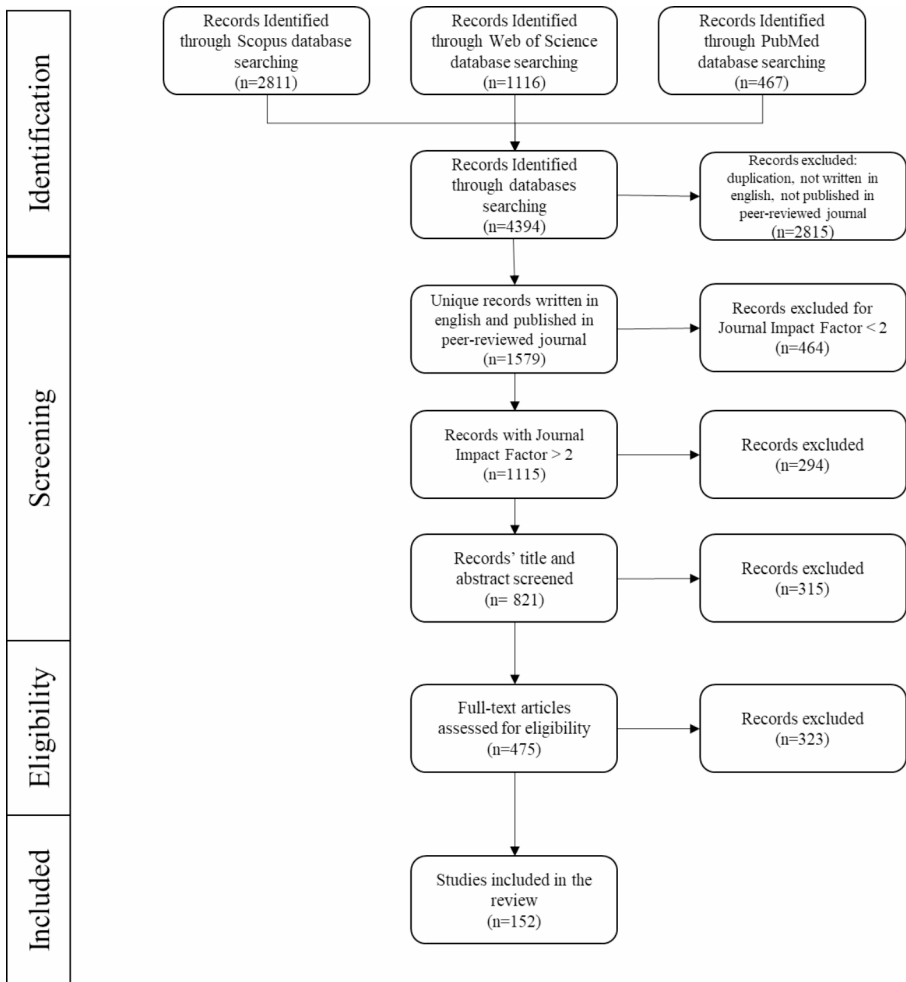


Fig. 1 PRISMA flow diagram for the identification, screening, eligibility, and inclusion of studies

decision-making process. Concerning the first pillar, the following keywords have been selected: “business intelligence” which is the process of acquiring and analysing data to make informed decisions (Chen et al. 2012); “data-driven” was introduced because of the growing interest in data-driven decision making (Provost and Fawcett 2013); “data mining” as it is the set of techniques and methodologies for extracting information from large amounts of data (Provost and Fawcett 2013); finally, the last term is “big data” as it concerns the huge amount of data that could be exploited for applying complex algorithms for decision making (Galetsi et al. 2020). Regarding the second pillar, namely the context of the research, the keyword “healthcare” has been selected. Finally, to retrieve the papers concerning the decision-making process, the keywords “decision making” and “decision support system” were chosen (Chahar 2021; Provost and Fawcett 2013). In order to identify as many records as possible, terms relating to risk and uncertainty were not included because they would have limited the results of the research since they would have excluded papers that do not explicitly address risk and uncertainty in the healthcare organization which is implicitly characterized by them (Brehmer 1992; Kuziemyisky 2016; Lipsitz 2012). For the same reasons, also the keywords concerning the main domains (structure, process, outcome) of the healthcare organization have not been included since they would have excluded papers that do not explicitly use these keywords to describe them. The final research string is shown in Table 1.

The set of retrieved papers contained 4,395 items (2,811 items from Scopus, 1116 items from Web Of Science and 467 from PubMed), running until the 31st of August 2024.

The following phase (i.e., screening) includes the use of exclusion and inclusion criteria to refine the set of papers (summarized in Table 1). Initially, exclusion criteria have been applied based on the type of publication through the databases’ filters, excluding articles not written in English, books and book chapters, conference proceedings, theses and editorials, and obtaining 2374 papers. Then, the set of retrieved papers has been screened in order to remove all the duplicated items, obtaining 1579 unique items, and the retrieved papers were limited to those published in a journal listed in the 2022 ISI Journal Citation Report with an impact factor of 2 or above consistently with other reviews (De Keyser et al. 2019), obtaining 1115 items. It has not been used any filter about the year of publication so as not to limit the results of the literature review to a given period. Then, the screening of titles and abstracts has been conducted excluding the articles not focused on the healthcare organization, or that

Table 1 Exclusion and inclusion criteria for the selection of papers

Search string	Inclusion Criteria	Exclusion Criteria
Keywords: (“data driven” OR “data-driven” OR “big data” OR “business intelligence” OR “data mining”) AND “healthcare” AND (“decision making” OR “decision support system”) •Search in: Title, Abstract and Keywords •Document type: Peer-reviewed Journals •Language: English	•All the articles that concern the use of data to support the decision-making process in the healthcare organizations •Articles that do not limit their analysis only to the development of algorithms.	•Articles not written in English. •Journal Impact Factor < 2 •Articles not published in peer-reviewed journals •Articles not focused on healthcare organizations.

are focused on the development and testing of algorithms (e.g., papers concerning the performance of algorithms), or that are not concerned with the use of data to support the decision-making process. Thus, all the quantitative and qualitative papers that were focused on the exploitation of data to support the decision-making process were included, achieving 475 papers for the eligibility stage. After a thoroughly analysis of the eligible papers, 152 have been used to conduct the literature review.

3.2 The framework

The framework leading this review aims to extract insights from collected papers to eventually answer the research question, based on two dimensions, “Healthcare Organization Domains” and “Risk Management Phases”. In particular, concerning the “Healthcare Organization Domains” the papers included in the review were examined considering the dimension borrowed from the Donabedian’s conceptual model on assessing the quality of healthcare organizations (1988), namely structure, process and outcome. The structure refers to professional and organizational resources associated with the delivery of healthcare services (e.g., material resources, human resources and organizational structure). The process represents the tasks done for the patient (e.g., hospital referrals) and the patient’s activities to receive the services. Finally, the outcome refers to the desired outcome of the care provided by the health practitioners (e.g. patient satisfaction and costs). Concerning the “Risk Management Phases”, the collected papers were examined following the ISO 31,000 (ISO 2018). This dimension of the framework is further distinguished into three categories, that correspond to the phases of the risk management process: the Risk Assessment phase (which is composed of risk identification, risk analysis and risk evaluation), the Risk Treatment phase, and the Risk Monitoring. The Risk Assessment phase aims to identify and analyse risks in order to identify and analyse the probabilities and the magnitude of risk, and to evaluate the risk priorities for guiding further actions. The Risk Treatment phase involves the selection and implementation of optimal strategies for mitigating the risks. Risk Monitoring involves the monitoring of risks to assure the quality and effectiveness of the strategies adopted and the need for new actions to mitigate risks. The framework is depicted in Fig. 2.

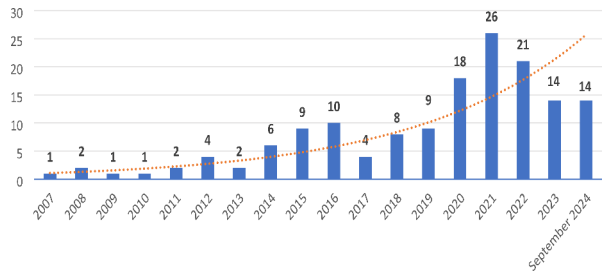
4 Results

The final step of the systematic literature review is the critical analysis of the collected articles. In the first sub-section, the descriptive findings are presented. These include the distribution and evolution of papers across time, sources, and employed research methodologies. In the second sub-section, the thematic analysis and the main findings of the systematic literature review are discussed according to the framework introduced.



Fig. 2 The framework

Fig. 3 Articles distribution per publishing year



4.1 Descriptive statistics

Figure 3 illustrates the distribution of papers per year, showing the increasing interest in the research topic. About three-quarters (70%) of the records were published in the last six years (2019–2024). The first article selected for this review is in 2007, and in the following years, the trend of articles published increased, spiking in 2021. The

topic started its ascent between 2019 and 2020, doubling the number of total articles per year. The first paper appeared in 2007, with modest publication numbers until a noticeable increase began around 2014. Interest in the topic surged between 2019 and 2020, with a doubling in the number of publications per year. The peak occurred in 2021, with 26 papers, which is a 44% increase compared to 2020's 18 papers. Although there was a slight dip in 2022 and 2023, the total number of papers remains relatively high, underscoring the ongoing relevance of this research domain. The upward trend visible in the figure suggests a sustained interest in the topic, validating the need for a thorough and systematic review of the literature.

As for the research methodologies adopted by the selected articles, we can group them into three different categories:

- Quantitative research (90%) – this category is characterised by studies that use quantitative research methods, such as regression models, artificial neural network and data analysis.
- Qualitative research (7%) – in this category are grouped both studies that employ qualitative research methods and conceptual papers, such as case studies.
- Mixed method research (3%) – this category is characterised by studies that use both quantitative and qualitative research methods.

The fact that most of the papers employed a quantitative methodology is consistent with the aim of this research, namely to understand how data can optimize decisions confronted with risk and uncertainty in the main domains (structure, process, outcome) of healthcare organizations.

Table 2 shows the journals in which the selected articles were published. In the table all journals whence at least two selected articles are listed, while all journals from which only one article was selected are grouped under the label “other”. It is of particular interest to note that although that nearly two-thirds of the journals (100 out of 152) are represented by only a single article, it was found that the journals belong mainly to three subject areas i.e., Management and Business, Medicine and Computer Science which shows the interdisciplinarity of this topic. The top contributor

Table 2 Number of papers per Journal

Journal	Number of Papers
Decision Support Systems	9
Scientific Reports	6
IEEE Access	5
Journal of Biomedical Informatics	5
Artificial Intelligence in Medicine	4
Annals of Operations Research	3
Big Data	3
BMC Health Services Research	3
Expert Systems with Applications	3
Health Informatics Journal	3
Industrial Management and Data Systems	3
Journal of Medical Systems	3
Sensors	3
<i>Other</i>	99

is “Decision Support Systems” (9 papers), followed by Scientific Reports (6 papers), IEEE Access (5 papers), the Journal of Biomedical Informatics (5 papers) and Artificial Intelligence in Medicine (4 papers).

4.2 Findings

In the following, the papers selected in the academic literature regarding how data can optimize decisions confronted with risk and uncertainty in the main domains (structure, process, outcome) of healthcare organizations are discussed.

Table 3 shows the distribution of the selected papers according to the two main dimensions of the framework described in Sect. 3, namely the “Healthcare Organization Domains” and the “Risk Management Phases”.

The following sections present the findings based on the defined framework in order to ensure clarity and logical flow, facilitating easier understanding and detailed examination of the potential of data in each domain. We begin with the use of data for risk assessment across the three domains of healthcare organizations. Next, we discuss risk treatment, and finally, we address risk monitoring. Table 4 provides a summary of the main findings identified in the academic literature about the use of data to optimize decisions affected by risk and uncertainty in healthcare organizations.

4.2.1 Risk assessment

4.2.1.1 Structure In the structure domain, data are mainly used to identify and assess the risks that threaten the *optimal use of resources*, (i.e., physical facilities, human resources, equipment and Information and Communication Technology (ICT)) and *organizational activities* (e.g., staff training and resources maintenance).

The use of data in the healthcare industry is opening up the optimisation of the usage and distribution of healthcare resources both through the employment of administrative (e.g., date of hospitalisation, date of discharge, medicines used) and patients’ health data (e.g. blood pressure, glucose level).

For instance, it is possible to estimate the number of patients needing hospitalisation and compare them with the total number of available beds, thus identifying whether there is a risk of shortage of available resources (i.e., beds) (Sebaa et al. 2018). Also the patient’s length of stay affects the resources that will be used by the healthcare organizations: the longer the patient is hospitalised, the more resources will be needed to treat him. Through the data, it is possible to identify and assess which are the factors that increase the length of stay (Kudyba and Gregorio 2010) and

Table 3 Papers categorized according to the framework

		Risk Management phases			
		Risk Assessment	Risk Treatment	Risk Monitoring	Total
Healthcare Orga- nization Domains	Structure	19	30	4	53
	Process	37	22	4	63
	Outcome	13	20	3	36
	Total	69	72	11	152

Table 4 Summary of the main findings

	Risk Assessment	Risk Treatment	Risk Monitoring
Structure	<ul style="list-style-type: none"> •Utilizing data to optimize resource usage and distribution •Estimating patient needs and identifying shortage risks •Assessing factors affecting length of stay and re-hospitalization risks •Predicting healthcare demand and mitigating no-show risks •Predicting disease occurrence and optimizing resource allocation •Managing resources during health emergencies and pandemic impacts 	<ul style="list-style-type: none"> •Optimizing planning, allocation, and management of resources •Minimizing gaps in equipment availability •Reducing readmissions and length of stay •Reducing resource waste •Optimizing medical equipment repair and replacement •Minimizing expected daily resource usage •Optimizing patient scheduling •Optimizing capacity planning for emergency outbreaks •Minimizing waiting times 	<ul style="list-style-type: none"> •Using real-time data transmission for optimizing bed and medical equipment scheduling •Analyzing data on patients' conditions for optimal resource use
Process	<ul style="list-style-type: none"> •Providing tools to assess health risks •Managing chronic diseases •Managing patients during epidemic outbreaks •Extracting information from EHRs and social network data for health insights •Identifying and assesses disease risk •Investigating patient health status •Predicting treatment outcomes •Using clinical trial data to assess treatment risks. 	<ul style="list-style-type: none"> •Using data for assessing and improving patient health conditions •Mitigating risks for specific diseases •Predicting at-risk patients to prevent disease occurrence •Using data to make informed diagnostic decisions and improve guidelines •Identifying optimal treatment plans •Simulating likely course of diseases to optimize care pathway 	<ul style="list-style-type: none"> •Real-time monitoring for daily practice and disease spread planning •Using data from sensors and equipment for remote collection •Supporting diagnosis by monitoring patients' heart conditions
Outcome	<ul style="list-style-type: none"> •Predicting in-patient falls •Predicting long-term survival •Assessing impact of waiting times and overcrowding on mortality •Analyzing social perceptions and patient feelings about healthcare service quality. •Assessing risk of depression affecting health outcomes 	<ul style="list-style-type: none"> •Reducing risks of complications and mortality •Identifying risks of medication nonadherence to minimize adverse outcomes •Supporting personalized treatment to mitigate hospitalization risks and therapy complications •Improving Quality of Life (QoL) post-treatment •Formulating policies to minimize outbreaks spread •Improving scheduling to reduce waiting times and patient dissatisfaction •Ensuring elderly patients' satisfaction with domestic care services 	<ul style="list-style-type: none"> •Focusing on monitoring health status over patient satisfaction •Enabling self-monitoring for patients and remote monitoring for physicians •Providing continuous information about healthcare service outcomes

therefore increase the consumption of resources. Once the hospitalisation period is over, the patient is discharged from the healthcare organization. However, it happens that the patient's health issues have not been fully addressed and after a short time the patient needs to be re-hospitalised (Ryan et al. 2015), thus exposing the healthcare organization to the risk of shortage of available resources. This may lead to an

excessive burden on the healthcare organization's resources, i.e., beds, medicines, and human resources.

Furthermore, data are used to assess the risk of lack of resources by predicting the expected number of arrivals in terms of the number of patients who will need healthcare services, thus allowing to plan the available resources with the expected demand (Feretzakis et al. 2024; Xu et al. 2013). Additionally, data-driven models can assist in managing healthcare processes, as seen in the case of BRCA-mutated women with breast cancer, where a data-driven DSS model provided more accurate cost estimations, enabling more effective resource allocation and decision-making throughout the treatment process (Basile et al. 2023).

It is also possible to exploit data to assess the risk of waste of healthcare resources, by predicting the number of patients who do not turn up for appointments with physicians (i.e., no-show risk patients). These missed appointments imply a waste of resources that had been already scheduled, such as human resources (i.e. doctors and nurses) and the equipment needed for the visit (Simsek et al. 2020, 2021). Data have been often used for the prediction of patient disease occurrence (Phillips-Wren et al. 2008) and population disease occurrence (Hrovat et al. 2014). This can support the identification of healthcare services that will be needed, thus avoiding un-useful service (i.e. risk of waste), and permit the evaluation of the best possible allocation of resources, thus preventing the indiscriminate consumption of resources (i.e. risk of shortage and waste of resources). For instance, data can be used for assessing resource consumption associated with lung cancer (Phillips-Wren et al. 2008) or for supporting the assessment of the risk of resources' waste determined by the unbalanced daily workload management, by comparing the scheduling and the expected demand (Perez et al. 2016).

Data can be also particularly useful in assessing risks faced by healthcare structures during health emergencies such as the COVID-19 pandemic (Kellner et al. 2023; Sarkar et al. 2021), as shown in the work of Kellner et al. (2023), where data from social networks like Twitter and Reddit were used to predict case numbers and hospitalization rates, providing early insights compared to official reports. Also, data have been particularly useful to assess the risk of shortage and waste of healthcare resources determined by the occurrence of COVID-19, predicting the daily number of patients (Manevski et al. 2020) and vaccine willingness (Riad et al. 2021). Some types of healthcare services were provided in smaller numbers due to the needs created by the COVID-19 pandemic, resulting in an imbalance in resource use planning. In this context, data allow the identification of services that have been provided in lower numbers due to the pandemic and thus the associated resources. As a result, it is possible to assess the risk of shortage and abundance of resources thus supporting decisions on the sizing of resources required (Segui et al. 2021).

Finally, concerning organizational activities, data can support the assessment of risks that affect payment methods and medical equipment maintenance. Healthcare organizations in many countries are paid for the delivery of healthcare services provided by public healthcare organizations through the health reimbursements, i.e. an amount of money paid for services and regulated by national health systems. The procedures for claiming these reimbursements can be carried out improperly and cause problems such as late payment or delayed diagnostic tests. In this context, data enable

the analysis of these processes and identify the causes of delayed reimbursements (Gerhardt et al. 2018). While, concerning medical equipment maintenance, data can be used to evaluate and predict the priority of their preventive maintenance, corrective maintenance, and replacement programs in order to assess the risk of resource shortage (Zamzam et al. 2021), as demonstrated by Rahman et al. (2023) with the use of a data-driven machine learning model to predict medical device failure, optimizing maintenance and reducing both intervention features and training time.

4.2.1.2 Process Data are used to identify, assess and prioritise risks at almost every stage of the *care process*: disease prevention, assessment of patient history where genetic and family status is studied, diagnosis of disease, selection of treatments, and prognosis (Croft et al. 2015). In particular, the use of data to assess risks in the process domain supports the decision-making process in both cross-sectional ways and specific stages of the care process. In the first case, physicians and managers of healthcare organizations have provided data-driven tools that make information and knowledge available to assess the health risks of patients without limiting the focus to a specific stage of the care process (Chignell et al. 2013; El-Sappagh and El-Masri 2014; Jung et al. 2015; Lin et al. 2019a, b; Raj et al. 2020). For instance, tools that collect and make available administrative and health data have been proposed to support physicians in managing patients affected by chronic disease (i.e. a health condition or illness that has long-term or persistent consequences, or a sickness that develops over time) (Dag et al. 2022; Yu et al. 2022), or to support the managers of healthcare organizations during epidemic outbreaks (Liu et al. 2016). Data-driven tools extract this information usually from data storage in a systematized collection of patient and population health and administrative data in digital format, the so-called electronic health records (EHRs) (Cascini et al. 2021; Dai et al. 2022; Zhang et al. 2011). To assist the decisions of physicians and healthcare managers in the risk assessment of patients' health status, it is possible to extract information from the clinical notes data stored in these EHRs (Groenhof et al. 2019; Song et al. 2024). For instance, Chen et al. (2024) leveraged electronic health records from 4540 patients to predict tuberculosis co-infection among HIV patients, demonstrating the value of EHRs in epidemiological investigations and clinical decision-making. In addition, it is also possible to extract insights about patients' health from social network data (Fiaidhi 2020). Therefore, data-driven tools are very relevant for healthcare organizations since they help to manage, analyse, and use the information to assess the risks that threaten patients' health.

Data can also be used to support the decisions of physicians and managers along specific stages of the care process. In the disease prevention stage, data are used to support physicians in identifying and assessing the risk that a patient may develop or be affected by a certain disease (Y.-K. Lin et al. 2017; Mioshi et al. 2014; Morales-Sánchez et al. 2024; Nasir et al. 2019), e.g. to support physicians in identifying patients at risk of acute or chronic infection disease (Zhou et al. 2022), or for the early detection of chronic disease patients (Moreira et al. 2020), at risk of heart failure disease (Nagamine et al. 2022) or for identifying patients with neurological or neuromuscular disorders (Ramon-Gonen et al. 2023). Considering the diagnosis stage,

data can support physicians in investigating the health status of patients (Baytas et al. 2016; Malik et al. 2020), e.g. assessing the risk of heart disease (Bashir et al. 2021; Singh et al. 2021) or diabetes occurring (Bukhari et al. 2021). While, data can be used for the treatment selection stage by predicting the likely outcomes both to assess risks for patients associated with specific treatments (Ejaz et al. 2022; Gongora-Salazar et al. 2022; Pinto et al. 2021) and to describe treatment expected outcomes (Delen et al. 2012; Fan et al. 2022). Moreover, physicians can use information extracted from clinical trials to assess the risk of exposure for patients to a treatment (Kuang et al. 2020; Lambay and Mohideen 2024; Mayer et al. 2021), e.g., to predict the risk of death in kidney transplant recipients (Topuz et al. 2018).

4.2.1.3 Outcome The analysis of selected papers reveals that data have been used in healthcare to enable the identification and assessment of *risks to patient health outcomes* and for assessing the *risk of patient dissatisfaction*.

Concerning patient health outcomes, data have been adopted for identifying and assessing the risks that the healthcare service or interventions have on the health status of patients (Asaria et al. 2016; Demirbaga and Aujla 2022; Kaykicioglu et al. 2024; Sarp et al. 2023), for example, to predict the probability of an in-patient falls (Lindberg et al. 2020) or to predict the long-term survival outcomes after heart transplantation (Dag et al. 2016) or to uncover disease subtypes and predict distinct risks of in-hospital mortality or treatment escalation in COVID-19 patients (Rangelov et al. 2023). Data can also enable the assessment of the impact of specific organizational characteristics such as the waiting time and overcrowding on hospitalised patient mortality (Clark and Normile 2012; Soffer et al. 2021).

Data have been also exploited for assessing the risks related to the satisfaction of the patients, by analysing social perceptions of healthcare service quality (Lee et al. 2021a, b) or by analysing patients' feelings and emotions about their experiences in hospitals from online reviews (Abirami and Askarunisa 2017). Through the use of data, it is also possible to assess the risk of depression that can impact the final health outcome (Jin et al. 2015). Similarly, Al Nuairi et al. (2023) developed a data-driven Bayesian belief network model to explore the impact of provider-related factors on patient experience, offering insights into how these factors influence overall satisfaction.

4.2.2 Risk treatment

4.2.2.1 Structure The analysis of selected papers reveals that data are widely used in healthcare to pursue strategies to reduce some *risks that may affect the physical and organizational settings* where healthcare occurs.

In particular, several scholars have used data for treating shortage risks of human, technical and organizational resources by providing information and knowledge to optimize their planning, allocation and management (Abroshan et al. 2022; Bianchin et al. 2022; Debjit et al. 2022; Delias et al. 2015; Kraus et al. 2024; Lavrač et al. 2007). For example, it is possible to extract or predict information (e.g., number of

admissions) that allows the minimization of gaps in the availability of equipment and supplies in healthcare facilities (Pakravan and Johnson 2021; Qiu et al. 2020; Rios-Zertuche et al. 2020), as well as predict patient service times in outpatient clinics, enabling more efficient scheduling and resource management (Golmohammadi et al. 2023) or optimize resource sharing between healthcare facilities during emergency situations (Zhang et al. 2024). It is also possible to use the data to mitigate the risk of shortage of available resources by optimising the number of readmissions (Demir 2014; Y.-W. Lin et al. 2019a, b; Zhu et al. 2015), and the length of stay (Dwyer-Matzky et al. 2020), which increase the burden on healthcare resources. For instance, data have been used for the simulation of the length of stay of patients, by supporting physicians and managers of healthcare organizations to reduce the risk of resources waste (Kovalchuk et al. 2018). Another way to mitigate the risk of a shortage of available resources is the use data to optimize decisions about medical equipment repair and replacement (Liao et al. 2021). In addition, data helps mitigate the risk of wasted resources by minimizing daily expected usage across various resources such as intensive care unit and operating room surgeries (Rabbani et al. 2023; Rath et al. 2017; Wang et al. 2023) as demonstrated by W. Zhang and Li (2024) to accurately predict ICU resource needs and avoid excessive reserves during health emergencies. This is achieved by optimizing the scheduling of patients (Mandelbaum et al. 2020; Nallathamby et al. 2021) and by assigning patients to physicians in an emergency department based on resource availability (Rosemarin et al. 2021). Still considering the mitigation of waste of resources, physicians and managers can exploit data to select new technologies and treatments that create real value in healthcare organizations (Thompson et al. 2016). More recently, the use of data to support the mitigation of the aforementioned risks (i.e., risk of shortage and waste of available resources) was considered during the COVID-19 pandemic. In this emergency context, available resources were scarce and data were useful to optimise the capacity planning of the COVID-19 vaccination, minimising the waiting time by using data to optimise the locations of the medical hubs, the distribution of the available vaccines and healthcare professionals (Z.-Y. Chen et al. 2023; Demir 2014; Le Lay et al. 2024; Markhorst et al. 2021). Machine learning-based forecasting models have also been applied to predict the progression of COVID-19, assisting healthcare professionals in optimizing resource management and patient care amidst the uncertainty (Devarajan et al. 2023). Moreover, data can be used also to optimise the work of healthcare staff, for instance, Poly et al. (2020) developed a machine learning model to predict the healthcare staff fatigue generated.

4.2.2.2 Process Some of the selected papers discuss the use of data for mitigating *issues affecting the stages of the care process*, both cross-sectional and in specific stages, as observed in the [risk assessment](#) section.

Concerning the cross-sectional way, data are used not only to get useful information for the assessment of health patients' conditions but also to improve it by promoting evidence-based medicine. Following these aims, physicians and managers use data to mitigate risks in healthcare organizations for particular disease conditions such as the cardiovascular ones (Raghu et al. 2015). Decision-makers of healthcare

organizations can also mitigate risks in the specific stages of the care process. The disease-prevention stage is widely supported by data. Indeed, in this stage, physicians can use data to predict patients at-risk of disease occurrence and take decisions to prevent it (Dreischulte et al. 2012; Duan et al. 2011; Laila et al. 2022; Layeghian Javan and Sepehri 2021; Wu et al. 2022); for instance using clinical notes to predict blood flow disorders (Afzal et al. 2018), using data to enable the early identification of patients to optimize the outcomes of cancer screening (Sharma et al. 2016) or using electronic health records for the early prediction of septic shock (Wang et al. 2022).

In the diagnosis stages, physicians can use the information extracted from the data to make more informed decisions (Shoaib et al. 2024; Wong-Lin et al. 2020) and even to improve the existing diagnosis guidelines (Antonelli et al. 2012), for instance, data can be used to discover the conditions under which decision-making strategies produce undesired or suboptimal outcomes (Meyer et al. 2014). Concerning the selection of treatment stage, physicians can identify the optimal treatment plan based on the data about the attributes of the patients (Carey et al. 2022; Chi et al. 2022; Newland et al. 2018), e.g. to optimize anaemia treatment in hemodialysis patients (Escandell-Montero et al. 2014). Furthermore, physicians can optimize the selection of treatments using data to consider complexity among the many variables that can influence the outcome, considering likely adverse side effects with a specific therapy (Sun et al. 2021) or in presence of comorbidities (Zolbanin et al. 2015). In this context, data allows physicians to make a personalised treatment plan, taking into account the patients' chronic diseases such as obesity, hypertension, and hyperglycemia (Valero-Ramon et al. 2020). Finally, physicians can use data to optimize treatment selection decisions by getting information about the likely course of the disease, e.g. from the likely course of oncological patients (Ji et al. 2019; Johnson et al. 2020).

4.2.2.3 Outcome The use of data for risk treatment in the outcome domain focuses mainly on *patients' health and satisfaction*. The use of data enables decisions that mitigate patients' health risks, mostly by reducing the risk of complications and mortality (Bhatti et al. 2019; Kamala et al. 2021; Khalili et al. 2024; Mahajan et al. 2019; O'Grady et al. 2021; Tao et al. 2021). For instance, data can be used to mitigate the risk of maternal and fetal mortality and morbidity rates by predicting the women at risk of developing gestational diabetes (Kopanitsa et al. 2023; Moreira et al. 2018) and also to minimise adverse health outcomes by analysing the reasons for medication nonadherence (i.e., the practice of patients failing to take medications as prescribed) (Xie et al. 2022). Data can also support the choice of personalised treatment to mitigate risks of hospitalization related to specific diseases (Yeh et al. 2011) or complications and side effects that may occur with certain therapies, e.g. chemotherapy (Dara et al. 2024; Mosa et al. 2020; Shirazibeheshti et al. 2023). Also, the Quality of Life (QoL) of the patients after treatments is of paramount importance for healthcare organizations, and data can support the decision of physicians and managers in order to reduce the incidence of decisions-error (Faria et al. 2015). While, during the COVID-19 outbreak several scholars and organizations proposed tools to provide data-driven real-time insights to improve patients' health outcomes (Tolk et al. 2021). As an example, decision-makers used data both to formulate social distancing policy to minimize the risks of infection spread (K. Chen et al. 2021a, b) and to

predict or diagnose diabetes patients to decrease the mortality rate of COVID-19, being particularly sensitive patients (Li et al. 2021).

Finally, concerning the satisfaction of the patients, for example, data can be exploited to mitigate patient dissatisfaction by improving the scheduling of healthcare activities mitigating the risks of long waiting times (Cho et al. 2019). Moreover, it is also possible to use data to support physicians in the decision-making process to ensure elderly patients' satisfaction in delivering domestic care services (Lam et al. 2021). Another application to optimize the satisfaction of patients is the recommendation of the optimal insurance healthcare plan based on the likely expected healthcare services needs (Stein 2016).

4.2.3 Risk monitoring

4.2.3.1 Structure Few of the selected works focused on the use of data for risk monitoring in the structure domain, primarily focusing on the optimal use of resources. In particular, data are used for beds and medical equipment by providing insights through real-time data transmission for the optimisation of resources scheduling and management (Demirdöğen et al. 2022; Picozzi et al. 2024; Wautelet et al. 2018). Furthermore, it is also possible to exploit data to monitor the impact of new organizational models for the optimal use of healthcare resources by analysing data concerning patients' conditions (Vanasse et al. 2020).

4.2.3.2 Process Physicians and managers of a healthcare organization can use data in the process domain to monitor risks related to patients' health, by supporting the decisions throughout the care process.

The whole care process can benefit from data to carry out monitoring activities in real-time, for instance, to support healthcare professionals' daily practice for patients with specific diseases (Mathe et al. 2009) or to monitor the spread of diseases to enable managers and physicians to plan actions (Chinnaswamy et al. 2019). In addition, the development of sensors and equipment for the real-time and remote collection of patient health data is enabling care process activities to be carried out without patients needing to visit facilities, as for elderly patients (Wong et al. 2017). For example, it is possible to support the diagnosis stage analysing data from the monitoring of patients' heart conditions (Hu et al. 2008).

4.2.3.3 Outcome Few papers that exploit data for risk monitoring activities in the outcome domain have been identified, and their focus is mainly concerned with the monitoring risks related to the patients' health status rather than their satisfaction. Both patients and physicians can benefit from the use of data collected through sensors and remote technologies for risk monitoring activities. From the patient perspective, data enable self-monitoring of health, while from the perspective of physicians, data enable the monitoring without physically visiting patients. Indeed, the main goal of using data for risk monitoring in this domain is to get continuous information

about the outcome of healthcare services (Rathore et al. 2016). For instance, it has been proposed a model to monitor patients' outcomes in terms of wellness by considering lifestyle variables (Agarwal et al. 2016).

5 Future research directions

In the following, the future research directions will be presented grounded on the findings with a focus on the specific dimensions of the framework defined in Sect. 3 (Fig. 2).

5.1 Risk assessment

In the *structure* domain, data support the identification and assessment of risks that threaten the *optimal use of resources* and *organizational activities* in the healthcare organization. Focusing on the resources, data are used to identify and assess the risk of resource shortage and waste by making predictions on variables such as the length of stay (Kudyba and Gregorio 2010) or the probability of re-hospitalisation (Ryan et al. 2015). While, concerning organizational activities, fewer studies have been identified, and they are focused on equipment maintenance (Zamzam et al. 2021) and payment methods (Gerhardt et al. 2018). However, the risk assessment to support staff training is particularly relevant in the current era, where several new technologies and innovations have been introduced in the healthcare industry. In this context, the use of data can enable the identification of improvement areas in healthcare staff education in order to detect and assess the risk of defective training. For instance, supporting educators in the training of healthcare professionals for the use of technological innovations, such as artificial intelligence for diagnostic or telemedicine for remote delivery of healthcare services. Despite the relevance, it has not been identified qualitative or quantitative research that used data to identify areas for improvement to guide educators, managers and decision-makers in planning training pathways for healthcare professionals. *For these reasons, more research on the use of data for supporting educational managers in the development of staff training pathways would be needed.*

In the *process* domain, data are widely used to optimise decisions affected by risk and uncertainty. It turns out that data enable risk assessment both in cross-sectional ways (Liu et al. 2016) and at specific stages of the care process. In particular, several scholars have explored risk assessment in the different stages of the care process: identification of patients at risk (Mioshi et al. 2014; Nasir et al. 2019), diagnosis (Bashir et al. 2021; Malik et al. 2020), choice of treatment (Fan et al. 2022). The relevance of these results lies in the fact that through data it is possible to identify and assess risks at various stages of the care process, thereby following the paradigm of providing the right care at the right time and in the right place. In this review, the findings highlighted how data support decisions about the stages of the care process, and thus disease-specific processes i.e., diagnosis, choice of treatment, etc. However, the process domain in the Donabedian model is not focused only on the disease-specific

processes, but also on processes that are not directly focused on the patient's disease, e.g., physician-patient communication and nursing care. In recent years, such processes are gaining increasing interest since they can further improve healthcare services delivery. In order to support the managers and physicians in these processes, a new paradigm is spreading. This paradigm involves the *engagement of patients* throughout the care process. In this context, the role of the patient would no longer be a passive one but rather he/she can become an active member of the whole care process, as the ultimate expert in his or her own physical and psychological health status. For instance, the engagement of patients could improve the commitment to follow a suggested treatment through the improvement of physician-patient communication. Despite the *non-disease-specific* processes seems a promising topic to improve the delivery of healthcare services, papers that use data to address it have not been identified in the literature. *Thus, it is recommended to explore through the conceptual and qualitative paper the use of data for assessing risks that affect non-disease-specific processes.*

Data are also used to support the risk assessment phase in the *outcome* domain by assessing the risks associated with decisions that impact both patients' health outcomes (Asaria et al. 2016; Clark and Normile 2012) and their satisfaction (Abirami and Askarunisa 2017; H. J. Lee et al. 2021a, b). Focusing on patients' health outcomes, both data related to patient characteristics and organizational characteristics are used to identify and assess risks. However, the use of data to identify and assess risks related to errors in healthcare service delivery was not identified in this review. This issue has become increasingly relevant in recent years, and the World Health Organization (WHO) estimates that adverse health outcomes resulting from errors in the delivery of healthcare services are one of the leading causes of death worldwide (World Health Organization, 2023). The use of data seems to be promising to identify these risks in healthcare organizations to support the decisions of physicians and managers. For instance, data could enable the early or even real-time identification of errors in care procedures. *For the aforementioned reasons, the authors suggest investigating how data can help healthcare organizations in identifying and assessing the risk of errors in the delivery of healthcare services.*

5.2 Risk treatment

In the *structure* domain, data are used to mitigate risks that threaten the physical and organizational settings in which healthcare occurs. In this context, data allow for optimised decision-making by dealing with risks of resource shortage (Dwyer-Matzky et al. 2020; Pakravan and Johnson 2021) and resource waste (Nallathamby et al. 2021; Rosemarin et al. 2021). However, in this domain, contrary to the findings in the risk assessment, no studies have been identified that exploit data to support decisions about organizational activities (e.g., staff training, and payment methods). In a complex system such as healthcare, organizational activities play a crucial role in guaranteeing that healthcare organizations work properly in order to deliver optimal services. And as already stated in this review, risk and uncertainty can lead to sub-optimal decision-making that will impact health outcomes. For instance, planning, scheduling and control of the patient flow are organizational activities in which

decisions are taken under risky and uncertain conditions. Healthcare managers can use data to deal with these conditions, e.g., exploiting data to support the planning and scheduling of activities to manage effectively the flow of patients by forecasting what and how much needs to be done (i.e., planning) and defining who and when the activities will be performed (i.e., scheduling). Thus, supporting the mitigation of risk and uncertainty in decision-making is a relevant issue also for organizational activities. *Future research should be performed to explore the use of data for optimising decisions in conditions of risk and uncertainty about organizational activities.*

In the *process* domain, data assist the decision-making by supporting physicians in the mitigation of risks throughout the care process: identification of patients at-risk (Afzal et al. 2018; Layeghian Javan and Sepehri 2021), diagnosis (Meyer et al. 2014; Wong-Lin et al. 2020), selection of treatment (Newland et al. 2018). In particular, in the selection of treatment, it turns out that data about the characteristics of patients are used, i.e. health-related data. While, other kinds of data such as those related to the patients' satisfaction or the economic costs for equivalent treatments, are still overlooked. These data could enable physicians to take more aware and holistic decisions by considering not only the health-related perspective but also other ones, such as the patient's satisfaction. *It is suggested that further research should explore the combination of health-related data with other types of data, such as satisfaction ones, to support the selection of treatment.*

Regarding the treatment of risk in the *outcome* domain, data are used to treat the risk for patients' health, e.g., to reduce complications and mortality rates (Kamala et al. 2021) and for ensuring the patients' satisfaction (Cho et al. 2019; Lam et al. 2021; Stein 2016). As emerged for risk assessment, the use of data to mitigate the risk of errors in healthcare service delivery was not identified. However, it is possible to use data to support physicians and healthcare managers to reduce the frequency of these errors. An example could be the use of data to reduce the incidence of unnecessary hospitalisation. The use of healthcare resources for unnecessary admissions deprives other patients who need those resources, leading to delays in treatment and ultimately suboptimal results. Although such an application has not been found in the literature, it is possible to predict necessary hospitalisations through data (Yeh et al. 2011). Thus, data could be a valuable support in reducing the risk of these errors. *For this reason, scholars should deeper explore the role of data to mitigate the risk of errors in the delivery of healthcare services.*

5.3 Risk monitoring

Concerning risk monitoring, few papers have been identified (8), despite the relevance of the topic in healthcare organizations. Indeed, nowadays, the spreading of mobile and wearable technologies is enabling the collection of data to monitor several risks that threaten the domain of healthcare organizations. In the domain of *structure*, scholars have focused their attention on monitoring the use of physical facilities and equipment to support scheduling (Wautelet et al. 2018) and testing new organizational models that ensure the optimal use of resources (Vanasse et al. 2020). No studies have been identified that use data to monitor the status of human resources in healthcare organizations, a topic that has become particularly relevant after the COVID-19 pandemic. In this context, physicians and healthcare staff have

been overburdened by the high number of hospitalisations and high-risk conditions related to the spread of the virus. Data could make it possible to monitor the condition of human resources, both from the psychological and physical points of view, so that action could be taken to ensure the optimal state of human resource conditions in healthcare organizations, reducing the so-called staff “burnout”. *It is recommended to explore through the conceptual and qualitative paper the use of data for monitoring risks that affect the physical and mental conditions of human resources.*

In the *process* domain, scholars have focused on the use of data to support the monitoring of risks that may affect the different phases of the care process. An increasing field of application is the use of data to perform remote monitoring activities of the care process, such as in the prognosis and diagnosis (Chinnaswamy et al. 2019; Wong et al. 2017). The delivery of remote healthcare services based on monitoring patients’ health status is considered relevant both in cases where patients are not able to reach the healthcare organizations due to patient characteristics (e.g., elderly patients) and when circumstantial conditions make it hard to reach them (e.g., pandemics). An interesting and overlooked field of application for risk monitoring activities through data is the non-urbanised areas (i.e., rural areas). These areas are characterised by the scarce presence of well-equipped healthcare organizations, so it is difficult for patients to reach the healthcare facilities. The use of data for carrying on remote activities would overcome this issue, ensuring the delivery of healthcare services. *Thus, a viable future research direction is the development of empirical studies for the development of systems that allow the use of data to monitor the risk of patients’ health status in rural areas.*

In the *outcome* domain, the authors identified papers in which data are used to monitor the outcome of healthcare services on patients’ health status, rather than monitoring patients’ satisfaction. Patients’ satisfaction is significant both for understanding performance in healthcare service delivery and because it affects patients’ health status. *For these reasons, researchers should explore the use of data to monitor patient satisfaction through empirical studies in healthcare organizations.*

Finally, the review unveils that data are not explicitly integrated with the risk management phases in the healthcare decision-making processes, consistent with the findings of the review performed by Dicuonzo et al. (2021) about the use of big data and artificial intelligence for supporting the risk management. This strengthens the relevance of this systematic literature review since it provides the first overview of the integration of data for decision-making under conditions of risk and uncertainty to ensure the delivery of high-quality services in healthcare organizations. In addition, the lacks identified in the literature encourage the development of new scientific projects to investigate the optimal integration of data into established risk management techniques.

6 Conclusion

This study aimed to provide evidence on the use of data to optimise decisions in conditions of risk and uncertainty considering three healthcare organizations’ domains.

The contributions of this study are twofold: theoretical and managerial. This study contributes to the extant literature on the use of data to optimise decisions under risk and uncertainty in healthcare organizations. It should be mentioned that the use and potential

of data in the healthcare setting have been already discussed in the literature (Islam et al. 2018; Salazar-Reyna et al. 2020) and also some studies have focused on the risks and uncertainty conditions systematically reviewing established risk management techniques (Liu et al. 2020) or on investigating emerging risks (Sardi et al. 2020) in the healthcare setting. However, to the best of the authors' knowledge, this is the first systematic literature review to address the problem of decision-making affected by risk and uncertainty in healthcare organizations from the perspectives of risk management, data and quality of healthcare services. Moreover, the analysis of the findings made it possible to identify gaps in the literature and thus, future research directions were defined to address them with regard to the phases of risk management and the domains of healthcare organizations where decisions are made.

Concerning managerial contribution, physicians and healthcare managers daily deal with decisions affected by risk and uncertainty in all the domains of healthcare organizations. Thus, this study offers some insights for healthcare professionals who are interested in understanding how to use data to optimize decisions affected by risk and uncertainty. For instance, healthcare managers may understand how data can support their decisions in both planning the needs of healthcare resources and the scheduling of activities, mitigating several risks, such as the risk of shortage. While, physicians can use the findings of the study to gain a better understanding of how data-driven decisions can support them at various stages of the care process, e.g., in the selection of treatment data allow them to consider at the same time many variables that can cause adverse outcomes. Policy-makers can also use the findings of this review to guide decisions affected by risk and uncertainty, such as the decisions about social distancing policies in healthcare facilities during the Covid-19 pandemic.

Although the authors used a systematic methodology to conduct this review, this work has some limitations. The literature analyzed was identified through specific keywords, so papers that did not use the keywords defined have not been included in this review. However, to mitigate this issue, the authors tested different sets of keywords in order to check that papers related to the topic were only excluded after a deeper analysis of the title and abstract and not a priori. One could say that not including keywords related to risk and uncertainty is a limitation of this study. However, it is only an apparent limitation as it turns out that in the healthcare setting the conditions of risk and uncertainty are not explicitly claimed in papers so a research keyword-based would have not retrieved these papers. For this reason, the authors preferred not to add these keywords in order to avoid a reduction of the set of papers, rather retrieve more papers and remove them only after reading them thoroughly. Furthermore, the filters used to perform the research on the database could also be considered a limitation. Indeed, consistently with other systematic reviews, only articles and reviews from peer-reviewed journals written in English were considered in this research. While, concerning the databases used for retrieving the papers, the authors selected Scopus, Web of Science and PubMed in order to ensure the review's comprehensiveness.

In conclusion, the authors expect that the framework proposed to discuss the papers, the findings and the future research directions represent a valuable contribution for scholars to carry on future research, as they could help to shape the increasing and progressing stream of research about the use of data to support healthcare decision-makers to take decisions affected by risk and uncertainty.

Funding Open access funding provided by Université Paris-Saclay.

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Data availability The data that support the findings of this study are available from the corresponding author upon request.

Declarations

Competing interests The authors have no competing interests to declare that are relevant to the content of this article.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Abirami AM, Askarunisa A (2017) Sentiment analysis model to emphasize the impact of online reviews in healthcare industry. *Online Inf Rev* 41(4):471–486. <https://doi.org/10.1108/OIR-08-2015-0289>. Scopus
- Abroshan M, Yip KH, Tekin C, van der Schaar M (2022) Conservative policy construction using Variational Autoencoders for Logged Data with missing values. *IEEE Trans Neural Networks Learn Syst*
- Aceto G, Persico V, Pescapé A (2020) Industry 4.0 and Health: internet of things, Big Data, and Cloud Computing for Healthcare 4.0. *J Industrial Inform Integr* 18:100129. <https://doi.org/10.1016/j.jii.2020.100129>
- Afzal N, Mallipeddi VP, Sohn S, Liu H, Chaudhry R, Scott CG, Kullo IJ, Arruda-Olson AM (2018) Natural language processing of clinical notes for identification of critical limb ischemia. *Int J Med Informatics* 111:83–89. <https://doi.org/10.1016/j.ijmedinf.2017.12.024>. Scopus
- Agarwal A, Baechle C, Behara RS, Rao V (2016) Multi-method approach to wellness predictive modeling. *J Big Data*. <https://doi.org/10.1186/s40537-016-0049-0>
- Ahmed Z, Kim M, Liang BT (2018) MAV-clic: management, analysis, and visualization of clinical data. *JAMIA Open*. <https://doi.org/10.1093/jamiaopen/ooy052>
- Al Nuairi A, Simsekler MCE, Qazi A, Sleptchenko A (2023) A data-driven bayesian belief network model for exploring patient experience drivers in healthcare sector. *Ann Oper Res*. <https://doi.org/10.1007/s10479-023-05437-9>
- Al-Ali AR, Zualkernan IA, Rashid M, Gupta R, Alikarar M (2017) A smart home energy management system using IoT and big data analytics approach. *IEEE Transactions on Consumer Electronics*, 63(4), 426–434. *IEEE Transactions on Consumer Electronics*. <https://doi.org/10.1109/TCE.2017.015014>
- Antonelli D, Baralis E, Bruno G, Chiusano S, Mahoto NA, Pettrigni C (2012) Analysis of diagnostic pathways for colon cancer. *Flex Serv Manuf J* 24(4):379–399. <https://doi.org/10.1007/s10696-011-9095-2>. Scopus
- Asaria M, Walker S, Palmer S, Gale CP, Shah AD, Abrams KR, Crowther M, Manca A, Timmis A, Hemingway H, Sculpher M (2016) Using electronic health records to predict costs and outcomes in stable coronary artery disease. *Heart*. <https://doi.org/10.1136/heartjnl-2015-308850>
- Babrow AS, Kasch CR, Ford LA (1998) The many meanings of uncertainty in illness: toward a systematic accounting. *Health Commun* 10(1):1–23. https://doi.org/10.1207/s15327027hc1001_1

- Bashir S, Almazroi AA, Ashfaq S, Almazroi AA, Khan FH (2021) A Knowledge-based Clinical Decision Support System utilizing An Intelligent Ensemble Voting Scheme for Improved Cardiovascular Disease Prediction. *IEEE Access*, 1–1. IEEE Access. <https://doi.org/10.1109/ACCESS.2021.3110604>
- Basile LJ, Carbonara N, Pellegrino R, Panniello U (2023) Business intelligence in the healthcare industry: the utilization of a data-driven approach to support clinical decision making. *Technovation* 120:102482. <https://doi.org/10.1016/j.technovation.2022.102482>
- Basile LJ, Carbonara N, Panniello U, Pellegrino R (2024a) How can technological resources improve the quality of healthcare service? The enabling role of big data analytics capabilities. *IEEE Transactions on Engineering Management*, 1–12. *IEEE Transactions on Engineering Management*. <https://doi.org/10.1109/TEM.2024.3366313>
- Basile LJ, Carbonara N, Panniello U, Pellegrino R (2024b) The role of big data analytics in improving the quality of healthcare services in the Italian context: the mediating role of risk management. *Technovation* 133:103010. <https://doi.org/10.1016/j.technovation.2024.103010>
- Baytas IM, Lin K, Wang F, Jain AK, Zhou J (2016) PhenoTree: Interactive Visual Analytics for Hierarchical Phenotyping from Large-Scale Electronic Health Records. *IEEE Trans Multimedia* 18(11):2257–2270. <https://doi.org/10.1109/TMM.2016.2614225>. Scopus
- Bhatti UA, Huang M, Wu D, Zhang Y, Mehmood A, Han H (2019) Recommendation system using feature extraction and pattern recognition in clinical care systems. *Enterp Inform Syst*. <https://doi.org/10.1080/17517575.2018.1557256>
- Bianchin G, Dall'Anese E, Poveda JI, Jacobson D, Carlton EJ, Buchwald AG (2022) Novel use of online optimization in a mathematical model of COVID-19 to guide the relaxation of pandemic mitigation measures. *Sci Rep* 12(1):1–14
- Brehmer B (1992) Dynamic decision making: human control of complex systems. *Acta Psychol* 81(3):211–241. [https://doi.org/10.1016/0001-6918\(92\)90019-A](https://doi.org/10.1016/0001-6918(92)90019-A)
- Bukhari MM, Alkhomees BF, Hussain S, Gumaei A, Assiri A, Ullah SS (2021) An Improved Artificial Neural Network Model for Effective Diabetes Prediction. *Complexity*, 2021. Scopus. <https://doi.org/10.1155/2021/5525271>
- Cagliano AC, Grimaldi S, Rafele C (2011) A systemic methodology for risk management in healthcare sector. *Saf Sci* 49(5):695–708. <https://doi.org/10.1016/j.ssci.2011.01.006>
- Carey J, Wang T, Yang L, Chan W, Whelan B, Silke C, O'Sullivan M, Rooney B, McPartland A, O'Malley G (2022) Conceptual design of the dual X-ray absorptiometry health informatics prediction system for osteoporosis care. *Health Inf J* 28(1):14604582211066465–14604582211066465
- Cascini F, Santaroni F, Lanzetti R, Failla G, Gentili A, Ricciardi W (2021) Developing a data-driven approach in order to improve the safety and quality of patient care. *Front Public Health* 9. Scopus <https://doi.org/10.3389/fpubh.2021.667819>.
- Chahar R (2021) Computational decision support system in healthcare: a review and analysis. *Int J Adv Technol Eng Explor*. <https://doi.org/10.19101/IJATEE.2020.762142>
- Champion C, Kuziemyky C, Affleck E, Alvarez GG (2019) A systems approach for modeling health information complexity. *Int J Inf Manag* 49:343–354. <https://doi.org/10.1016/j.ijinfomgt.2019.07.002>. Scopus
- Chen H, Chiang RHL, Storey VC (2012) Business Intelligence and Analytics: from Big Data to Big Impact. *MIS Q* 36(4):1165–1188. <https://doi.org/10.2307/41703503>
- Chen P-T, Lin C-L, Wu W-N (2020) Big data management in healthcare: adoption challenges and implications. *Int J Inf Manag* 53. <https://doi.org/10.1016/j.ijinfomgt.2020.102078>. Scopus
- Chen K, Pun CS, Wong HY (2021a) Efficient social distancing during the COVID-19 pandemic: integrating economic and public health considerations. *Eur J Oper Res*. Scopus <https://doi.org/10.1016/j.ejor.2021.11.012>.
- Chen T, Shang C, Su P, Keravnou-Papailiou E, Zhao Y, Antoniou G, Shen Q (2021b) A decision tree-initialised neuro-fuzzy Approach for clinical decision support. *Artif Intell Med* 111. <https://doi.org/10.1016/j.artmed.2020.101986>. Scopus
- Chen Z-Y, Sun M, Han X-X (2023) Prediction-driven collaborative emergency medical resource allocation with deep learning and optimization. *J Oper Res Soc*. <https://www.tandfonline.com/doi/abs/https://doi.org/10.1080/01605682.2022.2101953>
- Chen J, Liu L, Huang J, Jiang Y, Yin C, Zhang L, Li Z, Lu H (2024) LSTM-Based prediction model for Tuberculosis among HIV-Infected patients using structured Electronic Medical records: a retrospective machine learning study. *J Multidisciplinary Healthc*. <https://www.tandfonline.com/doi/abs/https://doi.org/10.2147/JMDH.S467877>

- Chi C-L, Wang J, Yew PY, Lenskaia T, Loth M, Pradhan PM, Liang Y, Kurella P, Mehta R, Robinson JG (2022) Producing personalized statin treatment plans to optimize clinical outcomes using big data and machine learning. *J Biomed Inform* 128:104029
- Chignell M, Rouzbahman M, Kealey R, Samavi R, Yu E, Sieminowski T (2013) Nonconfidential patient types in emergency clinical decision support. *IEEE Secur Priv* 11(6):12–18. <https://doi.org/10.1109/MSP.2013.103>. Scopus
- Chinnaswamy A, Papa A, Dezi L, Mattiacci A (2019) Big data visualisation, geographic information systems and decision making in healthcare management. *Manag Decis*. <https://doi.org/10.1108/MD-07-2018-0835>
- Cho M, Song M, Yoo S, Reijers HA (2019) An evidence-based decision support Framework for Clinician Medical Scheduling. *IEEE Access*. <https://doi.org/10.1109/ACCESS.2019.2894116>
- Clark K, Normile L (2012) Nursing informatics and data collection from the electronic medical record: study of characteristics, factors and occupancy impacting outcomes of critical care admissions from the Emergency Department. *Health Inf J*. <https://doi.org/10.1177/1460458212454023>
- Croft P, Altman DG, Deeks JJ, Dunn KM, Hay AD, Hemingway H, LeResche L, Peat G, Perel P, Petersen SE, Riley RD, Roberts I, Sharpe M, Stevens RJ, Van Der Windt DA, Von Korff M, Timmis A (2015) The science of clinical practice: Disease diagnosis or patient prognosis? Evidence about what is likely to happen should shape clinical practice. *BMC Med* 13:20. <https://doi.org/10.1186/s12916-014-0265-4>
- Dag A, Topuz K, Oztekin A, Bulur S, Megahed FM (2016) A probabilistic data-driven framework for scoring the preoperative recipient-donor heart transplant survival. *Decis Support Syst*. <https://doi.org/10.1016/j.dss.2016.02.007>
- Dag AZ, Akcam Z, Kibis E, Simsek S, Delen D (2022) A probabilistic data analytics methodology based on bayesian belief network for predicting and understanding breast cancer survival. *Knowl Based Syst* 242:108407
- Dai X, Park JH, Yoo S, D'Imperio N, McMahon BH, Rentsch CT, Tate JP, Justice AC (2022) Survival analysis of localized prostate cancer with deep learning. *Sci Rep* 12(1):1–10
- Dal Mas F, Massaro M, Rippa P, Secundo G (2023) The challenges of digital transformation in healthcare: an interdisciplinary literature review, framework, and future research agenda. *Technovation* 123:102716. <https://doi.org/10.1016/j.technovation.2023.102716>
- Dara ON, Ibrahim AA, Mohammed TA (2024) Advancing medical imaging: detecting polypharmacy and adverse drug effects with Graph Convolutional Networks (GCN). *BMC Med Imaging* 24(1):174. <https://doi.org/10.1186/s12880-024-01349-7>
- De Keyser B, Guiette A, Vandenbempt K (2019) On the use of paradox for generating theoretical contributions in management and organization research. *Int J Manage Reviews* 21(2):143–161
- Debjit K, Islam MS, Rahman MA, Pinki FT, Nath RD, Al-Ahmadi S, Hossain MS, Mumenin KM, Awal MA (2022) An Improved Machine-Learning Approach for COVID-19 prediction using Harris Hawks Optimization and feature analysis using SHAP. *Diagnostics* 12(5):1023
- Delen D, Oztekin A, Tomak L (2012) An analytic approach to better understanding and management of coronary surgeries. *Decis Support Syst* 52(3):698–705. <https://doi.org/10.1016/j.dss.2011.11.004>. Scopus
- Delias P, Doumpou M, Grigoroudis E, Manolitzas P, Matsatsinis N (2015) Supporting healthcare management decisions via robust clustering of event logs. *Knowl Based Syst*. <https://doi.org/10.1016/j.knsys.2015.04.012>
- Demir E (2014) A decision Support Tool for Predicting patients at risk of readmission: a comparison of classification trees, logistic regression, generalized additive models, and Multivariate Adaptive Regression splines. *Decis Sci*. <https://doi.org/10.1111/dec.12094>
- Demirbaga U, Aujla GS (2022) MapChain: a blockchain-based Verifiable Healthcare Service Management in IoT-based Big Data Ecosystem. *IEEE Trans Netw Serv Manage*
- Demirdöğen G, Işık Z, Arayıcı Y (2022) Determination of Business Intelligence and Analytics-based Healthcare Facility Management Key Performance indicators. *Appl Sci* 12(2). <https://doi.org/10.3390/app12020651>
- Devarajan JP, Manimuthu A, Sreedharan VR (2023) Healthcare Operations and Black Swan Event for COVID-19 Pandemic: A Predictive Analytics. *IEEE Transactions on Engineering Management*, 70(9), 3229–3243. *IEEE Transactions on Engineering Management*. <https://doi.org/10.1109/TEM.2021.3076603>

- Dicuonzo G, Donofrio F, Galeone G (2021) Big Data and Artificial Intelligence to Support Risk Management: A Systematic Literature Review. In M. S. Chiucchi, R. Lombardi, & D. Mancini (Eds.), *Intellectual Capital, Smart Technologies and Digitalization: Emerging Issues and Opportunities* (pp. 119–130). Springer International Publishing. https://doi.org/10.1007/978-3-030-80737-5_9
- Donabedian A (1988) The quality of care: how can it be assessed? *JAMA* 260(12):1743–1748. <https://doi.org/10.1001/jama.1988.03410120089033>
- Dreischulte T, Grant A, Donnan P, McCowan C, Davey P, Petrie D, Treweek S, Guthrie B (2012) A cluster randomised stepped wedge trial to evaluate the effectiveness of a multifaceted information technology-based intervention in reducing high-risk prescribing of non-steroidal anti-inflammatory drugs and antiplatelets in primary medical care: the DQIP study protocol. *Implement Sci*. <https://doi.org/10.1186/1748-5908-7-24>
- Duan L, Street WN, Xu E (2011) Healthcare information systems: data mining methods in the creation of a clinical recommender system. *Enterp Inform Syst*. <https://doi.org/10.1080/17517575.2010.541287>
- Dwyer-Matzky K, Pachamano D, Tilton V (2020) Accounting for capacity: a real-time optimization approach to managing observation unit utilization. *Nav Res Logist*. <https://doi.org/10.1002/nav.21907>
- Ejaz S, Khan SA, Hussain F (2022) Knowledge and Data-Driven Framework for Designing a computerized physician Order Entry System. *IEEE Access* 10:40953–40966
- El-Sappagh SH, El-Masri S (2014) A distributed clinical decision support system architecture. *J King Saud Univ - Comput Inform Sci*. <https://doi.org/10.1016/j.jksuci.2013.03.005>
- Escandell-Montero P, Chermisi M, Martinez-Martinez JM, Gomez-Sanchis J, Barbieri C, Soria-Olivas E, Mari F, Vila-Francés J, Stopper A, Gatti E (2014) Optimization of anemia treatment in hemodialysis patients via reinforcement learning. *Artif Intell Med* 62(1):47–60
- Fan B, Klatt J, Moor MM, Daniels LA, Sanchez-Pinto LN, Agyeman PK, Schlapbach LJ, Borgwardt KM (2022) Prediction of recovery from multiple organ dysfunction syndrome in pediatric sepsis patients. *Bioinformatics* 38(Supplement1):i101–i108
- Faria BM, Gonçalves J, Reis LP, Rocha Á (2015) A clinical support system based on quality of Life Estimation. *J Med Syst*. <https://doi.org/10.1007/s10916-015-0308-1>
- Feretzakis G, Sakagianni A, Anastasiou A, Kapogianni I, Bazakidou E, Koufopoulos P, Koumpouros Y, Koufopoulou C, Kaldis V, Verykios VS (2024) Integrating Shapley values into Machine Learning Techniques for Enhanced Predictions of Hospital Admissions. *Appl Sci* 14(13) Article 13. <https://doi.org/10.3390/app14135925>
- Ferroni P, Zanzotto FM, Scarpato N, Riandino S, Nanni U, Roselli M, Guadagni F (2017) Risk Assessment for venous thromboembolism in chemotherapy-treated ambulatory Cancer patients. *Med Decis Making* 37(2):234–242. <https://doi.org/10.1177/0272989X16662654>. Scopus
- Fiaidhi J (2020) Envisioning insight-driven learning based on Thick Data Analytics with Focus on Healthcare. *IEEE Access* 8:114998–115004. <https://doi.org/10.1109/ACCESS.2020.2995763>. Scopus
- Florence S, Shyamala Kumari C (2019) Big data and IoT in smart transportation system. *Int J Innovative Technol Exploring Eng*. <https://doi.org/10.35940/ijitee.i7597.078919>
- Galaitis SE, Cegan JC, Volk K, Joyner M, Trump BD, Linkov I (2021) The challenges of data usage for the United States' COVID-19 response. *Int J Inf Manag* 59. Scopus <https://doi.org/10.1016/j.ijinfomgt.2021.102352>.
- Galetsi P, Katsaliaki K, Kumar S (2020) Big data analytics in health sector: theoretical framework, techniques and prospects. *Int J Inf Manag* 50:206–216. <https://doi.org/10.1016/j.ijinfomgt.2019.05.003>. Scopus
- Gerhardt R, Valiati JF, dos Canto JV (2018) An investigation to identify factors that lead to Delay in Healthcare reimbursement process: a Brazilian case. *Big Data Res* 13:11–20. <https://doi.org/10.1016/j.bdr.2018.02.006>
- Golmohammadi D, Zhao L, Dreyfus D (2023) Using machine learning techniques to reduce uncertainty for outpatient appointment scheduling practices in outpatient clinics. *Omega* 120:102907. <https://doi.org/10.1016/j.omega.2023.102907>
- Gomes MAS, Kovaleski JL, Pagani RN, da Silva VL, de Pasquini TC S (2023) Transforming healthcare with big data analytics: technologies, techniques and prospects. *J Med Eng Technol* 47(1):1–11. <https://doi.org/10.1080/03091902.2022.2096133>
- Gongora-Salazar P, Glogowska M, Fitzpatrick R, Perera R, Tsiachristas A (2022) Commissioning [Integrated] Care in England: an analysis of the current decision context. *Int J Integr Care*, 22(4)

- Groenhouf TKJ, Rittersma ZH, Bots ML, Brandjes M, Jacobs JJJ, Grobbee DE, van Solinge WW, Visseren FLJ, Haitjema S, Asselbergs FW, de Jong PA, Verhaar MC, Visseren FLJ, Asselbergs FW, van der Kaaij NP, Höfer IE, de Borst G-J, Ruigrok YM, Hollander M, Group M (2019) A computerised decision support system for cardiovascular risk management ‘live’ in the electronic health record environment: development, validation and implementation—the Utrecht Cardiovascular Cohort Initiative. *Neth Heart J*. <https://doi.org/10.1007/s12471-019-01308-w>. of the U.-C. S.
- Han PKJ, Klein WMP, Arora NK (2011) Varieties of uncertainty in health care: a conceptual taxonomy. *Med Decis Making: Int J Soc Med Decis Mak* 31(6):828–838. <https://doi.org/10.1177/0272989X11393976>
- Han PKJ, Babrow A, Hillen MA, Gulbrandsen P, Smets EM, Ofstad EH (2019) Uncertainty in health care: towards a more systematic program of research. *Patient Educ Couns* 102(10):1756–1766. <https://doi.org/10.1016/j.pec.2019.06.012>
- Howard G, Walker A, Shaw DE, Hatton J (2015) Building analytic skills to drive improvements in patient care and organisational decision making: an information analysts’ development programme. *Clin-Gov*. <https://doi.org/10.1108/CGIJ-07-2015-0021>
- Hrovat G, Stiglic G, Kokol P, Ojsteršek M (2014) Contrasting temporal trend discovery for large healthcare databases. *Comput Methods Programs Biomed* 113(1):251–257. <https://doi.org/10.1016/j.cmpb.2013.09.005>. Scopus
- Hu F, Jiang M, Celentano L, Xiao Y (2008) Robust medical ad hoc sensor networks (MASN) with wavelet-based ECG data mining. *Ad Hoc Netw*. <https://doi.org/10.1016/j.adhoc.2007.09.002>
- Islam MS, Hasan MM, Wang X, Germack HD, Noor-E-Alam M (2018) A systematic review on healthcare analytics: application and theoretical perspective of data mining. *Healthcare (Basel)* 6:54. <https://doi.org/10.3390/healthcare6020054>
- ISO (2018) *ISO 31000:2018*. ISO. <https://www.iso.org/cms/render/live/en/sites/isoorg/contents/data/standard/06/56/65694.html>
- Ji A-B, Qiao Y, Liu C (2019) Fuzzy DEA-based classifier and its applications in healthcare management. *Health Care Manag Sci* 22(3):560–568. <https://doi.org/10.1007/s10729-019-09477-1>. Scopus
- Jin H, Wu S, Vidyanti I, Di Capua P, Wu B (2015) Predicting depression among patients with diabetes using longitudinal data: a multilevel regression model. *Methods Inf Med* 54(6):553–559. <https://doi.org/10.3414/ME14-02-0009>. Scopus
- Johnson M, Albizri A, Simsek S (2020) Artificial intelligence in healthcare operations to enhance treatment outcomes: a framework to predict lung cancer prognosis. *Ann Oper Res*. <https://doi.org/10.1007/s10479-020-03872-6>
- Jung H, Chung K-Y, Lee Y-H (2015) Decision supporting method for chronic disease patients based on mining frequent pattern tree. *Multimedia Tools Appl*. <https://doi.org/10.1007/s11042-013-1730-3>
- Kamala BA, Ersdal HL, Mduma E, Moshiro R, Ginary S, Østrem OT, Linde J, Dalen I, Søyland E, Bishanga DR, Bundala FA, Makuwani AM, Richard BM, Muzzazzi PD, Kamala I, Mdoe PF (2021) SaferBirths bundle of care protocol: a stepped-wedge cluster implementation project in 30 public health-facilities in five regions, Tanzania. *BMC Health Serv Res* 21(1). <https://doi.org/10.1186/s12913-021-07145-1>. Scopus
- Kayikcioglu E, Onder AH, Bacak B, Serel TA (2024) Machine learning for predicting colon cancer recurrence. *Surg Oncol* 54:102079. <https://doi.org/10.1016/j.suronc.2024.102079>
- Kellner D, Lowin M, Hinz O (2023) Improved healthcare disaster decision-making utilizing information extraction from complementary social media data during the COVID-19 pandemic. *Decis Support Syst* 172:113983. <https://doi.org/10.1016/j.dss.2023.113983>
- Khalili G, Zargoush M, Huang K, Ghazalbash S (2024) Exploring trajectories of functional decline and recovery among older adults: a data-driven approach. *Sci Rep* 14(1):6340. <https://doi.org/10.1038/s41598-024-56606-0>
- Kopanitsa G, Metsker O, Kovalchuk S (2023) Machine learning methods for pregnancy and childbirth risk management. *J Personalized Med* 13(6). <https://doi.org/10.3390/jpm13060975>
- Kovalchuk SV, Funkner AA, Metsker OG, Yakovlev AN (2018) Simulation of patient flow in multiple healthcare units using process and data mining techniques for model identification. *J Biomed Inform*. <https://doi.org/10.1016/j.jbi.2018.05.004>
- Kraus S, Jones P, Kailer N, Weinmann A, Chaparro-Banegas N, Roig-Tierno N (2021a) Digital Transformation: an overview of the current state of the art of Research. *SAGE Open* 11(3):21582440211047576. <https://doi.org/10.1177/21582440211047576>

- Kraus S, Schiavone F, Pluzhnikova A, Invernizzi AC (2021b) Digital transformation in healthcare: analyzing the current state-of-research. *J Bus Res* 123:557–567. <https://doi.org/10.1016/j.jbusres.2020.10.030>
- Kraus M, Feuerriegel S, Saar-Tsechansky M (2024) Data-Driven allocation of Preventive Care with Application to Diabetes Mellitus type II. *Manuf Service Oper Manage* 26(1):137–153. <https://doi.org/10.1287/msom.2021.0251>
- Kuang K, Cui P, Zou H, Li B, Tao J, Wu F, Yang S (2020) Data-driven variable decomposition for treatment effect estimation. *IEEE Trans Knowl Data Eng*
- Kudyba S, Gregorio T (2010) Identifying factors that impact patient length of stay metrics for healthcare providers with advanced analytics. *Health Inf J* 16(4):235–245. <https://doi.org/10.1177/1460458210380529>
- Kuziemyk C (2016) Decision-making in healthcare as a complex adaptive system. *Health Manage Forum* 29(1):4–7. <https://doi.org/10.1177/0840470415614842>
- Laila Ue, Mahboob K, Khan AW, Khan F, Taekeun W (2022) An ensemble approach to predict early-stage diabetes risk using machine learning: an empirical study. *Sensors* 22(14):5247
- Lam HY, Ho GTS, Mo DY, Tang V (2021) Enhancing data-driven elderly appointment services in domestic care communities under COVID-19. *Industrial Management and Data Systems*. Scopus. <https://doi.org/10.1108/IMDS-07-2020-0392>
- Lambay MA, Mohideen SP (2024) Applying data science approach to predicting diseases and recommending drugs in healthcare using machine learning models – A cardio disease case study. *Multimedia Tools Appl* 83(26):68341–68361. <https://doi.org/10.1007/s11042-023-18035-5>
- Lavrač N, Bohanec M, Pur A, Cestnik B, Debeljak M, Kobler A (2007) Data mining and visualization for decision support and modeling of public health-care resources. *J Biomed Inform* 40(4):438–447
- Layeghian Javan S, Sepehri MM (2021) A predictive framework in healthcare: case study on cardiac arrest prediction. *Artif Intell Med* 117:102099. <https://doi.org/10.1016/j.artmed.2021.102099>
- Le Lay J, Augusto V, Alfonso-Lizarazo E, Masmoudi M, Gramont B, Xie X, Bongue B, Celarier T (2024) COVID-19 Bed Management Using a Two-Step Process Mining and Discrete-Event Simulation Approach. *IEEE Transactions on Automation Science and Engineering*, 21(3), 3080–3091. *IEEE Transactions on Automation Science and Engineering*. <https://doi.org/10.1109/TASE.2023.3274847>
- Lee CKH, Tse YK, Ho GTS, Chung SH (2021a) Uncovering insights from healthcare archives to improve operations: an association analysis for cervical cancer screening. *Technol Forecast Soc Chang* 162:120375. <https://doi.org/10.1016/j.techfore.2020.120375>
- Lee HJ, Lee M, Lee H, Cruz RA (2021b) Mining service quality feedback from social media: a computational analytics method. *Government Inform Q* 38(2). <https://doi.org/10.1016/j.giq.2021.101571>. Scopus
- Li X, Zhang J, Safara F (2021) Improving the accuracy of diabetes diagnosis applications through a hybrid feature selection algorithm. *Neural Process Lett*. <https://doi.org/10.1007/s11063-021-10491-0>. Scopus
- Liao H-Y, Cade W, Behdad S (2021) Markov chain optimization of repair and replacement decisions of medical equipment. *Resources, Conservation and Recycling*, 171. Scopus. <https://doi.org/10.1016/j.resconrec.2021.105609>
- Lin Y-K, Chen H, Brown RA, Li S-H, Yang H-J (2017) Healthcare predictive analytics for risk profiling in chronic care: a bayesian multitask learning approach. *MIS Quarterly: Manage Inform Syst*. <https://doi.org/10.25300/MISQ/2017/41.2.07>
- Lin L, Liang W, Li C-F, Huang X-D, Lv J-W, Peng H, Wang B-Y, Zhu B-W, Sun Y (2019a) Development and implementation of a dynamically updated big data intelligence platform from electronic health records for nasopharyngeal carcinoma research. *Br J Radiol* 92(1102):20190255. <https://doi.org/10.1259/bjr.20190255>
- Lin Y-W, Zhou Y, Faghri F, Shaw MJ, Campbell RH (2019b) Analysis and prediction of unplanned intensive care unit readmission using recurrent neural networks with long short-term memory. *PLoS ONE*, 14(7), e0218942
- Lindberg DS, Prosperi M, Bjarnadottir RI, Thomas J, Crane M, Chen Z, Shear K, Solberg LM, Snigurska UA, Wu Y (2020) Identification of important factors in an inpatient fall risk prediction model to improve the quality of care using EHR and electronic administrative data: a machine-learning approach. *Int J Med Informatics* 143:104272
- Lipshitz R, Strauss O (1997) Coping with uncertainty: a naturalistic decision-making analysis. *Organ Behav Hum Decis Process* 69(2):149–163. <https://doi.org/10.1006/obhd.1997.2679>

- Lipsitz LA (2012) Understanding health care as a complex system: the foundation for unintended consequences. *JAMA* 308(3):243–244. <https://doi.org/10.1001/jama.2012.7551>
- Liu S, Poccia S, Candan KS, Chowell G, Sapino ML (2016) EpiDMS: Data management and analytics for decision-making from epidemic spread simulation ensembles. *J Infect Dis* 214:S427–S432. <https://doi.org/10.1093/infdis/jiw305>. Scopus
- Liu H-C, Zhang L-J, Ping Y-J, Wang L (2020) Failure mode and effects analysis for proactive healthcare risk evaluation: a systematic literature review. *J Eval Clin Pract* 26:1320–1337. <https://doi.org/10.1111/jep.13317>
- Mahajan A, Madhani P, Chitikeshi S, Selvagesan P, Russell A, Mahajan P (2019) Advanced Data Analytics for Improved decision-making at a Veterans Affairs Medical Center. *J Healthc Manag*. <https://doi.org/10.1097/JHM-D-17-00164>
- Malik KM, Krishnamurthy M, Alobaidi M, Hussain M, Alam F, Malik G (2020) Automated domain-specific healthcare knowledge graph curation framework: subarachnoid hemorrhage as phenotype. *Expert Syst Appl*. <https://doi.org/10.1016/j.eswa.2019.113120>
- Mandelbaum A, Momčilović P, Trichakis N, Kadish S, Leib R, Bunnell CA (2020) Data-driven appointment-scheduling under uncertainty: the case of an infusion unit in a cancer center. *Manage Sci*. <https://doi.org/10.1287/mnsc.2018.3218>
- Manevski D, Ružič Gorenjec N, Kežar N, Blagus R (2020) Modeling COVID-19 pandemic using bayesian analysis with application to Slovene data. *Math Biosci*. <https://doi.org/10.1016/j.mbs.2020.108466>
- Markhorst B, Zver T, Malbasic N, Dijkstra R, Otto D, van der Mei R, Moeke D (2021) A Data-Driven Digital Application to enhance the Capacity Planning of the COVID-19 vaccination process. *Vaccines* 9(10):1181. <https://doi.org/10.3390/vaccines9101181>
- Mathe JL, Martin JB, Miller P, Ledeczi A, Weavind LM, Nadas A, Miller A, Maron DJ, Sztipanovits J (2009) A model-integrated, guideline-driven, clinical decision-support system. *IEEE Softw*. <https://doi.org/10.1109/MS.2009.84>
- Mayer T, Marro S, Cabrio E, Villata S (2021) Enhancing evidence-based medicine with natural language argumentative analysis of clinical trials. *Artif Intell Med* 118. Scopus <https://doi.org/10.1016/j.artmed.2021.102098>.
- Meyer G, Adomavicius G, Johnson PE, Elidrissi M, Rush WA, Sperl-Hillen JAM, O'Connor PJ (2014) A machine learning approach to improving dynamic decision making. *Inform Syst Res*. <https://doi.org/10.1287/isre.2014.0513>
- Mioishi E, Hsieh S, Caga J, Ramsey E, Chen K, Lillo P, Simon N, Vucic S, Hornberger M, Hodges JR, Kiernan MC (2014) A novel tool to detect behavioural symptoms in ALS. *Amyotroph Lateral Scler Frontotemporal Degeneration*. <https://doi.org/10.3109/21678421.2014.896927>
- Moher D, Liberati A, Tetzlaff J, Altman DG (2009) Preferred reporting items for systematic reviews and Meta-analyses: the PRISMA Statement. *J Clin Epidemiol* 62(10):1006–1012. <https://doi.org/10.1016/j.jclinepi.2009.06.005>
- Morales-Sánchez R, Montalvo S, Riaño A, Martínez R, Velasco M (2024) Early diagnosis of HIV cases by means of text mining and machine learning models on clinical notes. *Comput Biol Med* 179:108830. <https://doi.org/10.1016/j.compbiomed.2024.108830>
- Moreira MWL, Rodrigues JJPC, Kumar N, Al-Muhtadi J, Korotaev V (2018) Evolutionary radial basis function network for gestational diabetes data analytics. *J Comput Sci* 27:410–417. <https://doi.org/10.1016/j.jocs.2017.07.015>. Scopus
- Moreira MW, Rodrigues JJ, Saleem K, Korotaev VV (2020) Computational learning approaches for personalized pregnancy care. *IEEE Network* 34(2):106–111
- Mosa ASM, Hossain AM, Yoo I (2020) A dynamic prediction engine to prevent chemotherapy-induced nausea and vomiting. *Artif Intell Med*. <https://doi.org/10.1016/j.artmed.2020.101925>
- Nagamine T, Gillette B, Kahoun J, Burghaus R, Lippert J, Saxena M (2022) Data-driven identification of heart failure disease states and progression pathways using electronic health records. *Sci Rep* 12(1):1–20
- Nallathamby R, Rene Robin CR, Miriam DH (2021) Optimizing appointment scheduling for out patients and income analysis for hospitals using big data predictive analytics. *J Ambient Intell Humaniz Comput* 12(6):5783–5795. <https://doi.org/10.1007/s12652-020-02118-4>. Scopus
- Nasir M, South-Winter C, Ragothaman S, Dag A (2019) A comparative data analytic approach to construct a risk trade-off for cardiac patients' re-admissions. *Industrial Manage Data Syst*. <https://doi.org/10.1108/IMDS-12-2017-0579>

- Newland P, Newland JM, Hendricks-Ferguson VL, Smith JM, Oliver BJ (2018) Recommendations and extraction of clinical variables of Pediatric multiple sclerosis using Common Data Elements. *J Neurosci Nurs*. <https://doi.org/10.1097/JNN.0000000000000368>
- Ngiam KY, Khor IW (2019) Big data and machine learning algorithms for health-care delivery. *Lancet Oncol* 20(5):e262–e273. [https://doi.org/10.1016/S1470-2045\(19\)30149-4](https://doi.org/10.1016/S1470-2045(19)30149-4)
- O’Grady MA, Lincourt P, Greenfield B, Manseau MW, Hussain S, Genece KG, Neighbors CJ (2021) A facilitation model for implementing quality improvement practices to enhance outpatient substance use disorder treatment outcomes: a stepped-wedge randomized controlled trial study protocol. *Implement Sci*. <https://doi.org/10.1186/s13012-020-01076-x>
- Odone A, Bossi E, Gaeta M, Garancini MP, Orlandi C, Cuppone MT, Signorelli C, Nicastro O, Zotti CM (2019) *Risk Management in healthcare: Results from a national-level survey and scientometric analysis in Italy*
- OECD (2020) *Digital Economy Outlook 2020* | OECD. <https://www.oecd.org/digital/oecd-digital-economy-outlook-2020-bb167041-en.htm>
- OECD (2019) *Health in the 21st Century: Putting Data to Work for Stronger Health Systems*. Organisation for Economic Co-operation and Development. https://www.oecd-ilibrary.org/social-issues-migration-health/health-in-the-21st-century_e3b23f8e-en
- OECD (2015) *Data-Driven Innovation: Big Data for Growth and Well-Being*. Organisation for Economic Co-operation and Development. https://www.oecd-ilibrary.org/science-and-technology/data-driven-innovation_9789264229358-en
- Pakravan MH, Johnson AC (2021) Electrification planning for healthcare facilities in low-income countries, application of a portfolio-level, multi criteria decision-making approach. *ISPRS Int J Geo-Information* 10(11). <https://doi.org/10.3390/ijgi10110750>. Scopus
- Patient Safety*. (n.d.). Retrieved April 10, (2022) from <https://www.who.int/news-room/fact-sheets/detail/patient-safety>
- Perez I, Brown M, Pinchin J, Martindale S, Sharples S, Shaw D, Blakey J (2016) Out of hours workload management: bayesian inference for decision support in secondary care. *Artif Intell Med*. <https://doi.org/10.1016/j.artmed.2016.09.005>
- Phillips-Wren G, Sharkey P, Dy SM (2008) Mining lung cancer patient data to assess healthcare resource utilization. *Expert Syst Appl* 35(4):1611–1619. <https://doi.org/10.1016/j.eswa.2007.08.076>
- Picozzi P, Nocco U, Pezzillo A, De Cosmo A, Cimolin V (2024) The Use of Business Intelligence Software to monitor key performance indicators (KPIs) for the evaluation of a computerized maintenance management system (CMMS). *Electronics* 13(12). <https://doi.org/10.3390/electronics13122286>
- Pinto A, Pereira S, Meier R, Wiest R, Alves V, Reyes M, Silva CA (2021) Combining unsupervised and supervised learning for predicting the final stroke lesion. *Med Image Anal* 69:101888
- Poly TN, Islam MM, Muhtar MS, Yang H-C, Nguyen PAA, Li Y-CJ (2020) Machine learning approach to reduce alert fatigue using a disease medication-related clinical decision support system: model development and validation. *JMIR Med Inf* 8(11):e19489
- Provost F, Fawcett T (2013) Big Data 1(1):51–59. <https://doi.org/10.1089/big.2013.1508>. Data Science and its Relationship to Big Data and Data-Driven Decision Making
- Qiu H, Luo L, Su Z, Zhou L, Wang L, Chen Y (2020) Machine learning approaches to predict peak demand days of cardiovascular admissions considering environmental exposure. *BMC Med Inf Decis Mak* 20(1):1–11
- Rabbani N, Ma SP, Li RC, Winget M, Weber S, Boosi S, Pham TD, Svec D, Shieh L, Chen JH (2023) Targeting repetitive laboratory testing with electronic health records-embedded predictive decision support: a pre-implementation study. *Clin Biochem* 113:70–77. <https://doi.org/10.1016/j.clinbiochem.2023.01.002>
- Raghu A, Praveen D, Peiris D, Tarassenko L, Clifford G (2015) Engineering a mobile health tool for resource-poor settings to assess and manage cardiovascular disease risk: SMARThealth study. *BMC Med Inf Decis Mak*. <https://doi.org/10.1186/s12911-015-0148-4>
- Rahman NHA, Hasikin K, Razak NAA, Al-Ani AK, Anni DJS, Mohandas P (2023) Medical Device Failure Predictions Through AI-Driven Analysis of Multimodal Maintenance Records. *IEEE Access*, 11, 93160–93179. IEEE Access. <https://doi.org/10.1109/ACCESS.2023.3309671>
- Raj P, Tejwani S, Sudha D, Muthu Narayanan B, Thangapandi C, Das S, Somasekar J, Mangalapudi S, Kumar D, Pindipappanahalli N, Shetty R, Ghosh A, Kumaramanickavel G, Chaudhuri A, Soumitra N (2020) Ophthathome™: an integrated knowledgebase of ophthalmic diseases for translating vision research into the clinic. *BMC Ophthalmol* 20(1). <https://doi.org/10.1186/s12886-020-0170-5-5>. Scopus

- Ramon-Gonen R, Dori A, Shelly S (2023) Towards a practical use of text mining approaches in electrodiagnostic data. *Sci Rep* 13(1):19483. <https://doi.org/10.1038/s41598-023-45758-0>
- Rangelov B, Young A, Lilaonitkul W, Aslani S, Taylor P, Guðmundsson E, Yang Q, Hu Y, Hurst JR, Hawkes DJ, Jacob J (2023) Delineating COVID-19 subgroups using routine clinical data identifies distinct in-hospital outcomes. *Sci Rep* 13(1):9986. <https://doi.org/10.1038/s41598-023-32469-9>
- Rath S, Rajaram K, Mahajan A (2017) Integrated anesthesiologist and room scheduling for surgeries: methodology and application. *Oper Res*. <https://doi.org/10.1287/opre.2017.1634>
- Rathore MM, Ahmad A, Paul A, Wan J, Zhang D (2016) Real-time Medical Emergency Response System: exploiting IoT and Big Data for Public Health. *J Med Syst*. <https://doi.org/10.1007/s10916-016-0647-6>
- Riad A, Huang Y, Abdulqader H, Morgado M, Domnori S, Koščik M, Mendes JJ, Klugar M, Kateeb E, IADS-SCORE (2021) Universal predictors of dental students' attitudes towards covid-19 vaccination: machine learning-based approach. *Vaccines* 9(10). <https://doi.org/10.3390/vaccines9101158>. Scopus
- Rios-Zertuche D, Gonzalez-Marmol A, Millán-Velasco F, Schwarzbauer K, Tristao I (2020) Implementing electronic decision-support tools to strengthen healthcare network data-driven decision-making. *Archives Public Health*. <https://doi.org/10.1186/s13690-020-00413-2>
- Rosemarin H, Rosenfeld A, Lapp S, Kraus S (2021) Lba: online learning-based assignment of patients to medical professionals. *Sensors* 21(9). <https://doi.org/10.3390/s21093021>. Scopus
- Ryan J, Hendler J, Bennett KP (2015) Understanding emergency department 72-hour revisits among Medicaid patients using electronic healthcare records. *Big Data* 3(4):238–248. <https://doi.org/10.1089/big.2015.0038>. Scopus
- Sardi A, Rizzi A, Sorano E, Guerrieri A (2020). Cyber risk in health facilities: a systematic literature review. *Sustainability* 12:7002. <https://doi.org/10.3390/su12177002>
- Salazar-Reyna R, Gonzalez-Aleu F, Granda-Gutierrez EMA, Diaz-Ramirez J, Garza-Reyes JA, Kumar A (2020). A systematic literature review of data science, data analytics and machine learning applied to healthcare engineering systems. *Management Decision* 60:300–319. <https://doi.org/10.1108/MD-01-2020-0035>
- Sarkar S, Pramanik A, Maiti J, Reniers G (2021) COVID-19 outbreak: a data-driven optimization model for allocation of patients. *Computers Industrial Eng* 161. Scopus <https://doi.org/10.1016/j.cie.2021.107675>.
- Sarp S, Kuzlu M, Zhao Y, Gueler O (2023) Digital Twin in Healthcare: A Study for Chronic Wound Management. *IEEE Journal of Biomedical and Health Informatics*, 27(11), 5634–5643. *IEEE Journal of Biomedical and Health Informatics*. <https://doi.org/10.1109/JBHI.2023.3299028>
- Sebaa A, Chikh F, Nouicer A, Tari A (2018) Architecture and System Design, a Case Study: Improving Healthcare Resources Distribution. *J Med Syst* 42(4):59. <https://doi.org/10.1007/s10916-018-0894-9>. *Medical Big Data Warehouse*:
- Sechi GM, Migliori M, Dassi G, Pagliosa A, Bonora R, Oradini-Alacreu A, Odone A, Signorelli C, Zoli A, Team A, C.-19 R (2020) Business Intelligence applied to Emergency Medical Services in the Lombardy region during SARS-CoV-2 epidemic. *Acta Biomed Atenei Parmensis* 91(2) Article 2. <https://doi.org/10.23750/abm.v91i2.9557>
- Seguí FL, Guillamet GH, Arolas HP, Marin-Gomez FX, Comellas AR, Morros AMR, Mas CA, Vidal-Alaball J (2021) Characterization and identification of variations in types of primary care visits before and during the covid-19 pandemic in catalonia: Big data analysis study. *Journal of Medical Internet Research*, 23(9). Scopus. <https://doi.org/10.2196/29622>
- Sharma A, Hostetter J, Morrison J, Wang K, Siegel E (2016) Lung, Colorectal, and Ovarian Cancer Screening Trial Dataset and Guide Screening Management for the Individual Patient. *J Digit Imaging*. <https://doi.org/10.1007/s10278-015-9826-0>. *Focused Decision Support: A Data Mining Tool to Query the Prostate*,
- Shirazibeheshti A, Etefaghian A, Khanizadeh F, Wilson G, Radwan T, Luca C (2023) Automated detection of patients at high risk of Polypharmacy including anticholinergic and sedative medications. *Int J Environ Res Public Health* 20(12). <https://doi.org/10.3390/ijerph20126178>
- Shoab M, Junaid A, Husnain G, Qadir M, Ghadi YY, Askar SS, Abouhawwash M (2024) Advanced detection of coronary artery disease via deep learning analysis of plasma cytokine data. *Front Cardiovasc Med* 11. <https://doi.org/10.3389/fcvm.2024.1365481>
- Simsek S, Tiaht T, Dag A (2020) Stratifying no-show patients into multiple risk groups via a holistic data analytics-based framework. *Decis Support Syst*. <https://doi.org/10.1016/j.dss.2020.113269>

- Simsek S, Dag A, Tiaht T, Oztekin A (2021) A Bayesian Belief Network-based probabilistic mechanism to determine patient no-show risk categories. *Omega* (United Kingdom). <https://doi.org/10.1016/j.omega.2020.102296>
- Singh P, Kaur A, Batth RS, Kaur S, Gianini G (2021) Multi-disease big data analysis using beetle swarm optimization and an adaptive neuro-fuzzy inference system. *Neural Comput Appl* 33(16):10403–10414. <https://doi.org/10.1007/s00521-021-05798-x>. Scopus
- Soffer S, Klang E, Barash Y, Grossman E, Zimlichman E (2021) Predicting in-hospital mortality at admission to the medical ward: a big-data machine learning model. *Am J Med* 134(2):227–234
- Song W, Latham NK, Liu L, Rice HE, Sainlaire M, Min L, Zhang L, Thai T, Kang M-J, Li S, Tejada C, Lipsitz S, Samal L, Carroll DL, Adkison L, Herlihy L, Ryan V, Bates DW, Dykes PC (2024) Improved accuracy and efficiency of primary care fall risk screening of older adults using a machine learning approach. *J Am Geriatr Soc* 72(4):1145–1154. <https://doi.org/10.1111/jgs.18776>
- Spiegelhalter DJ (2008) Understanding uncertainty. *Ann Fam Med* 6(3):196–197. <https://doi.org/10.1370/afm.848>
- Stein RM (2016) Real decision support for health insurance policy selection. *Big Data*. <https://doi.org/10.1089/big.2016.0012>
- Sun H, Arndt D, De Roo J, Mannens E (2021) Predicting future state for adaptive clinical pathway management. *J Biomed Inform*. <https://doi.org/10.1016/j.jbi.2021.103750>
- Tantalaki N, Souravlas S, Roumeliotis M (2019) Data-Driven decision making in Precision Agriculture: the rise of Big Data in Agricultural systems. *J Agricultural Food Inform* 20(4):344–380. <https://doi.org/10.1080/10496505.2019.1638264>
- Tao J, Yuan Z, Sun L, Yu K, Zhang Z (2021) Fetal birthweight prediction with measured data by a temporal machine learning method. *BMC Med Inf Decis Mak* 21(1). <https://doi.org/10.1186/s12911-021-01388-y>. Scopus
- Tartaglia R, Albolino S, Bellandi T, Bianchini E, Biggeri A, Fabbro G, Bevilacqua L, Dell'Erba A, Privitera G, Sommella L (2012) Adverse events and preventable consequences: retrospective study in five large Italian hospitals. *Epidemiol Prev* 36(3–4):151–161
- Thompson S, Varvel S, Sasinowski M, Burke JP (2016) From value assessment to value cocreation: informing clinical decision-making with medical claims data. *Big Data*. <https://doi.org/10.1089/big.2015.0030>
- Tolk A, Glazner C, Ungerleider J (2021) Computational decision support for the covid-19 healthcare coalition. *Comput Sci Eng* 23(1):17–24. <https://doi.org/10.1109/MCSE.2020.3036586>. Scopus
- Topuz K, Zengul FD, Dag A, Almelhmi A, Yildirim MB (2018) Predicting graft survival among kidney transplant recipients: a bayesian decision support model. *Decis Support Syst* 106:97–109. <https://doi.org/10.1016/j.dss.2017.12.004>. Scopus
- Valero-Ramon Z, Fernandez-Llata C, Valdivieso B, Traver V (2020) Dynamic models supporting personalised chronic disease management through healthcare sensors with interactive process mining. *Sens* (Switzerland). <https://doi.org/10.3390/s20185330>
- Vanasse A, Courteau J, Courteau M, Benigeri M, Chiu YM, Dufour I, Couillard S, Larivée P, Hudon C (2020) Healthcare utilization after a first hospitalization for COPD: a new approach of State Sequence Analysis based on the 6 W multidimensional model of care trajectories. *BMC Health Serv Res*. <https://doi.org/10.1186/s12913-020-5030-0>
- Wang J, Jia Y, Wang D, Xiao W, Wang Z (2022) Weighted IForest and siamese GRU on small sample anomaly detection in healthcare. *Comput Methods Programs Biomed* 218:106706
- Wang B, Li W, Bradlow A, Bazuaye E, Chan ATY (2023) Improving triaging from primary care into secondary care using heterogeneous data-driven hybrid machine learning. *Decis Support Syst* 166:113899. <https://doi.org/10.1016/j.dss.2022.113899>
- Wautelet Y, Kolp M, Heng S, Poelmans S (2018) Developing a multi-agent platform supporting patient hospital stays following a socio-technical approach: management and governance benefits. *Telematics Inform*. <https://doi.org/10.1016/j.tele.2017.12.013>
- Weerasinghe K, Scahill SL, Pauleen DJ, Taskin N (2022) Big data analytics for clinical decision-making: understanding health sector perceptions of policy and practice. *Technol Forecast Soc Chang* 174. Scopus <https://doi.org/10.1016/j.techfore.2021.121222>.
- Wong B, Ho GTS, Tsui E (2017) Development of an intelligent e-healthcare system for the domestic care industry. *Industrial Manage Data Syst*. <https://doi.org/10.1108/IMDS-08-2016-0342>
- Wong-Lin KF, McClean PL, McCombe N, Kaur D, Sanchez-Bornot JM, Gillespie P, Todd S, Finn DP, Joshi A, Kane J, McGuinness B (2020) Shaping a data-driven era in dementia care pathway through computational neurology approaches. *BMC Med*. <https://doi.org/10.1186/s12916-020-01841-1>
- World Health Organization (2023). Patient Safety [WWW Document], <https://www.who.int/news-room/fact-sheets/detail/patient-safety> (accessed 4.10.23)

- Wu J, Liu C, Xie L, Li X, Xiao K, Xie G, Xie F (2022) Early prediction of moderate-to-severe condition of inhalation-induced acute respiratory distress syndrome via interpretable machine learning. *BMC Pulm Med* 22(1):1–9
- Xie J, Liu X, Zeng D, D., Fang X (2022) UNDERSTANDING MEDICATION NONADHERENCE FROM SOCIAL MEDIA: A SENTIMENT-ENRICHED DEEP LEARNING APPROACH. *MIS Q*, 46(1)
- Xu M, Wong TC, Chin KS (2013) Modeling daily patient arrivals at Emergency Department and quantifying the relative importance of contributing variables using artificial neural network. *Decis Support Syst* 54(3):1488–1498. <https://doi.org/10.1016/j.dss.2012.12.019>
- Yeh J-Y, Wu T-H, Tsao C-W (2011) Using data mining techniques to predict hospitalization of hemodialysis patients. *Decis Support Syst*. <https://doi.org/10.1016/j.dss.2010.11.001>
- Yu G, Tabatabaei M, Mezei J, Zhong Q, Chen S, Li Z, Li J, Shu L, Shu Q (2022) Improving chronic disease management for children with knowledge graphs and artificial intelligence. *Expert Syst Appl* 201:117026
- Zamzam AH, Al-Ani AKI, Wahab AKA, Lai KW, Satapathy SC, Khalil A, Azizan MM, Hasikin K (2021) Prioritisation Assessment and Robust Predictive System for Medical Equipment: a comprehensive strategic maintenance management. *Front Public Health*, 9
- Zhang W, Li X (2024) A data-driven combined prediction method for the demand for intensive care unit healthcare resources in public health emergencies. *BMC Health Serv Res* 24(1):477. <https://doi.org/10.1186/s12913-024-10955-8>
- Zhang P, Zhu X, Shi Y, Guo L, Wu X (2011) Robust ensemble learning for mining noisy data streams. *Decis Support Syst* 50(2):469–479. <https://doi.org/10.1016/j.dss.2010.11.004>
- Zhang J, Tutun S, Anvazydz SF, Amini M, Sundaramoorthi D, Sundaramoorthi H (2024) Management of resource sharing in emergency response using data-driven analytics. *Ann Oper Res* 339(1):663–692. <https://doi.org/10.1007/s10479-023-05702-x>
- Zhou X, Lee EWJ, Wang X, Lin L, Xuan Z, Wu D, Lin H, Shen P (2022) Infectious diseases prevention and control using an integrated health big data system in China. *BMC Infect Dis* 22:344. <https://doi.org/10.1186/s12879-022-07316-3>
- Zhu K, Lou Z, Zhou J, Ballester N, Kong N, Parikh P (2015) Predicting 30-day hospital readmission with publicly available administrative database: a conditional logistic regression modeling approach. *Methods Inf Med*. <https://doi.org/10.3414/ME14-02-0017>
- Zolbanin HM, Delen D, Hassan Zadeh A (2015) Predicting overall survivability in comorbidity of cancers: a data mining approach. *Decis Support Syst*. <https://doi.org/10.1016/j.dss.2015.04.003>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Authors and Affiliations

Luigi Jesus Basile¹  · Nunzia Carbonara² · Umberto Panniello² · Roberta Pellegrino²

✉ Luigi Jesus Basile
luigijesus.basile@imt-bs.eu

Nunzia Carbonara
nunzia.carbonara@poliba.it

Umberto Panniello
umberto.panniello@poliba.it

Roberta Pellegrino
roberta.pellegrino@poliba.it

¹ Université Paris-Saclay, Univ Evry, IMT-BS, LITEM, Evry-Courcouronnes 91025, France

² Department of Mechanics, Mathematics and Management, Polytechnic University of Bari, Bari, Italy