SORT 47 (1) January-June 2023, 81-128 DOI: 10.57645/20.8080.02.2

Data science, analytics and artificial intelligence in e-health: trends, applications and challenges

Juliana Castaneda[1], Laura Calvet[2], Sergio Benito[3], Abtin Tondar[4] and Angel A. Juan[5]

Abstract

More than ever, healthcare systems can use data, predictive models, and intelligent algorithms to optimize their operations and the service they provide. This paper reviews the existing literature regarding the use of data science/analytics methods and artificial intelligence algorithms in healthcare. The paper also discusses how healthcare organizations can benefit from these tools to efficiently deal with a myriad of new possibilities and strategies. Examples of real applications are discussed to illustrate the potential of these methods. Finally, the paper highlights the main challenges regarding the use of these methods in healthcare, as well as some open research lines.

MSC: *68T01, 62-07.*

Keywords: *e-health, data science, analytics, artificial intelligence, machine learning.*

1. Introduction

In recent years, there has been a growing trend to digitize much of the data that used to be stored in hard copies. The healthcare industry, which has always been character-

[1] Dept. of Computer Science, Multimedia and Telecommunication, Universitat Oberta de Catalunya, 08018 Barcelona, Spain.

[2] Dept. of Computer Science, Multimedia and Telecommunication, Universitat Oberta de Catalunya, 08018 Barcelona, Spain.

[3] Dept. of Computer Science, Multimedia and Telecommunication, Universitat Oberta de Catalunya, 08018 Barcelona, Spain.

[4] Dept. of Computer Science, Multimedia and Telecommunication, Universitat Oberta de Catalunya, 08018 Barcelona, Spain.

[5] Dept. of Statistics and Operations Research, Universitat Politècnica de València, 03801 Alcoy, Spain.

Received: November 2021
Accepted: July 2022

ized by the generation of large amounts of data, has already begun this digital transformation (Raghupathi and Raghupathi, 2014). The term electronic health, or 'e-health', appeared in the 1990s by the influence of the Internet. The prefix 'e-' became popular to accompany different terms, such as e-mail or e-commerce, referring to various developments in information and communication technology (ICT). The term e-health has been defined by Eysenbach (2001) as: "an emerging field in the intersection of medical informatics, public health and business, referring to health services and information delivered or enhanced through the Internet and related technologies. In a broader sense, the term characterizes not only a technical development, but also a state-of-mind, a way of thinking, an attitude, and a commitment for networked, global thinking, to improve health care locally, regionally, and worldwide by using information and communication technology". This information could include several health-related concepts, as well as various stakeholders, roles, locations, and benefits (Oh et al., 2005). In this sense, ICTs are supporting tools for health-related activities. Within the main domains of e-health, we can find: telemedicine, clinical information systems, different types of medical networks, disease registries for different purposes (education, public health, patient/disease behavior, and healthcare management), mobile health, personalized health, and big data (Cowie et al., 2016). All in all, this allows us to consider e-health as a big data source, and a potential field for applying data science/analytics techniques, including predictive models, optimization algorithms, modeling and simulation, or any other technique that involves data processing for a defined purpose. In addition, new models, such as learning healthcare systems, have been recently developed to facilitate using medical data for improving healthcare (Enticott, Johnson and Teede, 2021).

Data science/analytics emerged as a hybridization of several disciplines, such as statistics, operations research/management science, data mining, computer science, data bases, machine learning, mathematics, and distributed systems. The combination of all the existing methodologies in this field makes the large amounts of data available valuable for individuals, organizations, and society (Van Der Aalst, 2016). Concepts such as artificial intelligence (AI), machine learning (ML), statistical learning, or data science are clearly interconnected, and they share many methods and techniques. For instance, it is possible to find predictive, regression, classification, and clustering models in all the previous concepts. Still, they are not exactly the same concept. Hence, AI is a wide area, being its main goal the development of machines capable of emulating human intelligence. With that purpose, it uses computer science algorithms (e.g., optimization and searching algorithms), statistical methods, ML models, computer vision techniques, etc. ML is usually seen as a subset of AI, and it focuses on a series of supervised methods (classification, regression, predictive models, etc.), unsupervised methods (clustering, dimensionality reduction, etc.), and reinforcement learning methods. Many of these methods are also employed in statistical learning. However, statistical learning is more focused on the statistical fundamentals of these methods, while ML is more oriented towards their computational and programming aspects. Data science, on the other hand also refers to many of the ML and AI methods as well as to the use of databases with large amounts of data, data gathering and pre-processing, and other analytical methods and algorithms (including, for instance, simulation models, time series analysis, etc.).

Managing big data in healthcare is a complex task because of its volume and the diversity of data types, and the speed at which they must be processed. There is an opportunity for data analysts to discover associations and understand patterns and trends within the data. Big data analytics can improve care, save lives, and reduce costs (Raghupathi and Raghupathi, 2014). Potential applications of data science/analytics methods and AI algorithms to e-health are almost unlimited: from the enhancement of interoperability in e-health systems (Gupta and Gupta, 2019; Razzaque and Hamdan, 2020) to the use of Internet of things (IoT) and algorithms for generating smart healthcare networks (Syed et al., 2019). This includes the use of healthcare data to provide smart medical services to citizens (Haldorai, Ramu and Murugan, 2019) or high-risk pregnancy home healthcare (Moreira et al., 2018). Typical applications of analytics/operations research/management science methods in healthcare –including the pharmaceutical industry as well– can be found in Beheshtifar and Alimoahmmadi (2015), Rais and Viana (2011), Saranga and Phani (2009), or Ahsan and Bartlema (2004).

Based on the Google Scholar and Elsevier Scopus databases, Figure 1 shows the time evolution of the number of scientific articles that contain all the following terms: "artificial intelligence", "e-health", and "data science". Notice the fast growth in the number of papers that combine all the aforementioned terms, which shows a clear tendency in the literature to consider the combined use of data analysis methods and AI algorithms in the e-health sector. Accordingly, one of the main goals of this work is to analyze the data science/analytics/AI methodologies implemented so far in e-health applications. This is achieved by identifying the main techniques and examples available in the scientific literature. A discussion on the benefits these techniques offer to the e-health sector, including a series of best practices, is another contribution of this paper. Finally, we also propose some open challenges yet to be fully explored. There are some recent reviews on related topics, such as those by Matheny, Whicher and Israni (2020) or Rong et al. (2020), among others. Our work contributes to this field by providing a holistic overview regarding the use of data science/analytics methods and AI algorithms in e-health and a discussion of the main open challenges. In our view, this can be equally useful for managers and researchers in the area.

The remaining of this paper is organized as follows: Section 2 offers a description of the primary needs of healthcare organizations and, in general, healthcare systems, as well as a text mining analysis that aims at identifying the most relevant keywords and hot topics. Section 3 provides an overview of the leading data science/analytics methods that can be applied in the e-health context. Section 4 performs a similar analysis in the case of AI algorithms. Several examples of real-life applications are analyzed in Section 5, while the main identified challenges and open research lines are discussed in Section 6. Finally, the main findings of this study are summarized in Section 7.

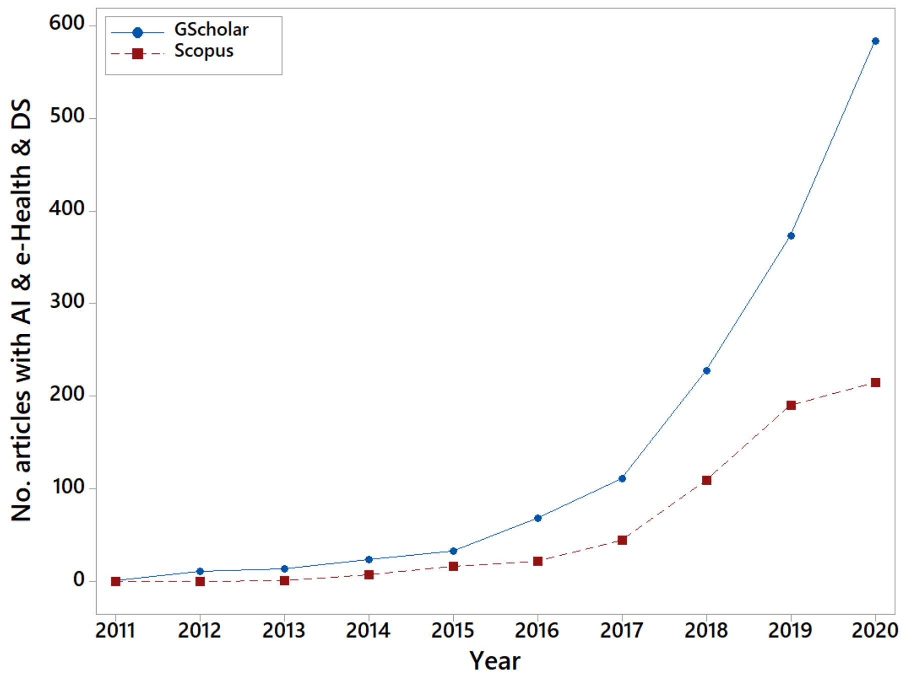

Figure 1. *Number of articles in Google Scholar and Scopus including all the terms: "artificial intelligence", "e-health", and "data science".*

2. Data science/analytics, AI, and current trends in healthcare

Healthcare is one of the biggest and fastest-growing industries in the world. Healthcare management has been changing from disease and volume-focused to patient and value-centered delivery systems in recent years (Huang et al., 2015). The efficient management, analysis, and use of big healthcare data are crucial for providing patient and value-centered care. Most common traditional data management protocols-which are currently in use in healthcare centers-cannot analyze big data efficiently since the complexity and volume of data in healthcare have significantly increased over the past three decades. Therefore, there is an ever-increasing demand for new innovative methods and tools of big data management to support the healthcare industry (Feldman, Davis and Chawla, 2015).

In the last year alone, several alliances have been taking place among some of the biggest pharmaceutical companies and the technology giants in AI use. For instance, Boehringer Ingelheim has announced an agreement with Google Quantum AI to support the research and application of use cases for quantum computing in pharmaceutical research and molecular dynamics simulations (Reinig, 2021). Novartis and Microsoft have reaffirmed their commitment to using AI for drug research and development (Zuest, 2019). AstraZeneca has announced partnerships with Alibaba to develop smart health services, screening tests, and AI-assisted diagnostic tools in China (Martuscelli, 2018). Capgemini has announced the signing of a contract with Bayer AG to accelerate its digitization (Connatty, 2019). IBM is close to predicting and diagnosing Alzheimer's, for which it has partnered with Pfizer to develop a model for early detection using AI (Terry, 2020). Almirall and Iktos (a company specialized in AI) have signed an agreement to accelerate the discovery of new drugs (Al Idrus, 2019). In the recent COVID-19 pandemic, a Canadian start-up AI company (BlueDot), which tracks and predicts the spread of infectious diseases, alarmed its customers about the spread of atypical pneumonia that was taking place near a shopping area in Wuhan, China. BlueDot was the first organization in predicting the spread of that disease, even nine days before the World Health Organization released its report about the outbreak of a new coronavirus in China (Bowles, 2020).

All this is reflected in the growing number of scientific articles that have been developed in recent years. The large volume of information available in this area creates the need to use intelligent techniques capable of processing large amounts of text and extracting valuable information from it, such as the topics most addressed. For this purpose, 633 scientific papers indexed in the Scopus database for the terms "artificial intelligence", "e-health", and "data science" are taken as a basis for analysis. We have completed a text mining process on their titles and abstracts. This analysis allows us to: *(i)* automatically identify the most cited keywords in all these papers; and *(ii)* automatically generate a list of hot topics in this research field. The latter goal has been achieved by employing a latent semantic analysis with the non-negative matrix factorization (NMF) algorithm. NMF is a feature extraction algorithm that combines attributes to produce meaningful topics. It decomposes multivariate data by creating a user-defined number of features, each of which is a linear combination of the original attribute set (Huang, Zhou and Zhang, 2012b). This algorithm enables the modeling of the topics, i.e., to extract the significant topics that recur in the corpus or media group of similar documents. Documents are decomposed into topics and topics into words. This technique requires that, in the case of texts, all documents to be analyzed have a similar length. We analyze abstracts and titles, which generally have a similar length in terms of number of words, so this requirement is met.

At a technical level, the NMF decomposes the matrix of visible variable (V), which is the input, into two smaller matrices, the document-topic matrix (W) and the topic-term matrix (H). Matrix V contains a count of the occurrence of each word (document by term frequency), matrix W, presents for each row one document by the non-normalized probabilities of topics. The W matrix presents for each row one document per non-standardized probabilities of the topics, and allows us to interpret that two terms that appear together frequently form a topic and each term gives more contextual meaning. Matrix H allows to establish the number of topics which are interpreted as every two terms appearing together frequently form a topic. Each term gives more contextual meaning to the term with which it is grouped, and if a term appears frequently in two themes, they are likely to be related. This allows us to establish the number of topics

(t) to determine the size of these matrices. In addition, it has the advantage that each topic is interpretable, which is not the case with other matrix decomposition methods such as principal component analysis and vector quantization that only use non-negative numbers (Snasel et al., 2007).

Both the data preprocessing and the implementation of the NMF algorithm have been performed with a Python script. Data preprocessing consists of the cleaning necessary before applying the algorithm. For this purpose, in the Python script an exclusion list is constructed with the words to be removed, such as articles, connectors, prepositions and determiners found in the data set. The "counter" function of the "collections" library is used, which is a container that counts the number of times that there is an equivalent value in the dataset to be analyzed. This allows us to know the number of words not significant for the analysis that are still in the input data of the algorithm. The cleaning process is repeated until the non-relevant words and their equivalents are not found in the dataset. The topic extraction model is with the NMF decomposition algorithm in the scikit-learn library (Pedregosa et al., 2011) as described by Liu (2017). The parameters implemented were $n\_components = 5$ and $n\_top\_words = 10$. Therefore, the results provide the 10 most popular words and the top 5 topics in the analyzed articles. The Python code is available at https://github.com/NMF_ehealth.git As shown in Figure 2, the most popular words in the set of analyzed articles are: "data", "health/healthcare", "research", "IoT", "security", "smart", "information", "system", and "learning".

Likewise, by configuring the NMF algorithm (Python scikit-learn version) to determine the top five topics, we obtain the results presented in Table 1. These topics are determined from organizing the sets of five keywords generated by the algorithm. Articles, connectors, prepositions and determiners are added to make sense of the five keywords in each set. These topics can be considered as some of the most popular emerging research lines in the literature, and they include keywords such as "machine learning", "IoT", "security", "blockchain", "privacy", "big data", "medical analytics", "cloud and fog computing", etc.

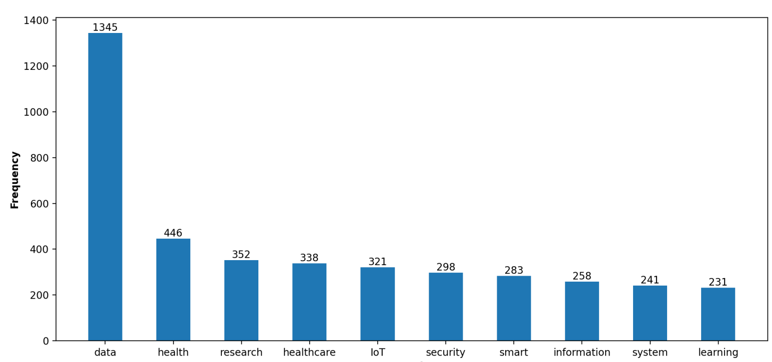

Figure 2. *Most common words in the articles indexed in Scopus including all the terms: "artificial intelligence", "e-health", and "data science".*

Table 1. *Topic modeling results obtained with the NMF algorithm for the articles indexed in Scopus including all the terms: "artificial intelligence", "e-health", and "data science".*

No.	Five-word set	Topic	References
1	learning, machine, deep, model, recognition	machine & deep learning for recognition	Shatte, Hutchinson and Teague (2019) Yu, Beam and Kohane (2018) Kavakiotis et al. (2017)
2	IoT, internet, things, devices, security	IoT devices security	Al-Garadi et al. (2020) Din et al. (2019) Makhdoom et al. (2018)
3	blockchain, technology, applications, consensus, research	blockchain technology in security and privacy applications	Chukwu and Garg (2020) Roy et al. (2018)
4	data, big, analytics, processing, medical	big data in healthcare medical analytics	Wang and Alexander (2020) Syed et al. (2019)
5	access, control, encryption, data, attribute	cloud and fog computing for privacy and security	Sun (2020); Dang et al. (2019) Mutlag et al. (2019) Puliafito et al. (2019)

A second text mining process on titles and abstracts is performed only for the terms "e-health" and "artificial intelligence" in the Scopus database, considering only articles, literature reviews, conference papers and book chapters. 403 documents are analysed using the same parameters as described above. Discounting the words "health" with 487 occurrences, "data" with 397, "medical" with 252, and "healthcare" with 247, we obtain the 10 most common words among the analysed documents shown in Figure 3. The most common words in the titles and abstracts of these documents are: "information", "patients", "systems", "e-health", "decision", "artificial", "clinical", "learning", "monitoring", and "smart". This reveals a trend towards AI research in healthcare related to patient data and clinical information, as well as monitoring, and decision-making. We have also extracted the top five topics using the NMF model. Table 2 presents the results showing that the most emerging lines of AI research relate to medical patient information, decision-making, IoT and cloud systems integration, the COVID-19 pandemic, and the use of machine learning in the area of diseases. As in the previous text mining process, these five topics are the result of the interpretation of the resulting set of five keywords generated by the algorithm.

3. Data science/analytics methods in e-health

The emergence of e-health has created a significant demand for data analysis of people's health and administrative healthcare processes. For this reason, this section presents an

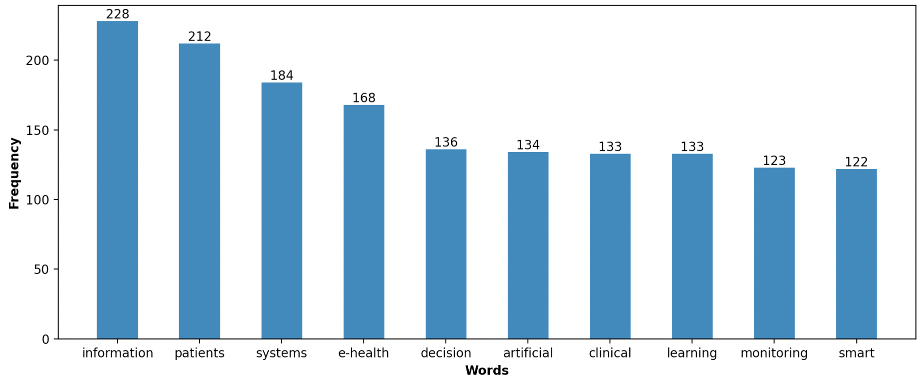

Figure 3. *Most common words in the titles and abstracts of indexed documents in Scopus including all the terms: "artificial intelligence" and "e-health".*

analysis of the methodologies and techniques of data science/analytics implemented in e-health systems, also considering AI methodologies. Data science allows us to apply both quantitative and qualitative methods to solve significant problems and predict results (Waller and Fawcett, 2013). Big data and data analysis go hand in hand because data is considered the raw material of data science (Larson and Chang, 2016). When applying these methods in e-health, one of the main goals is to extract knowledge from data in order to improve patient care (McIntosh et al., 2016). The data collected from software applications in e-health, apart from being available in large amounts, are also characterized by not having well-defined structures and being heterogeneous. Therefore, they require methods that order, filter, process, and extract patterns from large amounts of data in order to make them valuable. According to Van Der Aalst (2016), data science refers to data extraction, data preparation, data exploration, data transformation, storage and retrieval, computing infrastructures, all types of mining and learning, presentation of explanations and predictions, and the exploitation of results taking into account ethical, social, legal, and business aspects. Hernán, Hsu and Healy (2019) organize data science into three tasks: (i) description, i.e., referring to quantitative analysis ranging from elementary calculations to unsupervised learning algorithms and intelligent data visualization techniques; (ii) event prediction, i.e., using elementary predictive calculations, such as correlation of variables, and methods for recognizing patterns in data and supervised learning algorithms; and (iii) counterfactual prediction, i.e., using data to predict events in different scenarios involving causal inference. Table 3 summarizes the data science methods applied in different e-health contexts. It also mentions the general techniques with some possible applications.

According to Galetsi and Katsaliaki (2020), Descriptive Analytics techniques are used to identify problems and trends in the data. Within the examples of applications in e-health, they are mainly used to identify diseases or provide analysis on medical information obtained from different sources, such as mobile applications or medical test equipment. Wang and Hajli (2017) highlight the five main business capabilities of big

Table 2. *Topic modeling results obtained with the NMF algorithm analyzing the titles and abstracts of indexed documents in Scopus including all the terms: "artificial intelligence" and "e-health".*

No.	Five-word set	Topic	References
1	health, information, data, care, patients	patients health care information data	Mamdouh et al. (2020) Susanto (2017) Sethia et al. (2016)
2	decision, support, clinical, based, making	clinical-based decision-making support	Hu et al. (2016) Impedovo, Pirlo and Vessio (2018) Aldape-Pérez et al. (2018)
3	IoT, applications, things, internet, cloud	IoT / internet cloud applications	Zhu et al. (2015) Miori and Russo (2012) Lakshmanaprabu et al. (2019) Ruiz-Zafra et al. (2013)
4	covid, 19, pandemic, monitoring, detection	pandemic monitoring detection COVID-19	Channa, Popescu and Malik (2020) Chakkor et al. (2021) Lagos-Ortiz et al. (2020)
5	learning, machine, disease, data, using	machine learning using disease data	Kavakiotis et al. (2017) Pereira et al. (2019) Ferreri et al. (2018)

data analytics in healthcare, which are: (i) traceability in patient monitoring, including lab results, medication, historical data, and current status; (ii) analysis of structured and unstructured data, e.g., comparative analysis of images, voices, texts, etc; (iii) speed-up decision-making with automatic notifications or visual reporting; (iv) interoperability with the integration of heterogeneous data from different sources; and (v) prediction of patients behavior. These capabilities not only exemplify the possible applications of descriptive techniques in healthcare, but also point to the power they have to improve healthcare processes. For some authors, the fact that prediction is among the capabilities of descriptive analytics shows how the fields of data science and analytics are connected. Predictive analytics is the field of analytics in which future events are foreseen (Mishra and Silakari, 2012). As stated by Van Calster et al. (2019), the large investments in new AI technology reflects the value of this field to healthcare. This is due to its ability to diagnose individuals most likely to suffer from a disease, or to predict the evolution of diseases or viruses. This helps to make decisions in the treatment of patients, increasing the probability of recovery. It is also used in administrative issues to manage resources in seasons with a higher or lower probability of a high flow of patients in medical centers. From a general perspective, predictive techniques in e-health systems can make processes in the healthcare field more efficient.

Table 3. *Data science methods applied in e-health.*

Data science fields	Techniques	Examples of e-health applications	References
Descriptive analytics	Data retrieval & collection (statistics & data visualization)	Patients diagnostic, epidemic recognition, patients management with mobile apps, visual analysis tools for clinical data, reporting systems, social media analysis, and data warehouse tools.	Galetsi and Katsaliaki (2020) Wang and Hajli (2017)
Predictive analytics	Machine learning, probabilistic modeling, & statistical analysis	Disease prediction, software as a service, medical decision support system, patient flow and use of medical resource prediction.	Lepenioti et al. (2020) Van Calster et al. (2019) Mishra and Silakari (2012)
Prescriptive analytics	Logic-based modeling, evolutionary computation, mathematical programming, reinforcement learning, & simulation	Decision-making automation, scheduling problems, balanced assignment workload and resources, health management assessment, and cloud computing.	Javaid et al. (2021) Nandankar et al. (2021) Shah, Bhat and Khan (2021) Wijnhoven (2021) Bertsimas and Kallus (2020) Lepenioti et al. (2020)

There is also Prescriptive analytics, which is a field that prescribes optimal decisions for operations research (OR) and management science, together with machine learning techniques (Bertsimas and Kallus, 2020). In general, these techniques provide meaningful insights and support decision-making for organizations, thus giving them competitive advantages. This field is considered the evolution of the descriptive and predictive fields of data science, due to the capability of offering intelligent recommendations based on the analyzed and predicted data (Lepenioti et al., 2020). Since this is a field that is still in its infancy (Lepenioti et al., 2020), there are not too many works that integrate these data-driven techniques into e-health applications. Among the 25 papers available in the Google Scholar database for the first semester of 2021 and indexed under the terms "prescriptive analytics" and "e-health", some interesting applications are: *(i)* healthcare monitoring systems integrating a cloud IoT (Shah et al., 2021); *(ii)* a groundbreaking analysis system for large-scale healthcare data, which allows the use of fog computing and cloud systems to deal with data processing, storage, and classification problems (Nandankar et al., 2021); *(iii)* a dental 4.0 decision-support system to provide users with high-quality and personalized experiences (Javaid et al., 2021); and *(iv)* the implementation of an analytic clinical decision support system that analyzes medical data to predict the probability of sepsis in prematurely born infants, which can support physician decision-making on antibiotic stewardship (Wijnhoven, 2021).

With the goal of identifying some recent applications of data science methodologies in e-health, a similar study was carried out using the Scopus database. A total of 10 new papers were found for the first semester of 2021. These papers combined the terms "artificial intelligence", "e-health", and "data science" in the title, abstract, or keywords sections. Table 4 summarizes the most recent application fields of these techniques in

Table 4. *Identified data science applications in e-health within the literature available in Scopus database.*

Data science techniques	Application field
Data retrieval & collection (statistics & data visualization)	Epidemiological analysis (Pfeiffer and Stevens, 2015) Computer-assisted surgical skills (Vedula, Ishii and Hager, 2017) Preventive health management system (Neubert et al., 2019) Mental healthcare (Naslund et al., 2019) Correlations between clinical medicine subjects (Chen et al., 2020b)
Machine learning	Automatic behavior identification (Crocamo et al., 2020) Behavior assessment (Liang et al., 2020) Symptom classification (cardiology) (Oliver et al., 2018; Spanakis et al., 2017)
Modeling & simulation	Virus propagation risk analysis (Chatterjee, Gerdes and Martinez, 2020) Disease progression (Idrees and Sohail, 2021) Optimize emergency departments operations (Vanbrabant et al., 2019) Human organs simulation (Serra et al., 2021; Quarteroni et al., 2017)

e-health. Some of the applications that have been identified are those referring to the development of behavioral assessment systems (Spanakis et al., 2017; Crocamo et al., 2020) and to data management systems (Neubert et al., 2019), where data collection methods are implemented through different techniques. Spanakis et al. (2017) develop an adaptive feedback module of an e-coach application on eating behaviors. The authors use the acquisition of ecological momentary assessment (EMA) data through a mobile application. EMA is a set of methods that assess research subjects in their natural environment, in their current or recent states, at predetermined events of interest, and repeatedly over time (Moskowitz and Young, 2006). It allows data to be collected online and in real-time, thus generating more accurate and valuable results. As it is based on questionnaires, it reduces recall bias. It is a methodology that can easily be extended to include information from sensors, e.g.: assessing stress levels or GPS information to evaluate energy expenditure information (Spanakis et al., 2017). In addition, Crocamo et al. (2020) implement an automatic system that identifies keywords or *#hashtags* in Twitter publications that cite alcohol-related behaviors. They work with a systematic tracking process on Twitter using an ad-hoc Python script. After collecting the tweets, they classify all the information to filter out and identify the genuine users. Then, they implement an additional classifier focusing on the linguistic characteristics of the content. All data collection and processing methods are performed with different Python scripts based on the natural language toolkit framework and the scikit-learn library. For data management, Neubert et al. (2019) propose a decentralized system for preventive health based on multi-sensorial fusion (different devices), including heterogeneous data. They develop a mobile application that is the central data node for individual client monitoring, and also contains a data preparation system (data consolidation and synchronization, pre-processing, data selection, and reformatting of data sets) to be stored in the cloud in the required format. The data transfer between the mobile systems and the cloud is based on the formatting of the data to an arranged JavaScript object notation protocol.

In addition to data extraction, cleaning, and transformation, data visualization methodologies have also been applied in the area of e-health in order to extract essential and useful information in support of decision-making. An example of this is the "surgical data science" system. Vedula et al. (2017) conduct a literature review on computer-assisted objective technical capability assessment systems that allow for scalable and accurate assessment of surgeons. They also provide visualized feedback and automated training, thus improving surgical training and maintenance of surgical skills. These systems use data science techniques, such as summary features to represent data, time-series data representations, dictionaries or histogram-based representations, classification algorithms, etc. In more general areas of medical research, authors such as Chen et al. (2020b) explore and comprehensively compile topics from the clinical medicine literature by manually labeling the topics of diseases. In addition to searching in scientific databases and cleaning and integrating data in Microsoft Excel, they apply visualization techniques based on descriptive statistics (e.g., histograms, spatial distribution, etc.) developed in the R statistical software. The results allow them to reach interesting conclusions regarding the relationship to data-driven studies in the field of medicine and health. Pfeiffer and Stevens (2015) provide a literature review on the analysis of temporal and spatial data to support the management of complex animal health problems. According to their conclusions, the opportunity offered by digital technologies in animal and human health requires an interdisciplinary approach that, in addition to the various health areas, also includes information technology.

Data mining is a sub-discipline of data science characterized by discovering knowledge in large databases or extracting patterns from them (Kriegel et al., 2007). It lies at the interface of database technology, pattern recognition, machine learning, and other areas (Hand, 1998). Data mining techniques are also implemented in some areas of e-health, such as classification and clustering algorithms for human behaviors and heartbeat categories analysis (Liang et al., 2020). Artificial neural networks (ANN) are applied for heterogeneous and multidimensional data analysis in general healthcare and psychological interventions (Oliver et al., 2018), respectively. Data mining makes it possible to uncover patterns or trends in human behavior and illness trajectories that were previously not visible. Still, there is a need for large amounts of data in different healthcare areas, such as in mental health (Naslund et al., 2019) or surgical skills for assessment frameworks (Vedula et al., 2017).

The few applications available today in the literature show how implementing these methodologies supports several healthcare processes, ranging from physicians' training and evaluation systems to the use of diagnostic and follow-up systems, including different administrative processes. Furthermore, the number of articles identified indicates that the application of data science methods and the adoption of the terms e-health and data science together have not been widely used until now. The low number of articles explicitly including the term "data science" in their description reflects that the application of these methodologies in e-health is at an early stage yet. Furthermore, many of the articles are literary reviews seeking to establish theoretical foundations and to identify new research lines related to the applications in e-health, rather than the actual development of methodologies, case studies, and applications.

Table 5. *Number of papers per year by combining the term "e-health" with different AI subfields.*

subfield	Year																				Total
	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	
Deep learning	0	0	0	0	0	0	0	0	0	0	0	0	0	2	2	2	8	17	20	34	85
Reinforcement learning	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	2	3	3	5	15
Clustering	0	0	0	1	0	2	0	0	5	0	1	4	3	5	11	9	8	12	5	8	74
Data visualization	0	0	1	0	1	2	0	1	3	5	2	2	1	4	1	2	3	9	3	5	45
Artificial neural network	0	0	0	0	1	2	1	1	1	0	3	2	3	0	2	2	4	5	3	10	40
Natural language processing	1	0	0	0	0	0	1	1	1	0	2	4	4	3	3	2	3	9	4	5	43
Fuzzy logic	0	0	0	1	1	0	1	1	2	0	5	1	4	2	6	4	2	1	2	10	43
Bayesian networks	0	0	0	0	0	1	4	0	0	0	9	1	3	1	2	0	2	1	3	3	30

4. Artificial intelligence algorithms in e-health

As shown in Figure 1, the interest in AI applications to e-health has been quickly rising during the last decade. Among the several fields that can benefit from AI, the healthcare community shows a particular interest due to the amount of data and information that new technologies can provide. As pointed out by Oke (2008), inside the AI field we can find subfields that should be separately considered: machine learning, fuzzy logic, artificial life, data mining, Bayesian networks, knowledge engineering, ANN, reactive systems, semantic networks, computational language, natural language processing (NLP), etc. In particular, machine learning and data mining are subfields that in themselves encompass many subfields of AI, such as deep learning, reinforcement learning, clustering, data visualisation, etc. Data mining can extract usable information from immense raw data sources, primarily as a solid input for subsequent advanced downstream data analyses processes. On the other hand, machine learning has shown to have considerably broad applications in healthcare. Considering the above, Table 5 shows the evolution of the number of papers over the last two decades for the most popular subfields. The number of articles have been extracted from the Scopus database. Deep learning has been the most frequently employed approach in the last years, followed by clustering, data visualization, NLP, ANN, Bayesian networks, and reinforcement learning.

Thus, all these subfields lead to innovations and discoveries in all aspects of medicine, which may have persuaded researchers to focus more on exploring the potential of machine learning in healthcare research, such as the implementation on deep learning compared to data mining techniques like clustering and data visualization. However, these subfields are not worked on separately. Figure 4 presents the number of papers in which techniques from the different subfields analysed have been combined. This information is constructed after a combined search of the Scopus database for "e-health" and two of the subfields analysed. According to the figure, the subfields that have so far been studied together are clustering with all but deep learning and reinforcement learning. These last two subfields are worked together, with deep learning being the only one that is related to ANN, Bayesian networks, and NLP. Fuzzy logic also seems to be of great interest when working with ANN, clustering, and Bayesian networks.

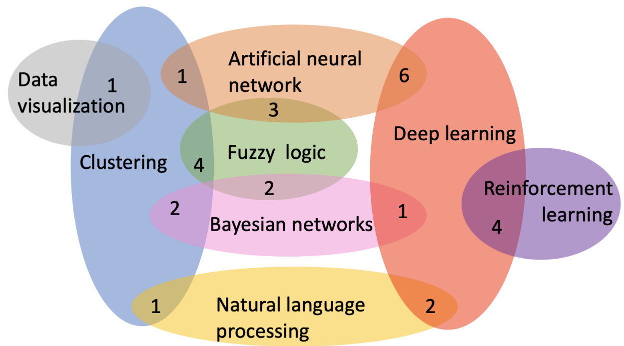

Figure 4. *Number of papers by combining the term "e-health" with two different AI subfields indexed in the Scopus database.*

This section will provide an overview of some of the aforementioned methodologies, as well as a discussion on their applications in e-health. The methodologies have been selected based on the number of publications during the last years. For a deeper discussion on how these methods are employed, the reader is referred to the works included in Table 6.

4.1. Machine learning

The growing use of devices that collect health data (e.g., wearables) favors the development of both, supervised and unsupervised methods, in diagnostics and diseases prediction. These devices allow collecting health data such as heart rate, number of steps, calories, sugar levels, hours of sleep, or images. All of these data can serve as inputs for the machine learning algorithms.

The main objective of algorithm development in the scientific community is to make computers completely autonomous in predicting results. However, in the health field, autonomous machine learning is far from being implemented. Consequently, the integration of a human expert in the circuit allows for better approaches in the complex field of e-health. The interactive machine learning approach can be defined as algorithms that can interact with human or computational agents to optimize the learning process (Holzinger, 2016). In e-health, the algorithms used in the interactive approach can be of particular interest in problems where the lack of data to train models or decision-making can be replaced by the help of the medical experts, where their knowledge and experience can be of significant contribution in solving the problem that would otherwise remain too complex. Some examples where this approach is used are described by Hund et al. (2015) and Lathrop (1994), where subspace clustering and protein folding problems are presented, respectively.

In contrast with the benefits mentioned, interactive machine learning approaches present an open question about their robustness: their evaluation is not only more diffi-

Table 6. *AI surveys by subfields.*

AI field	Topic	Reference
Machine learning	Summarizes a variety of machine learning research techniques in health informatics.	Dua, Acharya and Dua (2014)
	Studies the concept of interpretability in the artificial intelligence field in healthcare.	Al-Garadi et al. (2020)
	Reviews some ML algorithms used for developing efficient decision support for healthcare applications.	Shailaja, Seetharamulu and Jabbar (2018)
Reinforcement learning	Surveys of applications in healthcare, focusing on the discovery of new treatments, personalizing existing ones, and automated medical diagnosis.	Yu et al. (2021)
		Coronato et al. (2020)
		Gottesman et al. (2019)
Data mining	Reviews the utility of various DM techniques like regression, clustering, association, classification in healthcare.	Tomar and Agarwal (2013)
		Koh and Tan (2011)
		Jothi, Rashid and Husain (2015)
		Birnbaum (2004)
Artificial neural networks	Applications of ANN to health care organizational decision-making.	Lisboa and Taktak (2006)
		Shahid, Rappon and Berta (2019)
Natural language processing	Summarizes the applications, techniques, principal challenges of NLP in healthcare.	Friedman and Elhadad (2014)
	Discusses the main challenges of NLP in the healthcare sector.	Carrell et al. (2017)
	Reviews the NLP techniques used in healthcare, their applications and limitations.	Iroju and Olaleke (2015)
	Presents a system that combines both text mining and NLP.	Popowich (2005)
	Focuses on development and advances of NLP methods for clinical decision support.	Demner-Fushman, Chapman and McDonald (2009)

cult, but also these methods are not easy to replicate. The reason is that the contributions provided by human agents are subjective and cannot be easily imitated. Some applications of machine learning to diseases prediction can be found in Senders et al. (2018), Goldstein, Navar and Carter (2017), Weng et al. (2017), Churpek et al. (2016), Taylor et al. (2016), Kruppa et al. (2014), and Singal et al. (2013).

Reinforcement learning is a powerful and increasingly popular subfield of machine learning, which studies optimal sequential decision-making under uncertainty. This approach is highly relevant in the fields of dose optimization (Hrinivich and Lee, 2020; Tejedor, Woldaregay and Godtliebsen, 2020; Watts et al., 2020) and robotic-assisted surgery (Gao et al., 2020; Pore et al., 2021; Su, Huang and Hannaford, 2021). There are also other recent works on public health (Weltz, Volfovsky and Laber, 2022; Kwak, Ling and Hui, 2021), surgical decision-making (Datta et al., 2021), physical activity mobile health application (Liao et al., 2020), and cancer detection (Liu et al., 2019), among others.

4.2. Data mining

The application of data mining (Fayyad, Piatetsky-Shapiro and Smyth, 1996) can provide knowledge useful to support clinical decision-making. In the healthcare and biomedical fields, data mining deals with privacy and legal issues, as well as on improving the quality of data available in these disciplines. Paper-based or scanned-digital formats, heterogeneity of hospital information systems, or different lab tests for the same disease constitute big challenges for researchers. Here is where data mining can help by providing tools to extract information from large and unstructured data sources.

Data mining algorithms can be classified into two categories: descriptive and predictive (Tamilselvi and Kalaiselvi, 2013). The first one aggregates records with similarities, thus discovering unknown relationships in data. The second one deduces prediction rules from training data and applies these rules to unpredicted data.

4.3. Artificial neural networks and natural language processing

ANN are computer programs that try to reproduce how the human brain processes information, i.e.: learning through experience, by recognizing patterns and relationships in data. An ANN is constituted by hundreds of single units (artificial neurons), which have weighted inputs and one output. In healthcare, deep neural networks face big challenges that need to be addressed before they can be used in the daily life of patients. They require an exceptionally huge amount of information to function better than other strategies. They are extremely costly to train due to complex information models. Besides, deep learning needs costly GPUs and hundreds of machines. This enhances its costs of operation. There is not enough standard theory to direct people in choosing proper deep learning resources because it needs information on topology, preparing strategy, and other necessities. Therefore, it is hard to be used by less knowledgeable end users (Kim et al., 2020).

When managing clinical reports, one of the biggest challenges for analysts is that information is not structured. In fact, it is usually written in natural language as plain text. Hence, NLP techniques are responsible for extracting knowledge from this unstructured data, analyzing the information it contains, and providing it in a format that the electronic healthcare systems can easily understand. More than twenty years ago, Friedman and Hripcsak (1999) and Baud, Rassinoux and Scherrer (1992) had already highlighted the importance of NLP in medicine. Friedman and Elhadad (2014) ranks the top challenges regarding NLP and also concludes that software-generated recommendations should always be supervised by human decisions.

4.4. Use of AI in e-health

Many studies have demonstrated the wide applications of AI in e-health. As stated in Jiang et al. (2017), AI "is bringing a paradigm shift to healthcare". These authors provide a literature review on AI applications in healthcare, including areas such as cancer, neurology, and cardiology. A similar review is provided by Yu et al. (2018), who support the idea that AI has contributed to improving diagnosis and decision-making in healthcare, while recognize its potential in fast disease detection and customized treatments. Authors such as Davenport and Kalakota (2019) or Emanuel and Wachter (2019) highlight the possibilities that AI offers in areas such as automatic examination of radiology and pathology images. However, they suggest that the main challenge AI has to face in healthcare is not a technological one, but a cultural one –so that these techniques are adopted and employed in daily clinical practice. He et al. (2019) identify the main challenges regarding the practical implementation of AI into daily healthcare practice. Among these issues, they include data sharing and privacy as well as algorithms' transparency. According to Reddy, Fox and Purohit (2019), patient administration, clinical decision support, patient monitoring, and healthcare interventions are the healthcare areas in which the use of AI can be more beneficial.

5. Data science & AI best practices in e-health

Most e-health tools not only include the use of Internet-based applications, but also products, systems, and services. Health portals, telemedicine services, electronic health records, or health information networks are just some examples. As already pointed out in several works (Widmer et al., 2015; Triantafyllidis et al., 2015; Warmerdam et al., 2010), AI methods have an enormous potential for improving healthcare, reducing costs, and developing smart digital health interventions. Such enhanced interventions can lead to remarkable outcomes, both for patients and healthcare providers (Murray et al., 2016; Obermeyer and Emanuel, 2016). This section presents relevant application domains related to the application areas identified in Section 2. It is structured in five subsections: patient care, public health, healthcare management, COVID-19, and other topics.

5.1. Patient care

Regarding patient care, we identify the following areas of applications:

- *Diagnosis:* A large number of works propose an AI methodology to either perform or assist the expert with the diagnosis. For instance, Kermany et al. (2018) create a diagnostic tool based on a deep-learning approach for screening retinal and lung pathologies for age-related macular degeneration and pediatric pneumonia, respectively. The diagnostic results of the training that neural network framework was comparable to that of health professionals. Indeed, that AI-based diagnostic method was even able to provide more detailed diagnoses, especially at the onset of pathologies when the clinical symptoms might not be apparent to healthcare providers. Similarly, Guo et al. (2020) propose a deep learning model for real-time automated diagnosis of precancerous lesions and also to assist the diagnosis of esophageal cancer. The model has high sensitivity and specificity for both endoscopic images and videos. Additionally, NLP has been used in the diagnosis of diseases such as Aphasia, a neurological disorder (Rao and Venkatesh, 2021). It is also applied in medical decision support systems for the diagnosis of patients. This is performed by semantic analysis of their available medical records (Amato et al., 2018; Iram and Gill, 2018). Other solutions are focused on bringing the healthcare sector online. Virkar et al. (2021) develop a system that provides information on remote doctor connection and online identification and treatment, and describes the prediction of various diseases using a web application as an interface to store the data in a database. The NLP collects the necessary symptoms and sends them to the doctors. This is used to predict the exact disease the patient may have using the Random Forest algorithm. Thus, utilizing such AI-based frameworks can lead to earlier and more accurate diagnoses, which are crucial for curving the progression of diseases and selecting the most effective treatments for patients.

- *Personalized medications and diets*: The traditional "one-size-fits-all" approach is inefficient in many scenarios. AI is increasingly playing a pivotal role in personalized medications. An example is related to Antimicrobial resistance (AMR), a phenomenon in which microorganisms resist antimicrobial drugs. AMR poses a serious risk to preventing and treating infections caused by viruses, bacteria, and fungi. AMR will lead to nearly 10 million death and 100 billion dollars economic damage every year by 2050 (O'Neill, 2014). AI-based predictive models can be applied against AMR. These models can be used in predicting AMR and suggesting the best dose and time of antimicrobial treatment, as well as the best combinations of antimicrobial peptides and antibiotics for each patient (Lv, Deng and Zhang, 2020). Zeevi et al. (2015) use factors –such as gut microbiota, anthropometrics, blood parameters, and dietary habits– to create a gradient-based boosting regression for predicting the post-meal glycemic fluctuations in real life in a randomized controlled study. The results showed that the personalized diet created according to the post-meal glycemic simulation significantly optimized the post-meal glucose level in the study individuals. Such approaches can be used in various nutritional interventions in different diet-related medical problems, such as diabetes, obesity, and nonalcoholic fatty liver disease.

- *Personalized care:* AI also contributes to personalize care and interventions. For instance, Barrett et al. (2019) perform a study in which a "virtual doctor" is designed relying on AI, serious gaming, and patient coaching. This doctor boosts advanced and personalized self-care for heart failure patients, where the patients themselves perform standard care tasks. Burns et al. (2011) use a mobile-based, multi-component type of intervention that worked based on some machine learning models in order to predict the emotions, activities, mood, motivational and cognitive states –as well as social and environmental changes– for some patients who were suffering from depression. That intervention also enabled the patients to have access to the coaching feedback graphs provided by their caregivers. At the end of the study, nearly 90% of the patients indicated that the machine learning-based mobile intervention helped them treat their depression and control their mental symptoms, and they were satisfied with that.

- *Treatment optimization:* Planning a treatment tend to be a challenging and key task. Many authors have proposed AI-based applications to enhance this process. For instance, Wang et al. (2019) review AI-based applications for radiotherapy treatment planning aiming to minimize the normal tissue damage while persevering sufficient tumor control. Similarly, Cui et al. (2021) explore the use of random survival forest models to predict optimal regimen classes for individual patients and each line of therapy relying on baseline characteristics. The patients are adult females with HR+/HER2- breast cancer and the aim of the models is to maximize overall survival and time to treatment discontinuation based on electronic health records. Shamir et al. (2015) propose a clinical decision support systems based on machine learning clinical decision support systems based on machine learning (support vector machines, naïve Bayes, and random forest are considered) to optimize combined stimulation and medication therapies for Parkinson's disease.

- *Assisted or automated prescription:* In the context of a growing amount of available clinical data and an increasing focus on personalized care, assisted or automated prescription constitutes one of the most promising AI applications. For instance, Blansit et al. (2019) propose an approachable to prescribe imaging planes for cardiac MRI based on deep learning, which relies on the localization of anatomic landmarks. Barbieri et al. (2019) train a model to manage blood pressure, fluid volume, and dialysis dose in end-stage kidney disease patients. According to the authors, these models can help to anticipate patients' reactions through simulation, which can help choose the best treatment for each patient. Leaflets contain information on the administration of medicines, their composition, warnings or precautions for use that is difficult for patients to understand. Dascalu et al. (2019) have designed an intelligent platform based on NLP techniques for drug administration. Its functionalities include adding specific drugs to the user's profile, searching for possible contraindications or side effects and defining drug administration alerts on schedule, based on the doctor's prescription. This enables improved healthcare for users and ensures self-education.

- *Triage:* Determining the priority of patients' treatments by their condition or likelihood of recovery is critical in hospitals, especially during a crisis. For instance, Kim et al. (2018) propose a classification model for survival prediction to perform a precise triage. The authors compare the performance of different methodologies, for example, logistic regression, random forest, and deep neural networks, and design a consciousness index capable of remote monitoring through wearable devices.

- *Surgery:* According to Hashimoto et al. (2018), AI has the potential to revolutionize the way surgery is taught and practiced. The authors review machine learning applications, ANN, NLP, and computer vision. They also point out the models' low interpretability and the difficulty to determine causal relationships as the main limitations of AI. Riise, Mannino and Burke (2016) propose a new generalized model for surgery scheduling problems. Zhong et al. (2014) present a decision support system to develop master surgery schedules. Bai, Storer and Tonkay (2017) present a gradient-based algorithm in order to improve some aspects, like the cost incurred from patient, waiting time, blocking time, operating rooms overtime, etc.

- *Pregnancy management:* Management of the pregnancy aims to reduce child and maternal mortality by increasing pregnant women's access to high-quality health services. For instance, Moreira et al. (2019a) design smart mobile-health applications that use machine learning for pregnancy monitoring. These applications are able to predict high-risk situations during gestation. Predicting the risk of post-partum depression during pregnancy through biomedical and socio-demographic data analysis constitutes another task where AI may be potentially useful (Moreira et al., 2019b).

5.2. Public health

Best practices regarding the use of data science and AI in public health have also been documented in recent articles. Some of the most relevant ones are included next:

- *Forecasting demand for emergency department services:* Accurately predicting the demand of services in an emergency department is essential to provide efficient and high-quality services. Traditionally, these predictions have been made using historical data and experts' opinions, but some recent works have attempted to add Internet search data. Ho et al. (2019) analyze search volume data retrieved from Google Trends, and construct regression-based predictive models. The method used is relatively simple –multiple linear regression models–, but represents a useful tool to address congestion problems. Likewise, Martin et al. (2012) use the information from patients' phone calls to develop a model to alert health professionals that could predict when patients with critical conditions, such as those who suffered from cardiovascular and lung diseases, who also did not have appointments would go to the medical center for treatment. When his model was deployed over six months, it was able to predict nearly 70% of the unplanned events, which helped give more time to the health system to manage their resources better for admitting those patients. A number of experts have been working in this subject, developing forecasting strategies to preview the arrivals to emergency department (Billings et al., 2013; Kam, Sung and Park, 2010; Hoot and Aronsky, 2008; Hoot et al., 2007).

- *Screening:* In many countries, there are population-based screening programs, with the cancer screening program as one of the most common. These programs may have restrictions due to budget constraints and lack of experts. Chen et al. (2020a) describe an example of the potential of AI in this context. The authors present a wristband device based on AI to detect atrial fibrillation (AF). AF is a medical condition in which irregular heartbeat can lead to stroke or cardiac arrest. Detection of AF can be challenging as some patients may not have the symptoms while been screened. The wristband was equipped with sensors that measured single-channel electrocardiogram (ECG) and photo-plethysmography (PPG) and an AI algorithm designed to detect AF based on the ECG and PPG input. The reading accuracy, specificity, and sensitivity of wristband were 93%, 96%, and 88% for PPG, and 95%, 99%, and 87% for ECG. Some physicians also evaluated the wristband-recorded ECG. Their accuracy, specificity, and sensitivity of the physicians' judgment were 97%, 98%, and 97%, which were close to those of the wristband algorithm. The convenience of using this method has great potential for long-term screening patients that may suffer from AF, especially in individuals that may have minimal or inconsistent detectable AF symptoms. Another example is described in Bao et al. (2020), which explores an AI-assisted cytology system in a cervical cancer screening program. It improves sensitivity with clinically equivalent specificity.

- *Epidemics:* Data analysis is critical to track outbreaks and design effective strategies to curve epidemics. For example, Ray and Reich (2018) make predictions of infectious disease dynamics with ensemble methods. In particular, the authors predict influenza season timing and severity measures in the United States, both at the national and regional levels. Ganasegeran and Abdulrahman (2020) analyze the role of ineffectively preempting, preventing, and combating the threats of infectious disease epidemics, as well as facilitating the understanding of health-seeking behaviors and public emotions during epidemics. In the biomedical field, text analysis is performed to identify and extract disease symptoms and their associations from biomedical text documents retrieved from the PubMed database using NLP and information extraction techniques to identify feasible disease symptoms (Abulaish et al., 2019).

- *Facing fake news:* The spread of inaccurate information on the internet is a daily occurrence. The impact on people's lives is one of the majors concern in the field of public health. Most of the intelligent techniques tackle this problem are developed mainly using NLP and machine learning (Mesquita et al., 2020). For example, Pulido et al. (2020) conduct an analysis of social media such as Reddit, Facebook and Twitter where they identify that messages focused on false health information are mostly aggressive, those based on evidence of social impact are respectful and transformative. Parfenenko et al. (2020) propose an ANN for the classification of publications on medical care topic into true and misinformative in different WordPress forums to manage this type of publications.

5.3. Healthcare management

We list several types of applications in healthcare management. For each type, a brief description and a reference to a recent example are also provided.

- *Healthcare logistics:* Logistics constitute a strategic function of hospitals' management, senior citizens' rest homes, pharmaceutical companies, etc. The decision-making in this field aims to reduce errors, enhance process quality and reduce waiting times. An increasingly popular topic is home healthcare, which has been boosted thanks to unstoppable technological progress. An example is described in Fikar et al. (2016), where the authors develop a discrete-event driven metaheuristic for dynamic home service routing with synchronized trip sharing. Likewise, in Lostumbo et al. (2021), the authors propose a hybrid method, combining simulation with reliability analysis, to improve supply chains in the healthcare sector. A review of other works related to home healthcare logistics can be found in Fikar and Hirsch (2017), where several applications and different approaches are enumerated.

- *Resource forecasting and optimization:* The increasing demand for resources and the limited capacity in the healthcare sector has increased the use of tools based on forecasting, simulation and optimization for resource management. Some works optimize, for example, the management of hospital beds and the personnel required for each bed modeling the problem as an integer linear programming models, forecast the demand for specialists with ARIMA and linear regression models (Ordu et al., 2021). Others plan patient capacity and patient post-hospitalization fate using decision models based on survival trees (Garg et al., 2012). Ganguly and Nandi (2016) use analysis of variance (ANOVA) to identify drivers of demand, and autoregressive integrated moving average (ARIMA) to develop a forecasting model for optimal healthcare staff scheduling based on patient arrival rates. Similar models have been developed to predict patient visits in the emergency department (Khaldi, El Afia and Chiheb, 2019), and the demand for health diagnostic service such as endoscopy service (Harper, Mustafee and Feeney, 2017). Ellahham and Ellahham (2019) present a review on the applications of artificial intelligence to improve patient management and resource allocation in hospitals.

- *Medicine supply chain network design:* Logistics activities are essential for an efficient and sustainable distribution of medicines. Recently, Goodarzian, Hosseini-Nasab and Fakhrzad (2020) propose a hybrid particle swarm optimization and a genetic algorithm to achieve Pareto solutions for the design of a medicine supply chain network.

- *Treatment/surgery scheduling:* When scheduling surgeries or treatments, such as chemotherapy, a number of objectives and restrictions related to costs, elapsed times, available rooms, experts, and equipment, have to be optimized. For example, Belkhamsa, Jarboui and Masmoudi (2018) present two metaheuristics –an iterative local search approach and a hybrid genetic algorithm– to address the operating room surgery scheduling, with resource constraints in three stages, namely: preoperative, intraoperative, and postoperative stages. The goal is to minimize the maximum end time of the last activity in stage 3 and the total idle time in the operating rooms. In a different work, Martins et al. (2021) propose a metaheuristic optimization algorithm to support medical staff when assigning and scheduling treatments to cancer patients. The reader interested in healthcare scheduling problems is referred to Abdalkareem et al. (2021), which covers patients' admission scheduling problems, nurse scheduling problems, operation room scheduling problems, surgery scheduling problems, etc.

- *Healthcare facility location-allocation:* Making location-allocation decisions related to healthcare facilities may be a challenge because of multiple conflicting objectives and stakeholders. In this context, Wang, Shi and Gan (2018a) put forward a practicable hierarchical model to characterize the trade-off between social, economic, and environmental factors. The authors present a bi-level multi-objective particle swarm optimization algorithm to deal with the location decision and capacity adjustment.

- *Assessment of the hospital performance:* This constitutes a challenging task because of the high number of related indicators. For instance, Downing et al. (2017) build a semi-supervised machine learning algorithm, which highlights the similarities and differences between hospitals and detects hospital performance patterns for 1 614 U.S. hospitals.

- *Brand management and marketing:* Plenty of strategies in these fields are data-driven. For example, Oztekin (2018) develop data analytic models to help marketing managers identify locations to host peer-to-peer educational events for healthcare professionals.

- *Pricing and risk:* Data has always played an important role in insurance. The increasing amount of available data may improve the process of pricing and risk management. For example, Kshirsagar et al. (2020) assess machine learning models aiming to predict the per member per month cost of employer groups in their next renewal period. The authors conclude that these models may compute an accurate and fair price for health insurance products without losing interpretability.

- *Fraud detection:* Medicare fraud, waste, and abuse cause huge losses, but traditional detection methods tend to be time-consuming and have low accuracy. In recent years, AI has been extensively used for fraud detection, and several works can be found in the field of e-health. It is challenging because of class-imbalance. Zhang and He (2017) propose a method for medicare fraud detection. It consists of two parts: first, a spatial density-based algorithm, called improved local outlier factor; second, a robust regression to analyze the linear dependence between variables. Johnson and Khoshgoftaar (2019) choose the medicare fraud detection task to compare several deep learning methods built to deal with the class imbalance problem. The authors employ different data-level techniques, such as random over-sampling, random under-sampling, a hybrid one, and several algorithm-level techniques, such as a cost-sensitive loss function focal loss, and the mean false error loss.

- *Patient/user satisfaction:* Patient-oriented interactive tools are a great source of data that provide evidence for strategic planning for e-health development. An analysis of US hospital websites concludes that most hospitals need basic e-commerce tools for their patients/users (Huang and Chang, 2012). A study of more than 200 patients in the USA developed by Huang, Chang and Khurana (2012a), showed that they are interested in access to information such as medical records and lab results. In addition to the patient perspective, it is also noted that there are no ways to measure the success of e-health implementations in serving their patients/users. However, there are some examples of e-health systems that provide information on user/patient satisfaction. Silva et al. (2018) present a satisfaction and usability evaluation of a web-based clinical decision support system called HADA for antenatal care assisting in obstetric risk assessment. This provides a more effective and efficient use of resources and increases the capabilities of professionals and satisfaction for both professionals and patients. Similarly, Lan Hing Ting et al. (2021) study the development of the use of a robotic assistant for geriatric patients. Their results show that the implementation is feasible, as the performance and user satisfaction is promising. Rubrichi, Battistotti and Quaglini (2014) present a system for automatic evaluation of users' perception of the quality of an outpatient visit reminder system based on the short message service (SMS). The automatic interpretation of the content of these messages is useful for monitoring and improving health service performance.

5.4. COVID-19

The COVID-19 pandemic is severely affecting health systems and economies. In this subsection we review some recent examples of AI applications aimed at fighting the COVID-19 pandemic as well as other diseases.

Regarding the molecular aspects of this disease, AI can be used to identify and visualize the molecular structures of the SARS CoV-2 proteins, evaluate different existing drugs, and design new medicines that may help control the disease. Data science methods might also be critical for vaccine development, the design of more accurate diagnosis methods, and increasing our knowledge about the disease's molecular and clinical pathology. For example, Jumper et al. (2020) have designed a model called AlphaFold that predicts the three-dimensional structures of proteins based on their amino acid sequences. AlphaFold has been used to identify several SARS-Cov-2 proteins structures.

Richardson et al. (2020) have used the biomedical knowledge graph method and predicted that Baricitinib –a drug that is often used for the treatment of arthritis– can be used against COVID-19 since this drug inhibits the AP2-associated protein enzyme. As a result, it would be harder for the virus to enter host cells. Hofmarcher et al. (2020) have screened nearly 900 million compounds to estimate their efficacy for the inhibition of 3C and papain-like proteases using a long short-term memory model (Hochreiter and Schmidhuber, 1997). They used important factors, such as toxicity, predicted inhibitory effects, and proximity to known compounds in order to rank them. Finally, they selected 30 000 candidates for further screening.

Regarding the detection of coronavirus, AI methods for automated classification of COVID-19 on computed tomography scans are up-and-coming, as shown by numerous articles (Kundu et al., 2020; Elaziz et al., 2020; Li et al., 2020). Several companies have started developing AI-based apps that work as a COVID-19 health passport, which shows people's vaccination records and COVID-19 disease history and possible exposure to the virus based on interactions with people who might have been positive for the virus (Milliard, 2020).

In the social context of the pandemic, it has become important to control the spread of information online. More than 1 000 fake news were spread during COVID-19 generating a major social problem of misinformation about the disease (Naeem, Bhatti and Khan, 2021). Some intelligent techniques such as transformer-based algorithms, NLP, and supervised learning algorithms, have been implemented for data analysis, information extraction and identification of fake news related with COVID-19 pandemic (Gundapu and Mamidi, 2021). De Magistris et al. (2022) present an automatic fake news detection system based on different techniques from machine learning, deep learning and NLP to check medical news and, in particular, the reliability of publications related to the COVID-19 pandemic, the vaccine and the cure. Similarly, Mookdarsanit and Mookdarsanit (2021) develop a NLP model to identify Thai fake news related to COVID-19. A systematic review of articles indexed in journal citation report on e-health to combat COVID-19 developed by Alonso et al. (2021), provides a guide to deepen the applied work of data science and AI in e-health related to the pandemic. Similar analyses have also been conducted by other authors such as HassanAbady and Ganjali (2021); Monaghesh and Hajizadeh (2020); and Doraiswamy et al. (2020).

5.5. Other topics

The progress of personalized medicine is boosted by the development of *omics* technologies (such as genomics, transcriptomics, proteomics and metabolomics). DS and AI are essential to combine diverse types of omics data, analyzing them, and using the resulting models. Omics technology have the potential of providing a more complete view of biology and disease. Related applications may be found for diagnosis (Ma et al., 2020), prognosis (Poirion et al., 2021), and treatment. In this context, Karczewski and Snyder (2018) describe the utility of combining diverse types of data and the potential applications in human health and disease.

Another field that AI is starting to revolutionize is *drug discovery and design*. Heifetz (2022) describes a number of applications of AI, machine learning, and deep learning in drug design. These new approaches accelerate traditional drug design approaches such as: structure- and ligand-based, augmented and multi-objective de novo drug design, SAR and big data analysis, prediction of binding/activity, and ADMET, among others. This book covers cutting-edge techniques and lists the required software. Yang et al. (2019) explain the basic principles of learning tasks of techniques and describe the state-of-the art of AI-assisted pharmaceutical discovery, covering applications in structure- and ligand-based virtual screening, de novo drug design, physicochemical and pharmacokinetic property prediction, and drug repurposing. Similarly, Jing et al. (2018) discuss applications, limitations, and lines of future research, but focusing on deep learning –including convolutional neural networks, recurrent neural networks, and deep auto-encoder networks.

6. Insights and open challenges

This section offers a brief discussion based on the articles analyzed in the previous sections. We focus on drawing insights from the aforementioned articles, identifying open challenges, and proposing future lines of work.

6.1. Insights from the Literature

According to the results of our search in the Google Scholar and Scopus databases, the number of articles applying data science methods and AI algorithms in e-health has being systematically increasing over the last decade.

The raising interest in these methods and algorithms is a direct consequence of the huge amount of available healthcare data, as well as on the growing demand for new methods and tools that support decision-making in increasingly complex healthcare system. For instance, the growing design and use of smart devices that collect health data favors the use of data analytics and AI techniques in diagnostics and diseases prediction.

A wide range of methodologies have been adopted in e-health already, e.g.: modeling, simulation, statistics, machine learning, data visualization, etc. In particular, it is relevant to highlight the use of ANN in general healthcare and psychological interventions, as well as the use of machine learning to disease prediction. There are several authors that propose an interactive machine learning approach, in which algorithms interact with human or computational agents to optimize the learning process and improve the results. Given the large amount of unstructured data in the sector (e.g., clinical data and diagnostic information in text), NLP techniques are gaining prominence.

Among the different healthcare areas, those where the use of AI can be more valuable are patient administration, clinical decision support, patient monitoring, and healthcare interventions. The most popular topics studied with AI are related to cancer, depression, Alzheimer disease, heart failure, and diabetes. It is commonly accepted that the emergence of data science/analytics and AI represent a paradigm shift in healthcare. Today, the biggest challenge is a cultural one: it is needed that healthcare staff adopt and employ these techniques regularly. This can be achieved by reducing the gap between AI experts and healthcare staff. Among potential strategies, designing multidisciplinary curricula is the most promising one. Another challenge, regarding the practical implementation of these methods and algorithms, is the lack of a culture on data sharing among patients, hospitals, academia, and industry. Algorithms' transparency and interpretability constitute other barriers to the practical implementation of these approaches.

During the last years, the COVID-19 pandemic has evidenced the key role of data science/analytics methods and IA algorithms in e-health. A large number of authors have made related contributions aiming to identify and visualize the molecular structures of the SARS CoV-2 proteins, evaluate different existing drugs, design new medicines that may help to control the disease, and develop mechanism to detect coronavirus. Indeed, there are many big areas of applications where these analytical and computational tools are extensively being used, among others: patient care (e.g., in diagnosis, prescription, personalized medications and care, etc), public health (e.g., screening, epidemics, forecasting demand for emergency department services, etc), research and development (e.g., drug discovery, gene analysis and editing, etc), healthcare management (e.g., medicine supply chain network design, and treatment/surgery scheduling, etc).

6.2. Open challenges and future research lines

The growing volume of data in healthcare and the increasing complexity of decision-making processes is due, at least in part, to the variability in the evolution of diseases and their interaction with individuals, and gives rise to several challenges. The analysis performed in previous sections shows how data science/analytics and AI methods can efficiently support decision-making and disease diagnosis in different healthcare fields and their potential in the training of medical processes, disease control, and other many areas. According to the OECD (2020), the main challenges facing the integration of intelligent data-driven systems in the healthcare sector are associated with the heterogeneity of health and medical data. In this sector, data are not standardized, varying both between individuals and between populations and subfields. This can create cultural, racial, or geographic biases when transferring the applicability of models to patients or populations with different characteristics from the training data. There is also a risk that the quality and quantity of the data may be far from optimal, thus generating confusion between the noise and the real data during model training and preventing generalized models' development. Besides, much of the data provided in this sector is influenced by the human factor, as practitioners are typically responsible for providing the information. This results in errors, mistakes, and biases in the data, which affects the quality of the learning models.

One of the biggest barriers to the integration of data-driven systems in healthcare is the confidentiality and security of patient data. Information security laws around the world seek after a hazard- and process-oriented approach to guarantee the privacy, astuteness and accessibility of information and the strength of frameworks. This requires an intermittent handle to audit the adequacy of the security measures and their ceaseless enhancement. Information assurance is not a one-off action, but a task that must be inserted into all exercises relating to the management of health information systems. Similarly, information security could be an assignment and obligation of everybody involved in information handling, and ought to not be allotted solely to an information security officer or information administration division (Organization et al., 2021).

The safe and successful exchange of information must take into account protection issues. The implementation of the EU's Common Information Assurance Control (GDPR) has contributed to hindering the exchange of wellbeing information with analysts outside the EU/EEA. The Common Information Security Control (GDPR) addresses individual information security within the European Union (EU) and European Financial Area (EEA) and the universal exchange of information with areas outside the locality. Additionally, the use of GDPR has presented obstacles to this worldwide exchange of information with outside the EU/EEA, posing problems for academic analysts, healthcare professionals and others within the open division. The European Patients' Group has published data for patients on their rights to information and how exemptions to consent for research purposes (with specialized and authorized shields) should be monitored. The European Commission disseminates a master guideline for EU analysts on morality and information assurance, including universal information exchange. The Chamber of Universal Organizations of Therapeutic Sciences (CIOMS), in collaboration with WHO (CIOMS 2016), has created moral standards for health-related research consent (counting consent for unspecified future use) for the collection, capacity and use of organic tissue and related information. According to the GDPR, consent must be given unreservedly, in particular, educated and unambiguous. All this information is available at the report *International sharing of personal health data for research*, published by All European Academies (ALLEA), the European Academies' Science Advisory Council (EASAC), and the Federation of European Academies of Medicine (FEAM).

According to Blume (2015), the free development of information from the EEA is allowed in case there is a "suitability" choice for the beneficiary. The necessities for a nation to comply with the breadth guidelines are strict (Anghel and Drachenberg, 2019) and depend on whether solid safety standards are as of now connected inside that nation. Up to the present, he European Commission has recognized that some countries have satisfactory security (e.g. Andorra, Argentina, Canada (related to trade associations); Faroe Islands, Guernsey, Israel, Isle of Man, Japan, Shirt, Modern Zealand, Switzerland, and Uruguay). Therefore, there is no broad option for major research-intensive countries, such as China, Australia, USA and South Africa, and it is exceptionally unlikely to occur in countries that lack a legal framework to ensure protection, such as Australia or China. However, it is well known that intelligent methodologies must ensure algorithms transparency, robustness, and security. Risk management approaches (Sahoo et al., 2014) and best practices in applying the methodologies could ensure the protection and responsible use of data. Essential to this is political engagement –both nationally and internationally– in the healthcare sector through policy and regulatory reforms that compensate for data protection and the use of data in AI to deliver value to patients and society.

Another barrier is related to the rejection of the integration of AI tools into processes by medical practitioners. Healthcare professionals feel that digital tools could interfere with their patient care decisions. Understanding how they work is essential to influence their perspective on the use of smart technologies. Wang, Kung and Byrd (2018b) states that a data-driven healthcare organization needs data-driven environments that are well understood, reliable, accessible, and secure. Some strategies for the successful integration of smart technologies start from the foundation of strategic planning aimed at establishing a data-driven culture with a robust protocol that enables effective data utilization at all times. All healthcare personnel must acquire and foster a culture of sharing information horizontally –not only among themselves but also with providers and users, developers, and analysts. Likewise, this requires staff trained and skilled in intelligent data-driven technologies, who know the value of data in intelligent decision-making processes. Given the constant evolution of technologies, cloud computing becomes another requirement to meet the challenges of storage, processing, model development, and analysis of results.

There are several research lines that stem from the lack of multidisciplinary integration strategies between data science/analytics/AI scientists and healthcare experts, as well as the development and implementation of data-driven intelligent systems in the sector. The use of AI fosters the development of personalized systems for the treatment and prescription of patients. The development of personalized care systems allows to automate prescriptions and assist patients that require continuous monitoring. Rather than replacing doctors, this helps to enhance their efforts to care for their patients adequately. Other more generalized applications have come to light recently with the COVID-19 pandemic. Many of these show the potential of these models and tools in predicting and controlling virus dynamics and helping to understand and address the consequences in the healthcare sector and society. These applications can transcend to other health processes, such as predicting disease evolution, the scheduling of healthcare resources in different contexts, the study of long-term social impacts, etc.

7. Conclusions

This work has discussed the role of data science/analytics methods and AI algorithms in e-health, pointing out the emerging research topics, reviewing the existing literature, presenting some of the most popular methods and their applications to the healthcare industry, and highlighting the main challenges that need to be yet addressed in order to boost the use of data-intensive methods and algorithms in healthcare, and make the most out of them.

With the increasing use of mobile devices and sensors, large volumes of data can be collected now in real-time. Likewise, digitization processes are increasing the size of medical databases and the possibilities for searching and processing data in them, including enhanced data visualization services. Classification and clustering algorithms allow for more intelligent use of both preventive and reactive treatments. Regression models can also be employed to predict values related to illness trajectories. In the e-health context, one should consider interactive machine learning methods, which require human expertise to make efficient decisions when dealing with complex medical treatments. Apart from guaranteeing the quality of data and solving the heterogeneity-of-sources issue, other aspects that have to be considered when dealing with e-health data are related to the legal and privacy dimensions.

Artificial intelligence algorithms have shown to be effective in enhancing the quality of healthcare services (e.g., to speed up diagnostic processes, to personalize medications, to support surgery operations, etc.), increasing the efficiency of public health systems (e.g., by predicting the evolution of epidemics, by forecasting demand for hospital services, etc.), accelerating drug discovery, and supporting healthcare logistics and performance, among many other applications.

Regarding open challenges, these are probably more related to the need for a cultural change than for a technical evolution. In other words, the acceptance of a data- and algorithm-driven culture is needed instead of the traditional one, which is mainly based on the human subjective opinion. Hence, medical experts and healthcare managers need to get used to working hand in hand with data scientists, who will develop models, analyze data, and support the former when making complex decisions involving many variables. Likewise, the heterogeneity of data sources and the privacy and legal aspects associated with health data are relevant barriers that require a considerable effort to be reduced. All in all, the technology and the algorithms are quite advanced already to make better and more extensive use of them if those barriers are eliminated. Therefore, the next decade is expected to provide us with many novel applications of data science/analytics methods and AI algorithms in the healthcare industry.

Acknowledgments

This work has been partially supported by the Divina Pastora Seguros company.

References

Abdalkareem, Z. A., Amir, A., Al-Betar, M. A., Ekhan, P., and Hammouri, A. I. (2021). Healthcare scheduling in optimization context: A review. *Health and Technology*, 11:445–469.

Abulaish, M., Parwez, M. A. and Jahiruddin (2019). Disease: A biomedical text analytics system for disease symptom extraction and characterization. *Journal of Biomedical Informatics*, 100:103324.

Ahsan, M. and Bartlema, J. (2004). Monitoring healthcare performance by analytic hierarchy process: A developing-country perspective. *International Transactions in Operational Research*, 11(4):465–478.

Al-Garadi, M. A., Mohamed, A., Al-Ali, A. K., Du, X., Ali, I., and Guizani, M. (2020). A survey of machine and deep learning methods for Internet of things (IoT) security. *IEEE Communications Surveys & Tutorials*, 22(3):1646–1685.

Al Idrus, A. (2019). Almirall taps Iktos for AI drug discovery work. Fierce Biotech. https://www.fiercebiotech.com/biotech/almirall-taps-iktos-for-ai-drug-discovery-work#:~:text=Almirall%20is%20enlisting%20artificial%20intelligence.

Aldape-Pérez, M., Alarcón-Paredes, A., Yáñez-Márquez, C., López-Yáñez, I., and Camacho-Nieto, O. (2018). An associative memory approach to healthcare monitoring and decision making. *Sensors*, 18(8):2690.

Alonso, S. G., Marques, G., Barrachina, I., Garcia-Zapirain, B., Arambarri, J., Salvador, J. C., and de la Torre Díez, I. (2021). Telemedicine and e-health research solutions in literature for combatting covid-19: a systematic review. *Health and Technology*, 11(2):257–266.

Amato, F., Cozzolino, G., Mazzeo, A., and Romano, S. (2018). Intelligent medical record management: a diagnosis support system. *International Journal of High Performance Computing and Networking*, 12(4):391–399.

Anghel, S. and Drachenberg, R. (2019). Eprs— european parliamentary research service.

Bai, M., Storer, R. H., and Tonkay, G. L. (2017). A sample gradient-based algorithm for a multiple-OR and PACU surgery scheduling problem. *IISE Transactions*, 49(4):367–380.

Bao, H., Sun, X., Zhang, Y., Pang, B., Li, H., Zhou, L., Wu, F., Cao, D., Wang, J., Turic, B., and Wang, L. (2020). The artificial intelligence-assisted cytology diagnostic system in large-scale cervical cancer screening: A population-based cohort study of 0.7 million women. *Cancer Medicine*, 9(18):6896–6906.

Barbieri, C., Cattinelli, I., Neri, L., Mari, F., Ramos, R., Brancaccio, D., Canaud, B., and Stuard, S. (2019). Development of an artificial intelligence model to guide the management of blood pressure, fluid volume, and dialysis dose in end-stage kidney disease patients: Proof of concept and first clinical assessment. *Kidney Diseases*, 5(1):28–33.

Barrett, M., Boyne, J., Brandts, J., Brunner-La Rocca, H.-P., De Maesschalck, L., De Wit, K., Dixon, L., Eurlings, C., Fitzsimons, D., Golubnitschaja, O., Hageman, A., Heemskerk, F., Hintzen, A., Helms, T. M., Hill, L., Hoedemakers, T., Marx, N., McDonald, K., Mertens, M., Müller-Wieland, D., Palant, A., Piesk, J., Pomazanskyi, A., Ramaekers, J., Ruff, P., Schütt, K., Shekhawat, Y., Ski, C. F., Thompson, D. R., Tsirkin, A., Mierden, K. v. d., Watson, C., and Zippel-Schultz, B. (2019). Artificial intelligence supported patient self-care in chronic heart failure: A paradigm shift from reactive to predictive, preventive and personalised care. *EPMA Journal*, 10(4):445–464.

Baud, R., Rassinoux, A.-M., and Scherrer, J.-R. (1992). Natural language processing and semantical representation of medical texts. *Methods of Information in Medicine*, 31(02):117–125.

Beheshtifar, S. and Alimoahmmadi, A. (2015). A multiobjective optimization approach for location-allocation of clinics. *International Transactions in Operational Research*, 22(2):313–328.

Belkhamsa, M., Jarboui, B., and Masmoudi, M. (2018). Two metaheuristics for solving no-wait operating room surgery scheduling problem under various resource constraints. *Computers & Industrial Engineering*, 126:494–506.

Bertsimas, D. and Kallus, N. (2020). From predictive to prescriptive analytics. *Management Science*, 66(3):1025–1044.

Billings, J., Georghiou, T., Blunt, I., and Bardsley, M. (2013). Choosing a model to predict hospital admission: An observational study of new variants of predictive models for case finding. *BMJ Open*, 3(8):e003352.

Birnbaum, D. (2004). Application of data mining techniques to healthcare data. *Infection Control and Hospital Epidemiology*, 25(8):690–695.

Blansit, K., Retson, T., Masutani, E., Bahrami, N., and Hsiao, A. (2019). Deep learning–based prescription of cardiac MRI planes. *Radiology: Artificial Intelligence*, 1(6):e180069.

Blume, P. (2015). Eu adequacy decisions: the proposed new possibilities. *International Data Privacy Law*, 5(1):34.

Bowles, J. (2020). How canadian AI start-up bluedot spotted coronavirus before anyone else had a clue. Diginomica. https://diginomica.com/how-canadian-ai-start-bluedot-spotted-coronavirus-anyone-else-had-clue.

Burns, M. N., Begale, M., Duffecy, J., Gergle, D., Karr, C. J., Giangrande, E., and Mohr, D. C. (2011). Harnessing context sensing to develop a mobile intervention for depression. *Journal of Medical Internet Research*, 13(3):e55.

Carrell, D. S., Schoen, R. E., Leffler, D. A., Morris, M., Rose, S., Baer, A., Crockett, S. D., Gourevitch, R. A., Dean, K. M., and Mehrotra, A. (2017). Challenges in adapting existing clinical natural language processing systems to multiple, diverse health care settings. *Journal of the American Medical Informatics Association*, 24(5):986–991.

Chakkor, S., Baghouri, M., Cheker, Z., El Oualkadi, A., el Hangouche, J. A., and Laamech, J. (2021). Intelligent network for proactive detection of Covid-19 disease. In *2020 6th IEEE Congress on Information Science and Technology (CiSt)*, pages 472–478. IEEE.

Channa, A., Popescu, N. and Malik, N. Ur R. (2020). Managing covid-19 global pandemic with high-tech consumer wearables: A comprehensive review. In *2020 12th International Congress on Ultra Modern Telecommunications and Control Systems and Workshops (ICUMT)*, pages 222–228. IEEE.

Chatterjee, A., Gerdes, M. W., and Martinez, S. G. (2020). Statistical explorations and univariate timeseries analysis on COVID-19 datasets to understand the trend of disease spreading and death. *Sensors*, 20(11):3089.

Chen, E., Jiang, J., Su, R., Gao, M., Zhu, S., Zhou, J., and Huo, Y. (2020a). A new smart wristband equipped with an artificial intelligence algorithm to detect atrial fibrillation. *Heart Rhythm*, 17(5):847–853.

Chen, Y., Dong, Y., Zeng, Y., Yang, X., Shen, J., Zheng, L., Jiang, J., Pu, L., and Bao, Q. (2020b). Mapping of diseases from clinical medicine research: a visualization study. *Scientometrics*, 125(1):171–185.

Chukwu, E. and Garg, L. (2020). A systematic review of blockchain in healthcare: Frameworks, prototypes, and implementations. *IEEE Access*, 8:21196–21214.

Churpek, M. M., Yuen, T. C., Winslow, C., Meltzer, D. O., Kattan, M. W., and Edelson, D. P. (2016). Multicenter comparison of machine learning methods and conventional regression for predicting clinical deterioration on the wards. *Critical Care Medicine*, 44(2):368.

Connatty, S. (2019). Capgemini is selected by Bayer as a core strategic partner to transform its IT landscape. Intrado GlobeNewswire. https://www.globenewswire.com/news-release/2019/12/04/1956359/0/en/Capgemini-is-selected-by-Bayer-as-a-core-strategic-partner-to-transform-its-IT-landscape.html.

Coronato, A., Naeem, M., De Pietro, G., and Paragliola, G. (2020). Reinforcement learning for intelligent healthcare applications: A survey. *Artificial Intelligence in Medicine*, 109:101964.

Cowie, M. R., Bax, J., Bruining, N., Cleland, J. G., Koehler, F., Malik, M., Pinto, F., van der Velde, E., and Vardas, P. (2016). e-Health: A position statement of the European Society of Cardiology. *European Heart Journal*, 37(1):63.

Crocamo, C., Viviani, M., Bartoli, F., Carrà, G., and Pasi, G. (2020). Detecting binge drinking and alcohol-related risky behaviours from Twitter's users: An exploratory content-and topology-based analysis. *International Journal of Environmental Research and Public Health*, 17(5):1510.

Cui, Z. L., Kadziola, Z., Lipkovich, I., Faries, D. E., Sheffield, K. M., and Carter, G. C. (2021). Predicting optimal treatment regimens for patients with hr+/her2-breast cancer using machine learning based on electronic health records. *Journal of Comparative Effectiveness Research*, 10(9):777–795.

Dang, L. M., Piran, M., Han, D., Min, K., Moon, H., et al. (2019). A survey on Internet of things and cloud computing for healthcare. *Electronics*, 8(7):768.

Dascalu, D., Paraschiv, I. C., Nicula, B., Dascalu, M., Trausan-Matu, S., and Nuta, A. C. (2019). Intelligent platform for the analysis of drug leaflets using nlp techniques. In *2019 18th RoEduNet Conference: Networking in Education and Research (RoEduNet)*, pages 1–6. IEEE.

Datta, S., Li, Y., Ruppert, M. M., Ren, Y., Shickel, B., Ozrazgat-Baslanti, T., Rashidi, P., and Bihorac, A. (2021). Reinforcement learning in surgery. *Surgery*, 170(1):329–332.

Davenport, T. and Kalakota, R. (2019). The potential for artificial intelligence in healthcare. *Future Healthcare Journal*, 6(2):94.

De Magistris, G., Russo, S., Roma, P., Starczewski, J. T., and Napoli, C. (2022). An explainable fake news detector based on named entity recognition and stance classification applied to covid-19. *Information*, 13(3):137.

Demner-Fushman, D., Chapman, W. W., and McDonald, C. J. (2009). What can natural language processing do for clinical decision support? *Journal of Biomedical Informatics*, 42(5):760–772.

Din, I. U., Almogren, A., Guizani, M., and Zuair, M. (2019). A decade of Internet of Things: Analysis in the light of healthcare applications. *IEEE Access*, 7:89967–89979.

Doraiswamy, S., Abraham, A., Mamtani, R., Cheema, S., et al. (2020). Use of telehealth during the covid-19 pandemic: scoping review. *Journal of medical Internet research*, 22(12):e24087.

Downing, N. S., Cloninger, A., Venkatesh, A. K., Hsieh, A., Drye, E. E., Coifman, R. R., and Krumholz, H. M. (2017). Describing the performance of U.S. hospitals by applying big data analytics. *PLOS ONE*, 12(6):1–14.

Dua, S., Acharya, U. R., and Dua, P. (2014). *Machine learning in healthcare informatics*, volume 56. Springer.

Elaziz, M. A., Hosny, K. M., Salah, A., Darwish, M. M., Lu, S., and Sahlol, A. T. (2020). New machine learning method for image-based diagnosis of COVID-19. *PLOS ONE*, 15(6):e0235187.

Ellahham, S. and Ellahham, N. (2019). Use of artificial intelligence for improving patient flow and healthcare delivery. *J Comput Sci Syst Biol*, 12(303):2.

Emanuel, E. J. and Wachter, R. M. (2019). Artificial intelligence in health care: Will the value match the hype? *JAMA*, 321(23):2281–2282.

Enticott, J., Johnson, A., and Teede, H. (2021). Learning health systems using data to drive healthcare improvement and impact: A systematic review. *BMC health services research*, 21(1):1–16.

Eysenbach, G. (2001). What is e-health? *Journal of Medical Internet Research*, 3(2):e20.

Fayyad, U., Piatetsky-Shapiro, G., and Smyth, P. (1996). From data mining to knowledge discovery in databases. *AI Magazine*, 17(3):37–37.

Feldman, K., Davis, D., and Chawla, N. V. (2015). Scaling and contextualizing personalized healthcare: A case study of disease prediction algorithm integration. *Journal of Biomedical Informatics*, 57:377–385.

Ferreri, F., Bourla, A., Mouchabac, S., and Karila, L. (2018). e-addictology: an overview of new technologies for assessing and intervening in addictive behaviors. *Frontiers in psychiatry*, page 51.

Fikar, C. and Hirsch, P. (2017). Home health care routing and scheduling: A review. *Computers & Operations Research*, 77:86–95.

Fikar, C., Juan, A. A., Martinez, E., and Hirsch, P. (2016). A discrete-event driven metaheuristic for dynamic home service routing with synchronised trip sharing. *European Journal of Industrial Engineering*, 10(3):323–340.

Friedman, C. and Elhadad, N. (2014). Natural language processing in health care and biomedicine. In *Biomedical Informatics*, pages 255–284. Springer.

Friedman, C. and Hripcsak, G. (1999). Natural language processing and its future in medicine. *Acad Med*, 74(8):890–5.

Galetsi, P. and Katsaliaki, K. (2020). A review of the literature on big data analytics in healthcare. *Journal of the Operational Research Society*, 71(10):1511–1529.

Ganasegeran, K. and Abdulrahman, S. A. (2020). Artificial intelligence applications in tracking health behaviors during disease epidemics. In *Human Behaviour Analysis Using Intelligent Systems*, pages 141–155. Springer.

Ganguly, A. and Nandi, S. (2016). Using statistical forecasting to optimize staff scheduling in healthcare organizations. *Journal of Health Management*, 18(1):172–181.

Gao, X., Jin, Y., Dou, Q., and Heng, P.-A. (2020). Automatic gesture recognition in robot-assisted surgery with reinforcement learning and tree search. In *2020 IEEE International Conference on Robotics and Automation (ICRA)*, pages 8440–8446.

Garg, L., McClean, S. I., Barton, M., Meenan, B. J., and Fullerton, K. (2012). Intelligent patient management and resource planning for complex, heterogeneous, and stochastic healthcare systems. *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans*, 42(6):1332–1345.

Goldstein, B. A., Navar, A. M., and Carter, R. E. (2017). Moving beyond regression techniques in cardiovascular risk prediction: Applying machine learning to address analytic challenges. *European Heart Journal*, 38(23):1805–1814.

Goodarzian, F., Hosseini-Nasab, H., and Fakhrzad, M. (2020). A multi-objective sustainable medicine supply chain network design using a novel hybrid multi-objective metaheuristic algorithm. *International Journal of Engineering*, 33(10):1986–1995.

Gottesman, O., Johansson, F., Komorowski, M., Faisal, A., Sontag, D., Doshi-Velez, F., and Celi, L. A. (2019). Guidelines for reinforcement learning in healthcare. *Nature medicine*, 25(1):16–18.

Gundapu, S. and Mamidi, R. (2021). Transformer based automatic covid-19 fake news detection system. *arXiv preprint arXiv:2101.00180*.

Guo, L., Xiao, X., Wu, C., Zeng, X., Zhang, Y., Du, J., Bai, S., Xie, J., Zhang, Z., Li, Y., Wang, X., Cheung, O., Sharma, M., Liu, J., and Hu, B. (2020). Real-time automated diagnosis of precancerous lesions and early esophageal squamous cell carcinoma using a deep learning model (with videos). *Gastrointestinal Endoscopy*, 91(1):41 – 51.

Gupta, N. and Gupta, B. (2019). Big data interoperability in e-health systems. In *2019 9th International Conference on Cloud Computing, Data Science & Engineering (Confluence)*, pages 217–222. IEEE.

Haldorai, A., Ramu, A., and Murugan, S. (2019). Biomedical informatics and computation in urban e-health. In *Computing and Communication Systems in Urban Development*, pages 69–89. Springer.

Hand, D. J. (1998). Data mining: Statistics and more? *The American Statistician*, 52(2):112–118.

Harper, A., Mustafee, N., and Feeney, M. (2017). A hybrid approach using forecasting and discrete-event simulation for endoscopy services. In *2017 Winter Simulation Conference (WSC)*, pages 1583–1594. IEEE.

Hashimoto, D. A., Rosman, G., Rus, D., and Meireles, O. R. (2018). Artificial intelligence in surgery: Promises and perils. *Annals of Surgery*, 268(1):70.

HassanAbady, S. E. and Ganjali, R. (2021). Medical informatics applications in covid-19 crisis control: Protocol for systematic literature review. *Frontiers in Health Informatics*, 10(1):56.

He, J., Baxter, S. L., Xu, J., Xu, J., Zhou, X., and Zhang, K. (2019). The practical implementation of artificial intelligence technologies in medicine. *Nature Medicine*, 25(1):30–36.

Heifetz, A. (2022). *Artificial Intelligence in Drug Design*. Springer.

Hernán, M. A., Hsu, J., and Healy, B. (2019). A second chance to get causal inference right: A classification of data science tasks. *Chance*, 32(1):42–49.

Ho, A. F. W., To, B. Z. Y. S., Koh, J. M., and Cheong, K. H. (2019). Forecasting hospital emergency department patient volume using internet search data. *IEEE Access*, 7:93387–93395.

Hochreiter, S. and Schmidhuber, J. (1997). Long short-term memory. *Neural Computation*, 9(8):1735–1780.

Hofmarcher, M., Mayr, A., Rumetshofer, E., Ruch, P., Renz, P., Schimunek, J., Seidl, P., Vall, A., Widrich, M., Hochreiter, S., and Klambauer, G. (2020). Large-scale ligand-based virtual screening for SARS-Cov-2 inhibitors using a deep neural network. Technical report, SSRN.

Holzinger, A. (2016). Interactive machine learning for health informatics: When do we need the human-in-the-loop? *Brain Informatics*, 3(2):119–131.

Hoot, N. R. and Aronsky, D. (2008). Systematic review of emergency department crowding: Causes, effects, and solutions. *Annals of Emergency Medicine*, 52(2):126–136.

Hoot, N. R., Zhou, C., Jones, I., and Aronsky, D. (2007). Measuring and forecasting emergency department crowding in real time. *Annals of Emergency Medicine*, 49(6):747–755.

Hrinivich, W. T. and Lee, J. (2020). Artificial intelligence-based radiotherapy machine parameter optimization using reinforcement learning. *Medical physics*, 47(12):6140–6150.

Hu, C., Ju, R., Shen, Y., Zhou, P., and Li, Q. (2016). Clinical decision support for alzheimer's disease based on deep learning and brain network. In *2016 IEEE International Conference on Communications (ICC)*, pages 1–6. IEEE.

Huang, E. and Chang, C.-C. A. (2012). Patient-oriented interactive e-health tools on us hospital web sites. *Health marketing quarterly*, 29(4):329–345.

Huang, E., Chang, C.-c. A., and Khurana, P. (2012a). Users' preferred interactive e-health tools on hospital web sites. *International journal of pharmaceutical and healthcare marketing*.

Huang, T., Lan, L., Fang, X., An, P., Min, J., and Wang, F. (2015). Promises and challenges of big data computing in health sciences. *Big Data Research*, 2(1):2–11.

Huang, Z., Zhou, A., and Zhang, G. (2012b). Non-negative matrix factorization: A short survey on methods and applications. In *International Symposium on Intelligence Computation and Applications*, pages 331–340. Springer.

Hund, M., Sturm, W., Schreck, T., Ullrich, T., Keim, D., Majnaric, L., and Holzinger, A. (2015). Analysis of patient groups and immunization results based on subspace clustering. In *International Conference on Brain Informatics and Health*, pages 358–368. Springer.

Idrees, M. and Sohail, A. (2021). Bio-algorithms for the modeling and simulation of cancer cells and the immune response. *Bio-Algorithms and Med-Systems*, 17(1):55–63.

Impedovo, D., Pirlo, G., and Vessio, G. (2018). Dynamic handwriting analysis for supporting earlier parkinson's disease diagnosis. *Information*, 9(10):247.

Iram, A. and Gill, S. H. (2018). Similarity measuring for clustering patient's reports in telemedicine. In *International Conference on Intelligent Technologies and Applications*, pages 38–49. Springer.

Iroju, O. G. and Olaleke, J. O. (2015). A systematic review of natural language processing in healthcare. *International Journal of Information Technology and Computer Science*, 7(8):44–50.

Javaid, M., Haleem, A., Singh, R. P., and Suman, R. (2021). Dentistry 4.0 technologies applications for dentistry during COVID-19 pandemic. *Sustainable Operations and Computers*, 2:87–96.

Jiang, F., Jiang, Y., Zhi, H., Dong, Y., Li, H., Ma, S., Wang, Y., Dong, Q., Shen, H., and Wang, Y. (2017). Artificial intelligence in healthcare: Past, present and future. *Stroke and Vascular Neurology*, 2(4):230–243.

Jing, Y., Bian, Y., Hu, Z., Wang, L., and Xie, X.-Q. S. (2018). Deep learning for drug design: An artificial intelligence paradigm for drug discovery in the big data era. *The AAPS Journal*, 20(3):1–10.

Johnson, J. M. and Khoshgoftaar, T. M. (2019). Medicare fraud detection using neural networks. *Journal of Big Data*, 6(1):63.

Jothi, N., Rashid, N., and Husain, W. (2015). Data mining in healthcare–a review. *Procedia Computer Science*, 72:306–313.

Jumper, J., Tunyasuvunakool, K., Kohli, P., Hassabis, D., and the AlphaFold Team (2020). Computational predictions of protein structures associated with COVID-19. Deepmind website. https://deepmind.com/research/open-source/computational-predictions-of-protein-structures-associated-with-COVID-19.

Kam, H. J., Sung, J. O., and Park, R. W. (2010). Prediction of daily patient numbers for a regional emergency medical center using time series analysis. *Healthcare Informatics Research*, 16(3):158–165.

Karczewski, K. J. and Snyder, M. P. (2018). Integrative omics for health and disease. *Nature Reviews Genetics*, 19(5):299–310.

Kavakiotis, I., Tsave, O., Salifoglou, A., Maglaveras, N., Vlahavas, I., and Chouvarda, I. (2017). Machine learning and data mining methods in diabetes research. *Computational and structural biotechnology journal*, 15:104–116.

Kermany, D. S., Goldbaum, M., Cai, W., Valentim, C. C., Liang, H., Baxter, S. L., McKeown, A., Yang, G., Wu, X., Yan, F., Dong, J., Prasadha, M. K., Pei, J., Ting, M. Y., Zhu, J., Li, C., Hewett, S., Dong, J., Ziyar, I., Shi, A., Zhang, R., Zheng, L., Hou, R., Shi, W., Fu, X., Duan, Y., Huu, V. A., Wen, C., Zhang, E. D., Zhang, C. L., Li, O., Wang, X., Singer, M. A., Sun, X., Xu, J., Tafreshi, A., Lewis, M. A., Xia, H., and Zhang, K. (2018). Identifying medical diagnoses and treatable diseases by image-based deep learning. *Cell*, 172(5):1122–1131.

Khaldi, R., El Afia, A., and Chiheb, R. (2019). Forecasting of weekly patient visits to emergency department: real case study. *Procedia computer science*, 148:532–541.

Kim, D., You, S., So, S., Lee, J., Yook, S., Jang, D. P., Kim, I. Y., Park, E., Cho, K., Cha, W. C., Shin, D. W., Cho, B. H., and Park, H.-K. (2018). A data-driven artificial intelligence model for remote triage in the prehospital environment. *PLOS ONE*, 13(10):e0206006.

Kim, J., Lee, S., Hwang, E., Ryu, K. S., Jeong, H., Lee, J. W., Hwangbo, Y., Choi, K. S., Cha, H. S., et al. (2020). Limitations of deep learning attention mechanisms in clinical research: Empirical case study based on the korean diabetic disease setting. *Journal of Medical Internet Research*, 22(12):e18418.

Koh, H. C. and Tan, G. (2011). Data mining applications in healthcare. *Journal of Healthcare Information Management*, 19(2):65.

Kriegel, H.-P., Borgwardt, K. M., Kröger, P., Pryakhin, A., Schubert, M., and Zimek, A. (2007). Future trends in data mining. *Data Mining and Knowledge Discovery*, 15(1):87–97.

Kruppa, J., Liu, Y., Biau, G., Kohler, M., König, I. R., Malley, J. D., and Ziegler, A. (2014). Probability estimation with machine learning methods for dichotomous and multicategory outcome: Theory. *Biometrical Journal*, 56(4):534–563.

Kshirsagar, R., Hsu, L.-Y., Greenberg, C. H., McClelland, M., Mohan, A., Shende, W., Tilmans, N. P., Guo, M., Chheda, A., Trotter, M., Ray, S., and Alvarado, M.

(2020). Accurate and interpretable machine learning for transparent pricing of health insurance plans. *arXiv preprint arXiv:2009.10990*.

Kundu, S., Elhalawani, H., Gichoya, J. W., and Kahn Jr, C. E. (2020). Radiological Society of North America. How might AI and chest imaging help unravel COVID-19's mysteries? *Radiology: Artificial Intelligence*, 2(3):e200053.

Kwak, G., Ling, L., and Hui, P. (2021). Deep reinforcement learning approaches for global public health strategies for COVID-19 pandemic. *PLoS ONE*, 16(5):e0251550.

Lagos-Ortiz, K., Jácome-Murillo, E., Aguirre-Munizaga, M., and Medina-Moreira, J. (2020). Analysis of the services generated through mobile applications for an accurate diagnosis of epidemiological metrics related to covid-19. In *International Conference on Technologies and Innovation*, pages 151–165. Springer.

Lakshmanaprabu, S., Mohanty, S. N., Krishnamoorthy, S., Uthayakumar, J., Shankar, K., et al. (2019). Online clinical decision support system using optimal deep neural networks. *Applied Soft Computing*, 81:105487.

Lan Hing Ting, K., Voilmy, D., De Roll, Q., Iglesias, A., and Marfil, R. (2021). Fieldwork and field trials in hospitals: Co-designing a robotic solution to support data collection in geriatric assessment. *Applied Sciences*, 11(7):3046.

Larson, D. and Chang, V. (2016). A review and future direction of agile, business intelligence, analytics and data science. *International Journal of Information Management*, 36(5):700–710.

Lathrop, R. H. (1994). The protein threading problem with sequence amino acid interaction preferences is NP-complete. *Protein Engineering, Design and Selection*, 7(9):1059–1068.

Lepenioti, K., Bousdekis, A., Apostolou, D., and Mentzas, G. (2020). Prescriptive analytics: Literature review and research challenges. *International Journal of Information Management*, 50:57–70.

Li, K., Fang, Y., Li, W., Pan, C., Qin, P., Zhong, Y., Liu, X., Huang, M., Liao, Y., and Li, S. (2020). CT image visual quantitative evaluation and clinical classification of coronavirus disease (COVID-19). *European Radiology*, 30:4407–4416.

Liang, Y., Yin, S., Tang, Q., Zheng, Z., Elgendi, M., and Chen, Z. (2020). Deep learning algorithm classifies heartbeat events based on electrocardiogram signals. *Frontiers in Physiology*, 11:1255.

Liao, P., Greenewald, K., Klasnja, P., and Murphy, S. (2020). Personalized heartsteps: A reinforcement learning algorithm for optimizing physical activity. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 4(1):1–22.

Lisboa, P. J. and Taktak, A. F. (2006). The use of artificial neural networks in decision support in cancer: A systematic review. *Neural Networks*, 19(4):408–415.

Liu, Y. H. (2017). *Python Machine Learning By Example; Chapter 10: Discovering underlying topics in the newsgroups dataset with clustering and topic modeling*. Packt Publishing Ltd.

Liu, Z., Yao, C., Yu, H., and Wu, T. (2019). Deep reinforcement learning with its application for lung cancer detection in medical internet of things. *Future Generation Computer Systems*, 97:1–9.

Lostumbo, M., Saiz, M., Calvet, L., Juan, A. A., and Lopez, D. (2021). Combining simulation with reliability analysis in supply chain project management under uncertainty: a case study in healthcare. In *Proceedings of the 2021 winter simulation conference*. IEEE.

Lv, J., Deng, S., and Zhang, L. (2020). A review of artificial intelligence applications for antimicrobial resistance. *Biosafety and Health*, 3(1):22–31.

Ma, B., Meng, F., Yan, G., Yan, H., Chai, B., and Song, F. (2020). Diagnostic classification of cancers using extreme gradient boosting algorithm and multi-omics data. *Computers in Biology and Medicine*, 121:103761.

Makhdoom, I., Abolhasan, M., Lipman, J., Liu, R. P., and Ni, W. (2018). Anatomy of threats to the internet of things. *IEEE Communications Surveys & Tutorials*, 21(2):1636–1675.

Mamdouh, M., Awad, A. I., Hamed, H. F., and Khalaf, A. A. (2020). Outlook on security and privacy in ioht: Key challenges and future vision. In *The International Conference on Artificial Intelligence and Computer Vision*, pages 721–730. Springer.

Martin, C. M., Vogel, C., Grady, D., Zarabzadeh, A., Hederman, L., Kellett, J., Smith, K., and O'Shea, B. (2012). Implementation of complex adaptive chronic care: The Patient Journey Record system (PaJR). *Journal of Evaluation in Clinical Practice*, 18(6):1226–1234.

Martins, L. C., Castaneda, J., Juan, A. A., Barrios, B., Calvet, L., Tondar, A., and Sanchez, J. (2021). Supporting efficient assignment of medical resources in cancer treatments with simulation-optimization. In *Proceedings of the 2021 winter simulation conference*. IEEE.

Martuscelli, C. (2018). AstraZeneca announces China partnerships. MarketWatch. https://www.marketwatch.com/story/astrazeneca-announces-china-partnerships-2018-02-02.

Matheny, M. E., Whicher, D., and Israni, S. T. (2020). Artificial intelligence in health care: A report from the National Academy of Medicine. *JAMA*, 323(6):509–510.

McIntosh, A. M., Stewart, R., John, A., Smith, D. J., Davis, K., Sudlow, C., Corvin, A., Nicodemus, K. K., Kingdon, D., Hassan, L., Hotopf, M., Lawrie, S. M., Russ, T. C., Geddes, J. R., Wolpert, M., Wölbert, E., and Porteous, D. J. (2016). Data science for mental health: A UK perspective on a global challenge. *The Lancet Psychiatry*, 3(10):993–998.

Mesquita, C. T., Oliveira, A., Seixas, F. L., and Paes, A. (2020). Infodemia, fake news and medicine: Science and the quest for truth.

Milliard, M. (2020). Health passports and distancing tools among COVID-19 tech climbing Gartner Hype Cycle. Healthcare IT News. https://www.healthcareitnews.com/news/health-passports-distancing-tools-among-covid-19-tech-climbing-gartner-hype-cycle.

Miori, V. and Russo, D. (2012). Anticipating health hazards through an ontology-based, iot domotic environment. In *2012 Sixth International Conference on Innovative Mobile and Internet Services in Ubiquitous Computing*, pages 745–750. IEEE.

Mishra, N. and Silakari, S. (2012). Predictive analytics: A survey, trends, applications, oppurtunities & challenges. *International Journal of Computer Science and Information Technologies*, 3(3):4434–4438.

Monaghesh, E. and Hajizadeh, A. (2020). The role of telehealth during covid-19 outbreak: a systematic review based on current evidence. *BMC public health*, 20(1):1–9.

Mookdarsanit, P. and Mookdarsanit, L. (2021). The covid-19 fake news detection in thai social texts. *Bulletin of Electrical Engineering and Informatics*, 10(2):988–998.

Moreira, M. W., Rodrigues, J. J., Carvalho, F. H., Chilamkurti, N., Al-Muhtadi, J., and Denisov, V. (2019a). Biomedical data analytics in mobile-health environments for high-risk pregnancy outcome prediction. *Journal of Ambient Intelligence and Humanized Computing*, 10(10):4121–4134.

Moreira, M. W., Rodrigues, J. J., Kumar, N., Al-Muhtadi, J., and Korotaev, V. (2018). Nature-inspired algorithm for training multilayer perceptron networks in e-health environments for high-risk pregnancy care. *Journal of Medical Systems*, 42(3):51.

Moreira, M. W., Rodrigues, J. J., Kumar, N., Saleem, K., and Illin, I. V. (2019b). Postpartum depression prediction through pregnancy data analysis for emotion-aware smart systems. *Information Fusion*, 47:23 – 31.

Moskowitz, D. S. and Young, S. N. (2006). Ecological momentary assessment: What it is and why it is a method of the future in clinical psychopharmacology. *Journal of Psychiatry and Neuroscience*, 31(1):13.

Murray, E., Hekler, E. B., Andersson, G., Collins, L. M., Doherty, A., Hollis, C., Rivera, D. E., West, R., and Wyatt, J. C. (2016). Evaluating digital health interventions: Key questions and approaches. *American Journal of Preventive Medicine*, 51(5):843–851.

Mutlag, A. A., Abd Ghani, M. K., Arunkumar, N. A., Mohammed, M. A., and Mohd, O. (2019). Enabling technologies for fog computing in healthcare IoT systems. *Future Generation Computer Systems*, 90:62–78.

Naeem, S. B., Bhatti, R., and Khan, A. (2021). An exploration of how fake news is taking over social media and putting public health at risk. *Health Information & Libraries Journal*, 38(2):143–149.

Nandankar, P., Thaker, R., Mughal, S. N., Saidireddy, M., Linda, A., Kostka, J., and Nag, M. A. (2021). An IoT based healthcare data analytics using fog and cloud computing. *Turkish Journal of Physiotherapy and Rehabilitation*, 32:3.

Naslund, J. A., Gonsalves, P. P., Gruebner, O., Pendse, S. R., Smith, S. L., Sharma, A., and Raviola, G. (2019). Digital innovations for global mental health: Opportunities for data science, task sharing, and early intervention. *Current Treatment Options in Psychiatry*, 6(4):337–351.

Neubert, S., Geißler, A., Roddelkopf, T., Stoll, R., Sandmann, K.-H., Neumann, J., and Thurow, K. (2019). Multi-sensor-fusion approach for a data-science-oriented preventive health management system: Concept and development of a decentralized data collection approach for heterogeneous data sources. *International Journal of Telemedicine and Applications*, 2019(3):1–18.

Obermeyer, Z. and Emanuel, E. J. (2016). Predicting the future: big data, machine learning, and clinical medicine. *The New England Journal of Medicine*, 375(13):1216.

OECD (2020). Background paper for the G20 AI dialogue, digital economy task force. Technical report.

Oh, H., Rizo, C., Enkin, M., and Jadad, A. (2005). What is eHealth (3): A systematic review of published definitions. *Journal of Medical Internet Research*, 7(1):e1.

Oke, S. (2008). A literature review on artificial intelligence. *International Journal of Information and Management Sciences*, 19(4):535–570.

Oliver, E., Vallés-Pérez, I., Baños, R.-M., Cebolla, A., Botella, C., and Soria-Olivas, E. (2018). Visual data mining with self-organizing maps for ”self-monitoring” data analysis. *Sociological Methods & Research*, 47(3):492–506.

O'Neill, J. (2014). *Antimicrobial resistance: Tackling a crisis for the health and wealth of nations*. Review on Antimicrobial Resistance.

Ordu, M., Demir, E., Tofallis, C., and Gunal, M. M. (2021). A novel healthcare resource allocation decision support tool: A forecasting-simulation-optimization approach. *Journal of the operational research society*, 72(3):485–500.

Organization, W. H. et al. (2021). The protection of personal data in health information systems-principles and processes for public health. Technical report, World Health Organization. Regional Office for Europe.

Oztekin, A. (2018). Creating a marketing strategy in healthcare industry: A holistic data analytic approach. *Annals of Operations Research*, 270(1):361–382.

Parfenenko, Y., Verbytska, A., Bychko, D., and Shendryk, V. (2020). Application for medical misinformation detection in online forums. In *2020 International Conference on e-Health and Bioengineering (EHB)*, pages 1–4. IEEE.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., and Duchesnay, E. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12:2825–2830.

Pereira, C. R., Pereira, D. R., Weber, S. A., Hook, C., De Albuquerque, V. H. C., and Papa, J. P. (2019). A survey on computer-assisted parkinson's disease diagnosis. *Artificial intelligence in medicine*, 95:48–63.

Pfeiffer, D. U. and Stevens, K. B. (2015). Spatial and temporal epidemiological analysis in the big data era. *Preventive Veterinary Medicine*, 122(1-2):213–220.

Poirion, O. B., Jing, Z., Chaudhary, K., Huang, S., and Garmire, L. X. (2021). Deepprog: an ensemble of deep-learning and machine-learning models for prognosis prediction using multi-omics data. *Genome medicine*, 13(1):1–15.

Popowich, F. (2005). Using text mining and natural language processing for health care claims processing. *ACM SIGKDD Explorations Newsletter*, 7(1):59–66.

Pore, A., Corsi, D., Marchesini, E., Dall'Alba, D., Casals, A., Farinelli, A., and Fiorini, P. (2021). Safe reinforcement learning using formal verification for tissue retraction in autonomous robotic-assisted surgery. In *2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 4025–4031.

Puliafito, C., Mingozzi, E., Longo, F., Puliafito, A., and Rana, O. (2019). Fog computing for the internet of things: A survey. *ACM Transactions on Internet Technology (TOIT)*, 19(2):1–41.

Pulido, C. M., Ruiz-Eugenio, L., Redondo-Sama, G., and Villarejo-Carballido, B. (2020). A new application of social impact in social media for overcoming fake news in health. *International journal of environmental research and public health*, 17(7):2430.

Quarteroni, A., Lassila, T., Rossi, S., and Ruiz-Baier, R. (2017). Integrated heart-coupling multiscale and multiphysics models for the simulation of the cardiac function. *Computer Methods in Applied Mechanics and Engineering*, 314:345–407.

Raghupathi, W. and Raghupathi, V. (2014). Big data analytics in healthcare: Promise and potential. *Health Information Science and Systems*, 2(1):3.

Rais, A. and Viana, A. (2011). Operations research in healthcare: A survey. *International Transactions in Operational Research*, 18(1):1–31.

Rao, A. and Venkatesh, S. (2021). Identification of aphasia using natural language processing. *Journal of University of Shanghai for Science and Technology*, 23:1737–1747.

Ray, E. L. and Reich, N. G. (2018). Prediction of infectious disease epidemics via weighted density ensembles. *PLOS Computational Biology*, 14(2):e1005910.

Razzaque, A. and Hamdan, A. (2020). Artificial intelligence based multinational corporate model for EHR interoperability on an e-health platform. In *Artificial Intelligence for Sustainable Development: Theory, Practice and Future Applications*, pages 71–81. Springer.

Reddy, S., Fox, J., and Purohit, M. P. (2019). Artificial intelligence-enabled healthcare delivery. *Journal of the Royal Society of Medicine*, 112(1):22–28.

Reinig, M. (2021). Boehringer & google accelerate drug design with quantum computing. https://www.boehringer-ingelheim.com/press-release/partnering-google-quantum-computing. 03-02-2021.

Richardson, P., Griffin, I., Tucker, C., Smith, D., Oechsle, O., Phelan, A., and Stebbing, J. (2020). Baricitinib as potential treatment for 2019-nCoV acute respiratory disease. *Lancet*, 395(10223):e30.

Riise, A., Mannino, C., and Burke, E. K. (2016). Modelling and solving generalised operational surgery scheduling problems. *Computers & Operations Research*, 66:1–11.

Rong, G., Mendez, A., Assi, E. B., Zhao, B., and Sawan, M. (2020). Artificial intelligence in healthcare: Review and prediction case studies. *Engineering*, 6(3):291–301.

Roy, S., Ashaduzzaman, M., Hassan, M., and Chowdhury, A. R. (2018). Blockchain for IoT security and management: Current prospects, challenges and future directions. In *2018 5th International Conference on Networking, Systems and Security (NSysS)*, pages 1–9. IEEE.

Rubrichi, S., Battistotti, A., and Quaglini, S. (2014). Patients' involvement in e-health services quality assessment: a system for the automatic interpretation of sms-based patients' feedback. *Journal of biomedical informatics*, 51:41–48.

Ruiz-Zafra, Á., Benghazi, K., Noguera, M., and Garrido, J. L. (2013). Zappa: An open mobile platform to build cloud-based m-health systems. In *Ambient intelligence-software and applications*, pages 87–94. Springer.

Sahoo, S. S., Jayapandian, C., Garg, G., Kaffashi, F., Chung, S., Bozorgi, A., Chen, C.-H., Loparo, K., Lhatoo, S. D., and Zhang, G.-Q. (2014). Heart beats in the cloud: Distributed analysis of electrophysiological 'Big Data' using cloud computing for epilepsy clinical research. *Journal of the American Medical Informatics Association*, 21(2):263–271.

Saranga, H. and Phani, B. (2009). Determinants of operational efficiencies in the Indian pharmaceutical industry. *International Transactions in Operational Research*, 16(1):109–130.

Senders, J. T., Staples, P. C., Karhade, A. V., Zaki, M. M., Gormley, W. B., Broekman, M. L., Smith, T. R., and Arnaout, O. (2018). Machine learning and neurosurgical outcome prediction: A systematic review. *World Neurosurgery*, 109:476–486.

Serra, D., Romero, P., Lozano, M., García-Fernández, I., Liberos, A., Rodrigo, M., Berruezo, A., Bueno-Orovio, A., and Sebastian, R. (2021). Arrhythmic3D: A fast automata-based tool to simulate and assess arrhythmia risk in 3d ventricular models. In *2021 Computing in Cardiology (CinC)*, volume 48, pages 1–4.

Sethia, D., Gupta, D., Saran, H., Agrawal, R., and Gaur, A. (2016). Mutual authentication protocol for secure nfc based mobile healthcard. *IADIS International Journal on Computer Science & Information Systems*, 11(2).

Shah, J. L., Bhat, H. F., and Khan, A. I. (2021). Integration of Cloud and IoT for smart e-healthcare. In *Healthcare Paradigms in the Internet of Things Ecosystem*, pages 101–136. Elsevier.

Shahid, N., Rappon, T., and Berta, W. (2019). Applications of artificial neural networks in health care organizational decision-making: A scoping review. *PLOS ONE*, 14(2):e0212356.

Shailaja, K., Seetharamulu, B., and Jabbar, M. (2018). Machine learning in healthcare: A review. In *2018 Second international conference on electronics, communication and aerospace technology (ICECA)*, pages 910–914. IEEE.

Shamir, R. R., Dolber, T., Noecker, A. M., Walter, B. L., and McIntyre, C. C. (2015). Machine learning approach to optimizing combined stimulation and medication therapies for parkinson's disease. *Brain stimulation*, 8(6):1025–1032.

Shatte, A. B., Hutchinson, D. M., and Teague, S. J. (2019). Machine learning in mental health: A scoping review of methods and applications. *Psychological Medicine*, 49(9):1426–1448.

Silva, E. A. T., Gomez, I. F. L., Arango, J. F. F., Smith, J. W., Ocampo, S. U., and Hidalgo, J. E. (2018). Evaluation of satisfaction and usability of a clinical decision support system (cdss) targeted for early obstetric risk assessment and patient follow-up. *E-HEALTH 2018 ICT, Society And Human Beings 2018*, page 3.

Singal, A. G., Mukherjee, A., Elmunzer, B. J., Higgins, P. D., Lok, A. S., Zhu, J., Marrero, J. A., and Waljee, A. K. (2013). Machine learning algorithms outperform conventional regression models in predicting development of hepatocellular carcinoma. *The American Journal of Gastroenterology*, 108(11):1723.

Snasel, V., Gajdos, P., Abdulla, H. M. D., and Polovincak, M. (2007). Using matrix decompositions in formal concept analysis. *ISIM*, 252.

Spanakis, G., Weiss, G., Boh, B., Lemmens, L., and Roefs, A. (2017). Machine learning techniques in eating behavior e-coaching. *Personal and Ubiquitous Computing*, 21(4):645–659.

Su, Y.-H., Huang, K., and Hannaford, B. (2021). Multicamera 3d viewpoint adjustment for robotic surgery via deep reinforcement learning. *Journal of Medical Robotics Research*, 6(01n02):2140003.

Sun, P. (2020). Security and privacy protection in cloud computing: Discussions and challenges. *Journal of Network and Computer Applications*, 160:102642.

Susanto, H. (2017). Electronic health system: sensors emerging and intelligent technology approach. In *Smart Sensors Networks*, pages 189–203. Elsevier.

Syed, L., Jabeen, S., Manimala, S., and Elsayed, H. A. (2019). Data science algorithms and techniques for smart healthcare using IoT and big data analytics. In *Smart Techniques for a Smarter Planet*, pages 211–241. Springer.

Tamilselvi, R. and Kalaiselvi, S. (2013). An overview of data mining techniques and applications. *International Journal of Science and Research (IJSR), India Online ISSN*, 2(2):2319–7064.

Taylor, R. A., Pare, J. R., Venkatesh, A. K., Mowafi, H., Melnick, E. R., Fleischman, W., and Hall, M. K. (2016). Prediction of in-hospital mortality in emergency department patients with sepsis: A local big data–driven, machine learning approach. *Academic Emergency Medicine*, 23(3):269–278.

Tejedor, M., Woldaregay, A. Z., and Godtliebsen, F. (2020). Reinforcement learning application in diabetes blood glucose control: A systematic review. *Artificial Intelligence in Medicine*, 104:101836.

Terry, M. (2020). IBM and Pfizer believe machine learning can predict Alzheimer's risk. BioSpace. https://www.biospace.com/article/ibm-and-pfizer-believe-machine-learning-can-predict-alzheimer-s-risk/.

Tomar, D. and Agarwal, S. (2013). A survey on data mining approaches for healthcare. *International Journal of Bio-Science and Bio-Technology*, 5(5):241–266.

Triantafyllidis, A., Velardo, C., Chantler, T., Shah, S. A., Paton, C., Khorshidi, R., Tarassenko, L., Rahimi, K., and SUPPORT-HF Investigators (2015). A personalised mobile-based home monitoring system for heart failure: The SUPPORT-HF Study. *International Journal of Medical Informatics*, 84(10):743–753.

Van Calster, B., Wynants, L., Timmerman, D., Steyerberg, E. W., and Collins, G. S. (2019). Predictive analytics in health care: How can we know it works? *Journal of the American Medical Informatics Association*, 26(12):1651–1654.

Van Der Aalst, W. (2016). Data science in action. In *Process Mining*, pages 3–23. Springer.

Vanbrabant, L., Braekers, K., Ramaekers, K., and Van Nieuwenhuyse, I. (2019). Simulation of emergency department operations: A comprehensive review of kpis and operational improvements. *Computers & Industrial Engineering*, 131:356–381.

Vedula, S. S., Ishii, M., and Hager, G. D. (2017). Objective assessment of surgical technical skill and competency in the operating room. *Annual Review of Biomedical Engineering*, 19:301–325.

Virkar, S., Kadam, A., Mallick, S., Tilekar, S., and Raut, N. (2021). An e-health patient management system. In *12th International Conference on Advances in Computing, Control, and Telecommunication Technologies, ACT 2021*, volume 2021-August, pages 41–46.

Waller, M. A. and Fawcett, S. E. (2013). Data science, predictive analytics, and big data: A revolution that will transform supply chain design and management. *Journal of Business Logistics*, 34(2):77–84.

Wang, C., Zhu, X., Hong, J. C., and Zheng, D. (2019). Artificial intelligence in radiotherapy treatment planning: Present and future. *Technology in Cancer Research & Treatment*, 18:1533033819873922.

Wang, L. and Alexander, C. A. (2020). Big data analytics in medical engineering and healthcare: Methods, advances and challenges. *Journal of Medical Engineering & Technology*, 44(6):267–283.

Wang, L., Shi, H., and Gan, L. (2018a). Healthcare facility location-allocation optimization for China's developing cities utilizing a multi-objective decision support approach. *Sustainability*, 10(12):4580.

Wang, Y. and Hajli, N. (2017). Exploring the path to big data analytics success in healthcare. *Journal of Business Research*, 70:287–299.

Wang, Y., Kung, L., and Byrd, T. A. (2018b). Big data analytics: Understanding its capabilities and potential benefits for healthcare organizations. *Technological Forecasting and Social Change*, 126:3–13.

Warmerdam, L., Smit, F., van Straten, A., Riper, H., and Cuijpers, P. (2010). Cost-utility and cost-effectiveness of Internet-based treatment for adults with depressive symptoms: Randomized trial. *Journal of Medical Internet Research*, 12(5):e53.

Watts, J., Khojandi, A., Vasudevan, R., and Ramdhani, R. (2020). Optimizing individualized treatment planning for Parkinson's disease using deep reinforcement learning. In *2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, pages 5406–5409. IEEE.

Weltz, J., Volfovsky, A., and Laber, E. B. (2022). Reinforcement learning methods in public health. *Clinical Therapeutics*.

Weng, S. F., Reps, J., Kai, J., Garibaldi, J. M., and Qureshi, N. (2017). Can machine-learning improve cardiovascular risk prediction using routine clinical data? *PLOS ONE*, 12(4):e0174944.

Widmer, R. J., Collins, N. M., Collins, C. S., West, C. P., Lerman, L. O., and Lerman, A. (2015). Digital health interventions for the prevention of cardiovascular disease: A systematic review and meta-analysis. In *Mayo Clinic Proceedings*, volume 90, pages 469–480. Elsevier.

Wijnhoven, F. (2021). Challenges of adopting human-centered intelligent systems: An organizational learning approach. In *Human Centred Intelligent Systems*, pages 13–25. Springer.

Yang, X., Wang, Y., Byrne, R., Schneider, G., and Yang, S. (2019). Concepts of artificial intelligence for computer-assisted drug discovery. *Chemical Reviews*, 119(18):10520–10594.

Yu, C., Liu, J., Nemati, S., and Yin, G. (2021). Reinforcement learning in healthcare: A survey. *ACM Computing Surveys (CSUR)*, 55(1):1–36.

Yu, K.-H., Beam, A. L., and Kohane, I. S. (2018). Artificial intelligence in healthcare. *Nature Biomedical Engineering*, 2(10):719–731.

Zeevi, D., Korem, T., Zmora, N., Israeli, D., Rothschild, D., Weinberger, A., Ben-Yacov, O., Lador, D., Avnit-Sagi, T., Lotan-Pompan, M., Suez, J., Mahdi, J. A., Matot, E., Malka, G., Kosower, N., Rein, M., Zilberman-Schapira, G., Dohnalová, L., Pevsner-Fischer, M., Bikovsky, R., Halpern, Z., Elinav, E., and Segal, E. (2015). Personalized nutrition by prediction of glycemic responses. *Cell*, 163(5):1079–1094.

Zhang, W. and He, X. (2017). An anomaly detection method for medicare fraud detection. In *2017 IEEE International Conference on Big Knowledge (ICBK)*, pages 309–314.

Zhong, L., Luo, S., Wu, L., Xu, L., Yang, J., and Tang, G. (2014). A two-stage approach for surgery scheduling. *Journal of Combinatorial Optimization*, 27(3):545–556.

Zhu, N., Diethe, T., Camplani, M., Tao, L., Burrows, A., Twomey, N., Kaleshi, D., Mirmehdi, M., Flach, P., and Craddock, I. (2015). Bridging e-health and the internet of things: The sphere project. *IEEE Intelligent Systems*, 30(4):39–46.

Zuest, P. (2019). Novartis and Microsoft announce collaboration to transform medicine with artificial intelligence. Novartis. https://www.novartis.com/news/media-releases/novartis-and-microsoft-announce-collaboration-transform-medicine-artificial-intelligence.