Review

Impact of Big Data Analysis on Health

Albérico Travassos Rosário [1,*] and Joana Carmo Dias [2]

[1] GOVCOPP, Universidade Europeia; alberico@ua.pt
[2] COMEGI, Universidade Europeia; joana.carmo.dias@universidadeeuropeia.pt
* Correspondence: Albérico Travassos Rosário, alberico@ua.pt

Abstract: Big data analytics tools are the use of advanced analytic techniques targeting large and diverse volumes of data that include structured, semi-structured, and unstructured data from different sources and in different sizes from terabytes to zetabytes. The health sector is faced with the need to generate and manage large data sets from various health systems, such as electronic health records and clinical decision support systems. This data can be used by providers, clinicians, and policymakers to plan and implement interventions, detect disease more quickly, predict outcomes, and personalize care delivery. However, little attention is paid to the connection between big data analytics tools and the health sector. Thus, a systematic review of the bibliometric literature (LRSB) was developed to study how the adoption of big data analytics tools and infrastructures will revolutionize the healthcare industry. The review integrated 77 scientific and/or academic documents indexed in SCOPUS presenting up-to-date knowledge on current insights on how big data analytics technologies influence the healthcare sector and the different big data analytical tools used. The LRSB provides findings related to the impact of Big Data analytics on the health sector by introducing opportunities and technologies that provide practical solutions to various challenges.

Keywords: big data analytics, healthcare, data technologies, decision making, information management, EHR.)

1. Introduction

Big data analytics tools are used to process, manage, and store complex datasets that exceed the capacity of traditional data processing systems. As the global population embraces technologies, companies such as Google, IBM, and Facebook have presented technological infrastructures that extract valuable data from large data volumes collected from their users [1]. This data can be used to revolutionize the health sector by promoting evidence in developing and implementing treatment delivery models and decision-making regarding critical changes in the health industry [2]. By applying machine learning (ML) and artificial intelligence (AI) algorithms, big data analytics supports early disease detection, personalization of interventions, lower medical cost, and higher-quality care that lead to improved outcomes [3]. In addition, big data can be successfully implemented to manage chronic diseases [4] and reduce health disparities [5]. Thus, investing in the right big data analytics tools and infrastructures can improve clinicians, providers, and policymakers capability to plan and implement patient-centered and evidence-based interventions characterized by high-quality care, low costs, and higher outcomes.

With this study, we intend to understand how big data analytics technologies influence the healthcare sector and to identify the different big data analytical tools used. A systematic review of the bibliometric literature (LRSB) was developed following the four-step procedure proposed by Mariano et al. [6] to define the protocol, collect references, evaluate data, and interpret the findings. This LRSB provides findings related to the impact of Big Data analytics on the health sector by introducing opportunities and technologies that provide practical solutions to various challenges. Thus, this research shows several practical implications that big data analytics has in healthcare related to the opportu-

 doi:10.20944/preprints202203.0407.v1

nities created through the use of modern technologies for all stakeholders, including patients and care providers. Also, this research has made a theoretical contribution by highlighting big data analytics as practical tools to support and revolutionize the health sector.

2. Theoretical Background

In recent years, various institutions have increasingly generated huge volumes of structured, semi-structured, and unstructured data, referred to as big data. A wide range of devices, applications, and research activities generate heterogeneous data every day that needs to be stored, managed, or processed [7]. As the global population grows and health patterns change rapidly, healthcare providers and clinicians are expected to develop and implement treatment models that evolve based on these changes [8]. Some decisions made throughout these processes require big data to ensure their practicality and efficiency in addressing global health needs and demands [9]. Thus, like other sectors of the global economies, the healthcare sector is confronted by the need to generate and manage big data from various health systems such as electronic health records, the genome database, and clinical decisions support systems [10]. This data can be used by providers, clinicians, and policymakers to plan and implement interventions, detect diseases faster, predict outcomes, support therapeutic decisions, and personalize care delivery. Thus, embracing big data analytics tools and infrastructures will revolutionize the health care sector by supporting patient-centered and evidence-based care.

2.1. Defining Big Data Analytics

Definitions of big data are uncertain and confusing due to varying understanding of the terms among different people. For instance, Gandomi and Haider [11] indicated that some executives define big data from what it is while others focus on what it does. However, the common understanding of the term 'Big Data' encompasses collecting, analyzing, processing, and visualizing large data sets [12]. Some scholars have defined big data from the data quality dimensions (DQDs), introducing the 3 Vs (volume, velocity, variety) approach. Volume refers to the data magnitudes, such as petabytes or terabytes used to report data sizes. Variety encompasses a dataset's structural heterogeneity, for example, data can be structured, semi-structured, or unstructured. Velocity refers to the rate of generating data and the speed of analyzing and using it in decision-making. Gandomi and Haider [11] identified one definition of big data based on 3Vs, which states, "Big data is high-volume, high-velocity and high-variety information assets that demand cost-effective, innovative forms of information processing for enhanced insight and decision making" (p.138). DQDs in big data definitions emphasize the data characteristics needed to classify the large amounts of data gathered in big data as "quality" data [13]. While this dimension provides critical insights, it is insufficient in defining big data due to disagreements on whether the 3Vs are exclusively Big Data characteristics or data quality characteristics.

Alternative definitions take multiple approaches. For instance, Emmanuel and Stanier [14] identified three approaches: a comparative approach to definition, attribute-based definitions, and environmental definitions. The different definitions are based on themes identified in Big Data literature.

2.1.1. A comparative approach to the definition

The comparative approach compares big data characteristics and relational or traditional data characteristics. According to Emmanuel and Stanier [14], data management systems used at the beginning of the Big Data era were founded on the relational data model, whose characteristics were used to understand the new paradigm through comparison. This notion means that the scholars introducing Big Data based their definitions on the similarities and differences between the new big data technologies and traditional systems [15]. For instance, McKinsey Institute defined Big Data as "datasets that cannot be processed, stored and analyzed by traditional data management technologies" [14]

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 31 March 2022 doi:10.20944/preprints202203.0407.v1

(p.2). Instead of providing a formal definition of Big Data, this approach focuses on technological progress by comparing current tools and devices with the previously used ones [16]. It indicates that Big Data are products of progressive development in science and technologies and possibly indicates that they may change with time given the rapid technological advancements [17]. The relational aspect indicates that Big Data technologies can handle larger data volumes than traditional data management systems since they have higher processing capabilities and functionality.

3.1.2. Attribute-based definitions

This definition focuses on the characteristics of data, including the 3Vs Volume, Velocity, and Variety. Volume has been recognized as the most visible attribute of big data since it refers to the amount of data generated from various sources [14]. It refers to the magnitude of data gathered and analyzed in Big Data, which can also be perceived as the scale and size of data. However, factors such as the type of data can influence the definition of volume. For instance, Gandomi and Haider [11] explain that what is considered as data today may fail to meet the threshold in the future due to changes in storage capacities and amounts of data captured. Velocity as an attribute of data refers to the speed at which data is captured, analyzed, and stored. The development of digital technologies such as sensors and smartphones has influenced the rate at which data is generated and processed globally, leading to the need for evidence-based planning and real-time analytics [15]. Thus, Emmanuel and Stanier [14] indicated that velocity is defined as "real time or near real-time and to the streaming of data" (p.2). The importance of velocity as an attribute of Big Data is reflected in the increased demand for fast and timely data collection and analysis to ensure its effective commercialization. Variety refers to "structural heterogeneity in a dataset," classified as structured, semi-structured, and unstructured [11] (p.138). This attribute indicates that data is generated, processed, and stored in various formats, including texts, images, videos, or audio. The 3Vs provide data characteristics that differentiate Big Data from traditional data.

In recent years, scholars have extended 3Vs to include more characteristics such as value, veracity, and variability. IBM coined the fourth characteristic, veracity, to indicate that some data sources are unreliable yet important [11]. For example, some consumers can publish feedback and opinions about a product or brand on social media, which can be uncertain yet important for analyzing consumers' perceptions and engagement. Thus, Big Data tools and technologies are designed to handle imprecise and uncertain data [18]. Variability indicates that Big Data has varying flow rates, including peaks and troughs. Data analytic tools should have the capacity to analyze and store information regardless of changing variability. Oracle developed the value characteristic of big data to refer to its "low-value density", indicating that the value of the data generated is low compared to its volume [11] (p.140). Thus, the large data volumes must be analyzed to separate valuables that can be acted upon from valueless data. Most data-based decisions depend on the high-value data acquired after analysis.

3.1.3. Environmental definitions

This Big Data definition approach indicates that accurate definitions expand beyond data attributes, including processes, architectures, and applications. According to the National Institute of Standards and Technology (NIST), defining datasets in Big Data requires "a scalable architecture for efficient storage, manipulation, and analysis" [14] (p.3). this approach identifies the 3Vs as critical data features, which vary based on data context and process models. Thus, using them to define Big Data is insufficient [19]. Therefore, the proposed approach integrates these attributes, applications, architectures, and processing to create an overview of Big Data rather than providing a formal definition [18]. By accommodating the operations aspect of Big Data, this approach provides a more comprehensive approach that accounts for factors such as intended outcomes and cost-effectiveness, which are primary features of Big Data in commercial settings.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 31 March 2022 doi:10.20944/preprints202203.0407.v1

3.2. Big Data in Health Care

Computer programming has created a significant difference between traditional and big-data health analyses. In the traditional healthcare systems, health facilities depend on other industries such as IT for data analysis. However, the current generation of large data volumes requires the facilities and the health sector to embrace functional operating systems to ensure timely analysis of patient data for effective treatments [20]. Thus, healthcare providers, professionals, and policymakers recognize the growing significance of big data analytics in healthcare. Dantanarayana et al. [21] explain that the health community is currently experiencing an increase in complex and large data content resulting from corresponding data gathered through its information systems. In addition, service providers have expanded their services in the form of nursing and primary care, increasing the differentiation and diversity of health care services [22]. Thus, the health environment has become increasingly complex and challenging, demanding new technological approaches such as big data analytics that solve these major issues.

Additionally, the digitization of healthcare information creates opportunities and possibilities for improved care quality, health outcomes, and lower costs. Godbole and Lamb [23] explain that the health sector is achieving these goals by using data analytics to analyze various data sources such as electronic health records (EHRs), Enterprise Resource Planning (ERP), data from sensors, and other monitoring devices, clinician notes, and medical images. The combined information provides critical insights about each patient, allowing the personalization of services and medication [24]. Besides, the diversity of big data technologies allows hospitals to collect and analyze demographic, social, behavioral, and environmental data needed for evidence-based practice and information practice on unique trends that may not have otherwise been identified in the traditional data collection and analysis systems [25]. Thus, embracing big data analytics will transform health care by fostering efficiency, enabling patients to manage their health, enhancing care quality, and promoting sustainable healthcare approaches.

Unlike other business sectors, big data in the healthcare sector is still in its early stages. Dantanarayana et al. [21] associated the slow progress with multiple challenges unique to the industry including, the variety and velocity of healthcare data and the inability to accommodate the large volumes. For instance, healthcare data comes from multiple diverse sources, including clinical records, financial systems, genomic data, and administrative systems [26]. In addition, since most healthcare workers are within health-related fields, the sector struggles with acquiring skilled analysts to transform the large data volumes into valuable and applicable information [27]. Thus, embracing the new big data technologies in the health sector will require changes in the infrastructure and hiring employees with the needed skills and knowledge.

4. Materials and Methods

The research utilizes a Systematic Bibliometric Literature Review (LRSB) methodology to gather and analyze the impact of big data analysis on health sector [28, 29, 30, 31]. Okoli [32] (p.880) defined literature review as "a systematic, explicit, [comprehensive], and reproducible method for identifying, evaluating, and synthesizing the existing body of completed and recorded work produced by researchers, scholars, and practitioners". Snyder [33] built on this definition, indicating that this integration of findings and perspectives from multiple empirical kinds of research makes literature review a powerful methodology since it answers research questions in ways that other single studies cannot. These explanations form the rationale for the methodology selection since the researcher aims to provide valid, reliable, and repeatable findings and conclusions that can be used to improve care practice and decision-making.

This study employs a step-by-step approach to ensure the use of well-defined and rigorous criteria for identifying, appraising, and synthesizing existing data. The researcher used the four-step procedure proposed by Mariano et al. [6] to define the protocol, collect references, evaluate data, and interpret the findings.

 doi:10.20944/preprints202203.0407.v1

Under defining the protocol step, the researcher identified the research objectives and questions to guide the study. The primary research question for the study was: How big data analytics technologies have influenced the health sector? The objective was to identify the various big data analytics tools used in the health sector and contribute to its revolution. In addition, the research identified various challenges and opportunities associated with embracing big data technologies and their impact on practice.

To collect references the researcher used the database of indexing scientific articles SCOPUS, the most important peer review in the academic world. The use of Scopus alone is due to the fact that it is the main article base for academic journals/magazines, covering around 19,500 titles from more than 5,000 international publishers, including coverage of 16,500 peer-reviewed journals in the fields scientific, technical, and medical and social sciences. Thus providing a very real view of the researched subjects with scientific and/or academic relevance. The bibliographic search was limited to academic and scientific documents, including journal articles, books and book chapters, and conference articles published until November 2021.

The literature search began with using the keyword "Big data analytics" to screen abstracts and titles, where 8844 documents were identified. Mariano et al. [6] recommended repeating steps until a final list of the best quality sources is identified. Thus, the specific keyword "health" was added to the search query to narrow down the search and reduce the number of search results.

The data evaluation stage involved analyzing all the references identified for inclusion. The four steps used for the data evaluation were title evaluation, abstract evaluation, diagonal reading, and full-text reading. These stages ensure that the articles selected match the research question and objectives.

The findings interpretation phase involved summarizing themes, comparing results, and synthesizing data. The search and evaluation procedure are summarized in Table 1.

Table 1. Process of systematic LRSB

Database Scopus	Screening	Publications
Meta-search	keyword: Big data analytics	8,844
Inclusion Criterion	keyword: Big data analytics, Health	77
Screening	keyword: Big data analytics, Health Published until December 2021	77

Source: own elaboration

The 77 scientific and/or academic documents are subsequently analyzed narratively to deepen the content and the possible derivation of common themes that directly answer the article's research question [28, 29, 30, 31]. Of the 77 scientific and/or academic documents selected, 54 are Conference Papers; 18 Articles; 4 Reviews; and 1 Book Chapter.

4. Literature analysis: themes and trends

Peer-reviewed documents on the topic were selected for the period 2014-2021. The year 2019 was the year with the highest number of peer-reviewed publications on the subject, reaching 14.

Figure 1 summarizes the peer-reviewed literature published for the period 2014-2021. As we can see interest in the subject has varied over time.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 31 March 2022 doi:10.20944/preprints202203.0407.v1

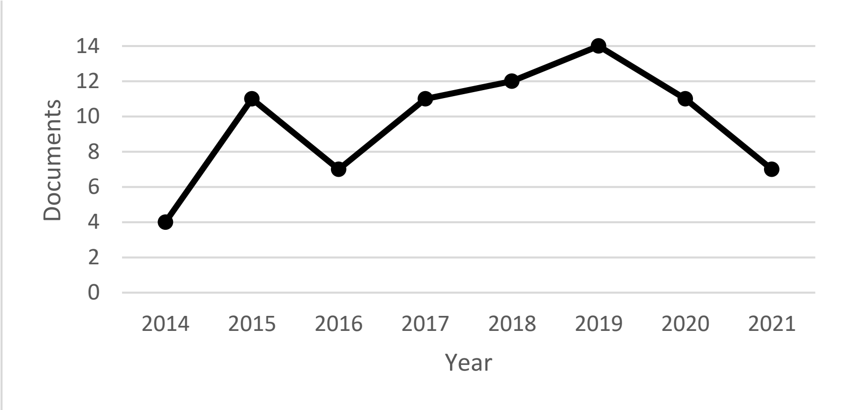

Figure 1. Documents by year. Source: own elaboration.

In Table 3 we analyzed for the Scimago Journal & Country Rank (SJR), the best quartile, and the H index by publication. Renewable And Sustainable Energy Reviews is the most quoted publication with 3,520 (SJR), Q1, and H index 295.

There are a total of 10 journals in Q1, 6 journals in Q2, 3 journals in Q3, and 37 journals without publications in Q4. Journals from best quartile Q1 represent 18% of the 56 journals titles; best quartile Q2 represents 11%, best quartile Q3 represents 5%, and finally, 37 of the publications representing 66%, the data are not available.

As evident from Table 2, the significant majority of articles on Impact of Big Data Analysis on Health rank on the Q1 best quartile index.

Table 3. Scimago journal & country rank impact factor

Title	SJR	Best Quartile	H index
Renewable And Sustainable Energy Reviews	3,520	Q1	295
Fertility And Sterility	2,270	Q1	208
Journal Of Industrial Information Integration	2,040	Q1	24
Science Of The Total Environment	1,800	Q1	244
Sustainable Cities And Society	1,650	Q1	61
Journal Of Dentistry	1,500	Q1	114
Future Generation Computer Systems	1,260	Q1	119
IEEE Transactions On Biomedical Engineering	1,150	Q1	200
International Journal Of Machine Learning And Cybernetics	0,680	Q1	44
IEEE Access	0,590	Q1	127
International Journal Of Environmental Research And Public Health	0,750	Q2	113
Sensors	0,640	Q2	172
Enterprise Information Systems	0,600	Q2	47
Mobile Networks And Applications	0,450	Q2	85
Journal Of Systems And Information Technology	0,300	Q2	25
Interdisciplinary Science Reviews	0,240	Q2	24
Studies In Health Technology And Informatics	0,300	Q3	58
Sensors And Materials	0,260	Q3	28
IETE Journal Of Research	0,220	Q3	25

 doi:10.20944/preprints202203.0407.v1

Procedia Computer Science	0,330	-*	76
Proceedings 2017 IEEE 19th Conference On Business Informatics CBI 2017	0,420	-*	11
Proceedings IEEE Symposium On Computer Based Medical Systems	0.260	-*	36
2017 IEEE 5th International Conference On Serious Games And Applications For Health Segah 2017	0,250	-*	10
Proceedings 2018 IEEE International Conference On Big Data Big Data 2018	0,250	-*	15
19th AIAA Non Deterministic Approaches Conference 2017	0,230	-*	7
Proceedings 2019 IEEE World Congress On Services Services 2019	0,230	-*	4
Proceedings 2017 5th International Conference On Enterprise Systems Industrial Digitalization By Enterprise Systems Es 2017	0,200	-*	6
Proceedings Of The International Astronautical Congress Iac	0,190	-*	15
ACM International Conference Proceeding Series	0,180	-*	123
Proceedings 2019 22nd International Conference On Control Systems And Computer Science Cscs 2019	0,140	-*	4
Proceedings 2019 IEEE International Conference On Smart Manufacturing Industrial And Logistics Engineering Smile 2019	0,100	-*	1
Proceedings Of The Annual Hawaii International Conference On System Sciences	0	-*	83
Proceedings Of The ASME Turbo Expo	0	-*	44
IEEE International Conference On Automation Science And Engineering	0	-*	28
Proceedings 2014 IEEE International Conference On Big Data IEEE Big Data 2014	0	-*	20
Proceedings 2015 IEEE International Congress On Big Data Bigdata Congress 2015	0	-*	16
International Conference On Electrical Electronics And Optimization Techniques Iceeot 2016	0	-*	16
2016 IEEE 18th International Conference On E Health Networking Applications And Services Healthcom 2016	0	-*	13
Pacific Asia Conference On Information Systems Pacis 2015 Proceedings	0	-*	10
15th International Conference On Advances In ICT For Emerging Regions Icter 2015 Conference Proceedings	0	-*	4
2015 12th International Conference And Expo On Emerging Technologies For A Smarter World Cewit 2015	0	-*	4

2016 3rd Mec International Conference On Big Data And Smart City Icbdsc 2016	0	-*	9
Proceedings 2015 2nd IEEE International Conference On Advances In Computing And Communication Engineering Icacce 2015	0	-*	9
Proceedings Of 2014 International Conference On Electrical Engineering And Computer Science Iceecs 2014	0	-*	5
Proceedings Of The 2014 International Conference On Artificial Intelligence Icai 2014 Worldcomp 2014	0	-*	5
Lecture Notes In Computer Science Including Subseries Lecture Notes In Artificial Intelligence And Lecture Notes In Bioinformatics	-*	-*	-*
2015 17th International Conference On E Health Networking Application And Services Healthcom 2015	-*	-*	-*
2020 IEEE 4th Conference On Energy Internet And Energy System Integration Connecting The Grids Towards A Low Carbon High Efficiency Energy System Ei2 2020	-*	-*	-*
2nd International Conference On Computational Systems And Information Technology For Sustainable Solutions Csitss 2017	-*	-*	-*
Asonam 2014 Proceedings Of The 2014 IEEE ACM International Conference On Advances In Social Networks Analysis And Mining	-*	-*	-*
Data Democracy At The Nexus Of Artificial Intelligence Software Development And Knowledge Engineering	-*	-*	-*
Icevt 2019 Proceeding 6th International Conference On Electric Vehicular Technology 2019	-*	-*	-*
Iciiecs 2015 2015 IEEE International Conference On Innovations In Information Embedded And Communication Systems	-*	-*	-*
Proceedings CAMP 2021 2021 5th International Conference On Information Retrieval And Knowledge Management Digital Technology For Ir 4 0 And Beyond	-*	-*	-*
Society Of Petroleum Engineers Abu Dhabi International Petroleum Exhibition And Conference 2020 Adip 2020	-*	-*	-*
Towards Sustainable Technologies And Innovation Proceedings Of The 27th Annual Conference Of The International Association For Management Of Technology Iamot 2018	-*	-*	-*

Note: *data not available. Source: own elaboration

The subject areas covered by the 66 scientific articles were: Computer Science (46); Engineering (28); Medicine (15); Mathematics (12); Decision Sciences (9); Health Profes-

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 31 March 2022 doi:10.20944/preprints202203.0407.v1

sions (9); Social Sciences (9); Business, Management and Accounting (5); Energy (5); Environmental Science (3); Materials Science (3); Physics and Astronomy (3); Earth and Planetary Sciences (2); Psychology (2); Arts and Humanities (1); Biochemistry, Genetics and Molecular Biology (1); Dentistry (1).

The most quoted article was "Smart clothing: Connecting humans with clouds and big data for sustainable health monitoring" from Chen et al. (2016) with 237 quotes, published in the Service Industries Journal 0,450 (SJR), the best quartile (Q2), and with H index (85). The published article focuses on the study of design details, key technologies, and practical implementation methods of the smart clothing system.

In Figure 2 we can analyze the evolution of articles citations published between ≤2014 and 2021. The number of quotes shows positive net growth with an R2 of 89% for the period ≤2014-2021, with 2020 reaching 220 citations.

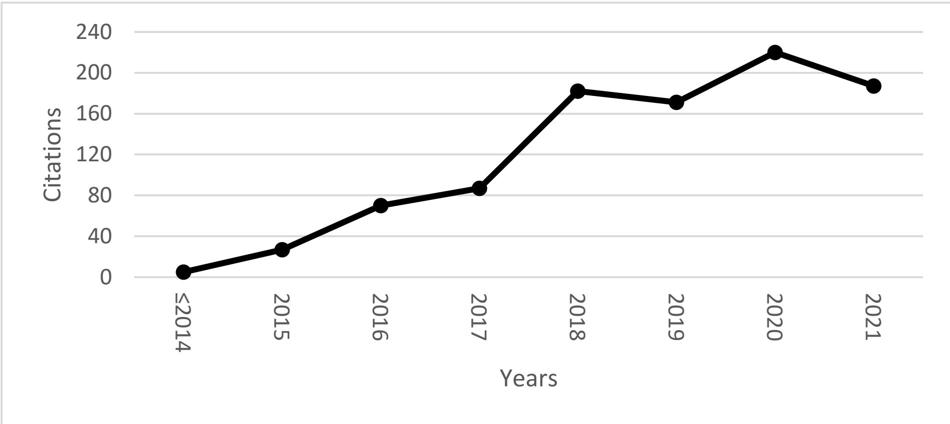

Figure 2. Evolution of citations between ≤2014 and 2021. Source: own elaboration

The h-index was used to ascertain the productivity and impact of the published work, based on the largest number of articles included that had at least the same number of citations. Of the documents considered for the h-index, 12 have been cited at least 12 times.

In Appendix A, the citations of all scientific articles from the ≤2014 to 2021 period are analyzed, with a total of 954 citations, of the 77 publications 14 were not cited.

Appendix B examines the self-citation of the document during the period ≤2014 to 2021, 33 documents were self-cited 169 times, the article "Digital fitness: Four principles for successful development of digital initiatives" by Melville (2015) published in the Paper presented at the Proceedings of the Annual Hawaii International Conference on System Sciences was cited 38 times.

In Figure 3, a bibliometric study was performed to examine the development of scientific information by the main keywords. The study of bibliometric outputs by the scientific software VOSviewe aims at identifying the main research keywords as "big data analytics", "data analytics" and "Health". The research relied upon the studied articles on the use and influence of big data analytics tools in healthcare. The correlated keywords can be viewed in Figure 4 allowing to making clear the network of keywords that appear together/linked in each scientific article, as well as knowing the topics studied to identify future research trends. Also, Figure 5 illustrates co-citations with a unit of analysis of cited references.

 doi:10.20944/preprints202203.0407.v1

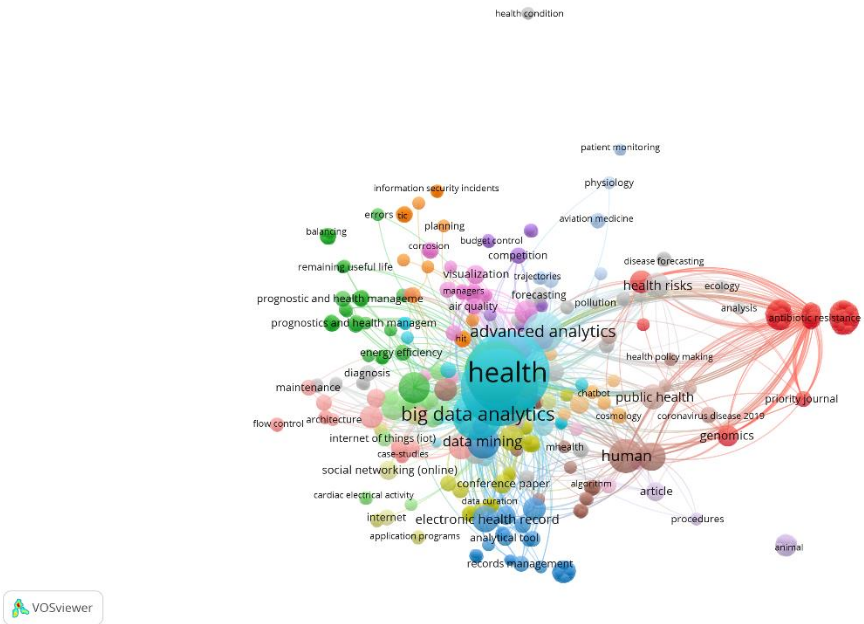

Figure 3. Network of all keywords.

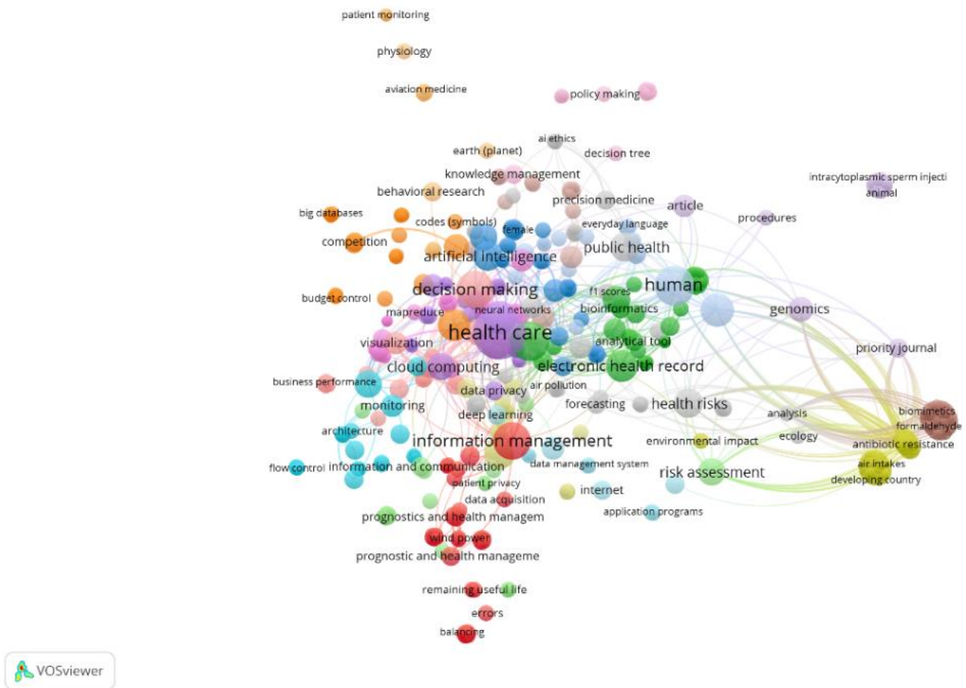

Figure 4. Network of Linked Keywords.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 31 March 2022 doi:10.20944/preprints202203.0407.v1

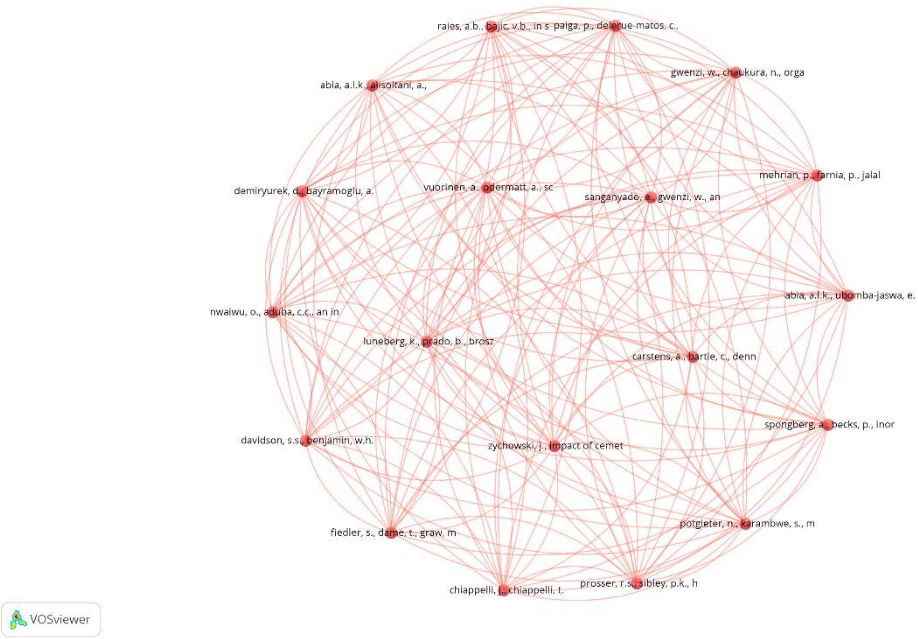

Figure 5. Networks bibliographic coupling.

5. Theoretical perspectives

5.1. The Impact of Big Data Analytics in Healthcare

Big Data has the potential to revolutionize the healthcare sector by promoting accurate patient diagnosis and treatment. With Big Data tools, analysts will collect accurate information within the health informatics systems, which can be analyzed and implemented to improve care delivery models and approaches [34]. Based on the notion that value is derived from balancing patient impacts (outcomes) and healthcare spending (cost), big data can have a significant impact on the patient-centered framework in five dimensions or pathways illustrated in Figure 6 [35]. By improving patient outcomes concerning the five pathways, the healthcare system will provide more value to the patients and other stakeholders, leading to significant transformations.

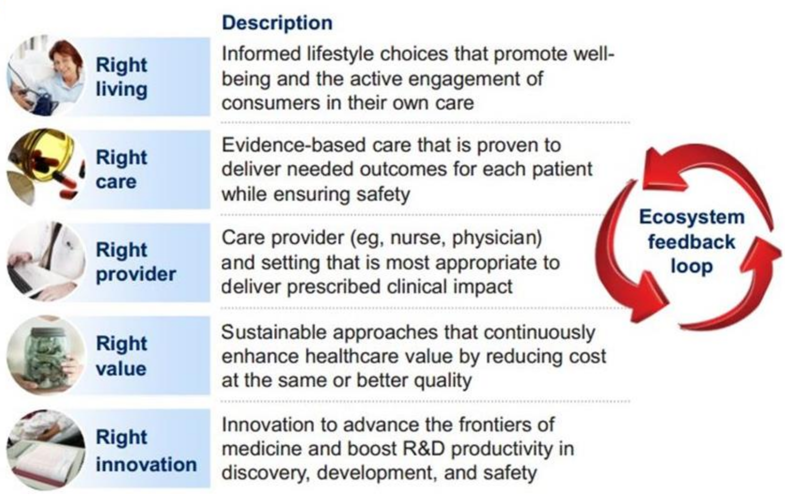

Figure 6. Big Data-Related Value Pathways in Healthcare. Source: Groves et al. (2013) [35].

5.1.1. Right living

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 31 March 2022 doi:10.20944/preprints202203.0407.v1

The right-living pathway refers to patients living a healthier and better life by actively participating in their treatment. With digital technologies, patients have access to medical information, which can be used to make decisions and better choices that enhance their wellbeing [35]. For instance, informed patients are likely to choose the right path for their diet, exercise, and preventive care, leading to better health outcomes. In addition, Kowalski [36] explained that patients use the internet and social networking sites to publish reviews about health services and products. As a result, healthcare facilities and agencies use big data analytics tools to extract valuable information from these reviews to improve their services and products, thus increasing patient access to quality care (Barik et al., 2018). Besides, the integration of patient feedback can increase their satisfaction and trust in healthcare procedures, professionals, and providers [24]. This argument indicates patients' active engagement in value creation in modern healthcare systems, leading to positive behaviors.

5.1.2. Right care

The right-care pathway ensures that all patients have access to effective treatment. It is based on providers' access to similar data and aligned objectives to ensure appropriate planning and elimination of redundancy. Groves et al. [35] explained that the right care pathway involves a coordinated approach where caregivers within the network have access to similar information and focus on achieving shared goals. Kuiler and McNeely [37] explained that Big Data health analytics and health informatics facilitate the creation of knowledge-focused systems that influence the distribution of health information and data access to support professional research, the establishment of evidence-based policies, and caregiver and patient participation in treatments. Big Data analytics allow healthcare stakeholders to leverage large complex datasets obtained from various sources by encouraging the creation of frameworks that support electronic health information exchange [38]. For the information-based systems to effectively function, all the key players including clinicians, financial personnel, and administration, must continuously update information on the system. This exchange provides valuable data used for decision-making in care provision.

5.1.3. Right provider

The right-provider pathway requires matching professionals with patient problems to ensure high outcomes. In this case, only the high-performing professional with the right knowledge and skills needed to treat a patient's specific problems should be assigned the task [35]. This pathway helps avoid misdiagnosis and wrong prescription problems that can have detrimental consequences [39]. In addition, Brennan and Bakken [40] indicated that big data and science challenges nurse educators to design curricula that match the current knowledge and skills demand. By transforming the education systems, big data will ensure that the professionals attending patients have adequate skills and knowledge regarding their specific health issues and enhance access to appropriate informational resources [41]. These activities will increase patients' access to the right providers.

5.1.4. Right value

The right-value pathway indicates that big data analytics tools will improve healthcare value while enhancing its quality. As the availability of new information improves with big data technologies, healthcare providers will better understand effective and appropriate care approaches [41]. Under this pathway, care providers are expected to consider multiple alternatives to achieve cost-effectiveness, such as reimbursement policies and eliminating waste and corruption within the system. Ojha and Mathur [42] contributed to this dimension by indicating that big data provides insights into hidden facts and figures to present solutions to lowering healthcare costs. The large volumes of data collected and analyzed can be used to identify various measures that can be implemented

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 31 March 2022 doi:10.20944/preprints202203.0407.v1

to increase financial transparency and reduce costs of care [43]. In addition, the infrastructures and models will help create an ecosystem feedback loop that will continuously improve value co-creation.

5.1.5. Right innovation

The right-innovation pathway focuses on enhancing innovations such as boosting R&D and advancing medicine. Through innovations, care providers can identify new interventions and strategies for delivering care, thus improving healthcare outcomes and efficiency [44]. Thus, big data analytics are associated with advancements in the provision of patient services that promote patient health and wellbeing and can lead to policy changes, including insurance. Big Data contributes to this by providing the information needed to identify and exploit opportunities, such as improving traditional medical protocols or clinical trials [45]. Right innovation can lead to timely identification of new diseases and treatments that will continuously evolve the medicine [46]. Thus, Big Data has significantly impacted medical innovations and led to the continuous growth of the systems, procedures, and care models, leading to operational efficiency, reduced costs, collaboration, and increased patient outcomes.

5.2. The Impact of Big Data on Healthcare Stakeholders

This section may be divided by subheadings. It should provide a concise and precise description of the experimental results, their interpretation, as well as the experimental conclusions that can be drawn.

5.2.1. Patients

Patients' access to medical knowledge from platforms such as hospital websites, clinical forums, and social networking sites increases their capacity for self-management. One major characteristic of Big Data is decentralized information that allows all stakeholders access to essential information to support informed decision-making [47]. Patients aware of their health conditions, medications, and potential home remedies are more likely to adopt appropriate behaviors and attitudes, increasing health outcomes [48]. In addition, data analytics tools increase the range of affordable healthcare services and facilitate personalized recommendations [49]. The data mined and analyzed provides critical insights on social, medical, and environmental information associated with individual patients, leading to the implementation of appropriate treatments and prevention care.

5.2.2. Healthcare insurers

The new analytical tools provide health insurers with opportunities to introduce new and targeted care plans. For instance, they can analyze geographical data to identify occurring diseases in various locations and introduce health plans at minimal premium costs [47]. In addition, Big Data analytics can provide insights on the correlation between demographic characteristics such as gender, age, incomes, and family history to develop and implement care plans [50]. Predictive modeling techniques can help insurance companies analyze unstructured data to predict patterns of factual claims and eliminate fraud among customers and within the organization [51]. Real-time data acquired through big data analytics can help analyze insurance customers' behaviors and use the data to develop and implement new innovative business models that transform work processes and boost customer experiences.

5.2.3. Medical practitioners

Big data provides medical practitioners with massive valuable data that provides a holistic view of patients' progress and treatment response. The data gathered can include patient history, laboratory results, medical imaging data, diseases classification codes, and other critical information captured through sensors and monitoring devices [47]. In addi-

 doi:10.20944/preprints202203.0407.v1

tion, Big Data increases clinicians' capability to combine data from multiple sources, including socioeconomic data, medical equipment, and public health statistics [1] (p.51). Consequently, medical practitioners are empowered to conduct targeted health investigations that lead to higher patient outcomes [52]. Besides, the increased access to data and technologies allows medical practitioners to improve their understanding of various medical phenomena such as emerging disease patterns and develop appropriate knowledge and skills to handle them. Thus, Big Data has significantly contributed to medical practitioners' competitiveness.

5.2.4. Pharma and clinical researchers

Big Data has significantly contributed to healthcare reformation through pharmaceutical and clinical research. For instance, predictive models based on big data analytics are used to understand drug and biological processes, leading to effective drug designs [1]. In addition, pharma companies can use diverse data sources to conduct trials to evaluate the outcomes of designed drugs [53]. In clinical research, big data analytical tools are used to collect information from electronic databases for analysis [47]. The data is obtained from unmodified daily clinical practice routines. Under big-data clinical trial (BCT), the data is collected from electronic medical record systems, health insurance systems, and administrative registries for infectious and chronic diseases [54]. The primary advantage of using big data in clinical and pharmaceutical research is that the outcomes directly reflect the clinical effectiveness of the studied variables or designs since the researchers use accurate and timely data gathered from diverse sources [55]. The evidence acquired through research findings can be used to promote patient-centered and evidence-based practice. Thus, through improved research, big data is transforming the healthcare system by providing critical data needed to transition from the traditional data management systems to the new big data systems.

5.3. Application of Big Data Frameworks in Health Informatics

Recent research and developments in the healthcare sector emphasize implementing various big data frameworks to enhance handling large data volumes. Under this section, four frameworks: EHRs, machine learning, predictive analytics, and IoT and their contribution to health informatics.

5.3.1. Electronic Health Records (EHRs)

Electronic health records (EHRs) present capabilities to enhance healthcare quality and efficiencies in health information technology. EHRs present capabilities to reduce health disparities in populations and include patient records stored in digital forms [50]. They include patient information such as contact information, medical histories and tests, and treatment plans [56]. EHRs provide considerable benefits that include improved efficiencies, positive patient outcomes, and population health and can analyze huge quantities of patient information, making the technologies beneficial in improving population health [57]. In healthcare, the most widespread healthcare application of big data is the EHRs. The analysis of patient data ensures that patient records are available and shared via secure information systems to private and public healthcare providers [58]. The information systems comprise modifiable files that allow doctors to make changes based on new medical test results without duplications or further paperwork [59]. Due to their complexity and ability to handle a considerable amount of patient information extracted from various sources for various purposes, EHRs can be considered intrinsically big data health resources.

5.3.2. Machine Learning in Healthcare

Machine Learning (ML) is an important layer of secure Big Data-based healthcare framework. The ML consists of other sub-modules such as early diagnosis, epidemic forecasting models, and data analytics that depend on machine learning methods, including genetic algorithms and support vector machines [47]. In healthcare, ML is a concept that is closely similar to data mining. The two approaches entail scanning data to identify patterns. However, the Machine Learning approach uses it to enhance the program's understanding. Unlike data mining, ML changes the program's function depending on the identified data patterns [60]. Therefore, Machine Learning is commonly identified as a learning process based on the input data to develop enough experience to present the required output. The ML approaches can be supervised or unsupervised [1]. The supervised approach entails predicting missing data set based on the information and data provided during the training process, while unsupervised learning does divide data into training and testing sets [61]. Data are grouped into subsets of similar objects based on their common and various features, thus enabling the earning process [62]. The Machine Learning output can help identify, diagnose, and treat disease, streamline healthcare tasks, plan, prepare, and execute tasks such as surgery. Therefore, ML presents capabilities to improve patient care delivery strategies.

5.3.3. Predictive Analytics in Healthcare

In healthcare, Predictive Analysis involves using forecasting techniques and predictive analytics to determine conditions with high occurrence probability in the future. Several modeling techniques are used to make robust predictions [63]. Such models include hierarchical linear models and the algorithms such as Artificial Intelligence and ML [64]. Predictive analytics has become a crucial component of personalized patient care. Predictive analytics can use the medical history, patient's social risk factors, environment, and genetics to customize a patient's treatment process in personalized healthcare [47]. Some of the key uses of predictive analysis include diagnosis, prognosis, improving care quality, reducing healthcare costs, remote monitoring, reducing adverse events, developing treatment courses, and supporting clinical decisions [1]. Predictive analytics is a crucial category of big data analytics that incorporates statistical methods, including data mining and machine learning, that use the historical patient's data to predict the future [65]. The predictive models used are helping healthcare providers identify patients at risk for readmission and helping healthcare providers make crucial patient treatment decisions based on the available data [66]. Predictive analysis requires a greater understanding of Machine Learning methods, which are crucial in predicting future health conditions. Based on current and historical data, predictive analytics can help to anticipate and reduce healthcare risks and chances of patient deterioration.

5.3.4. Internet of Things (IoT) in smart health

Smart health care comprises three application domains classified the healthy living, home care, and acute care. Healthy living entails tracking fitness, preventing diseases, and food monitoring [67]. Home Care includes mobile health, self-management, assisted living, and telemedicine. Acute Care includes hospitalization, specialty clinic, nursing homes, and community hospitals. IoT in smart health care provides a combination of connectivity, ambient intelligence, and ubiquitous communications that enhance such application domains' achievements [68]. It ensures the connection of all real-world communications in a cyber-physical paradigm [69]. In addition, it provides an integration of electronic devices, tablets, smartphones, and other real elements that can share data and information either wirelessly or physically [67]. IoT technologies enhance the management of devices and are associated with several benefits and attributes such as increased connectivity, personalized monitoring systems, sensing, location, and identification [71]. The technology presents capabilities to enhance managing chronic diseases and monitoring daily fitness goals [72]. Within smart healthcare systems, IoT infrastructures can help to bridge the gap between patient and their doctors [73]. The infrastructures can collect and transmit crucial medical information from patients [74]. The healthcare providers can use remote access to monitor patients and provide remote consultation services [75]. IoT com-

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 31 March 2022 doi:10.20944/preprints202203.0407.v1

bines technologies and digital tools such as sensors, local area networks, the internet, micro-controllers, actuators, and cloud computing technologies to obtain accurate patient information and results.

5.4. Opportunities of Big Data in Health Care

The Centers for Medicare & Medicaid Services (CMS) and other key players in the healthcare sector uses predictive modeling to prevent fraud. White [76] describes predictive modeling as a process that use "statistical techniques and historical data to estimate the probability of future results" (p.16). With the information analyzed, CMS contractors can differentiate fraudulent and authentic claims before paying for services [77]. In addition, predictive modeling can be used to understand how patients benefit from care plans [78]. For example, the model allows practitioners to identify population groups more likely to benefit from certain care plans (Li et al., 2015) [66]. Therefore, it can be used to improve resources allocation [79]. Healthcare providers can further use predictive modeling to determine disease patterns and their impact on various population groups based on characteristics such as geographical location, age, gender, or socioeconomic status. Thus, the providers can develop and implement targeted investigations and preventions to reduce the disease burden.

Another opportunity facilitated by big data is to make healthcare green. The diverse operations within the healthcare industry have raised concerns over its environmental impacts [23]. For instance, energy and water use and electronic waste (eWaste) can enormously impact the environment, prompting health institutions and regulators to advocate for green alternatives [80]. Combining big data analytics and cloud computing can help reduce energy consumption [81]. For example, developing energy-efficient systems [82] and application architecture such as virtual data storage and servers replacing traditional IT infrastructure can lower IT power consumption [83]. Besides, the traditional data systems are significantly reliant on paperwork, adversely impacting forests and the environment. Thus, Godbole and Lamb [23] indicated that embracing Electronic Medical Records (EMRs) can help hospitals avoid using approximately 1,044 tons of paper annually and reduce carbon dioxide emissions by 1.7 million tons in the US. Therefore, embracing big data and new technologies provide effective solutions to the prevailing environmental and climate concerns.

Additionally, big data technologies are promoting the use of telemedicine, thus increasing access to medical resources. Although telemedicine is not new, it was only accessible to people in the cities and urban areas [84]. However, with big data, telemedicine is now accessible to more people, including rural communities, consequently promoting chronic disease management and enhancing elderly care [23]. In addition, reducing tra veling to specialists can reduce gas emissions in the transportation industry since health services can be accessed from a patient's home or local facility.

5.5.Challenges of Big Data in Health Care

Despite the significant opportunities created by Big Data in the healthcare sector, embracing the developments is faced with multiple problems. Data segmentation is a primary problem affecting the effective use of big data in healthcare [85]. For instance, the financial and operational management teams store and manage administrative information such as reimbursement claims and costs, while the EHRs store clinical data, including diagnostic tests, progress notes, and medical history [76]. This segmentation limits access and exchange of information needed to make clinical decisions. For instance, combining clinical and socioeconomic data is identified as a critical opportunity for enhancing targeted medical investigations, personalized interventions, and modifying resource allocation policies [86]. However, storing data in silos throughout the organization hinders these opportunities. Each department has access to a certain amount of data relevant to their tasks instead of all the data within the data management systems [87]. Therefore, exploiting the opportunities presented by big data will require health organizations to embrace data systems that allow unrestricted flow and access of information.

Another challenge undermining efforts to leverage big data analytics in healthcare is concern about protecting patients' privacy. Data analytics revolutionizing the industry requires sharing data between key stakeholders, such as public health institutions, players, and providers [44]. However, these independent players, such as insurance companies and care providers, must protect patients from direct or indirect identification under the Health Insurance Portability and Accountability Act [88]. However, White [76] indicates that sharing patient information without identifying them while still maintaining its usefulness is challenging. Thus, in collaborative practices, these institutions remove identifying data such as names, geographical locations, contact information, and medical record numbers, among others [89]. This removal of some important information leads to insufficient datasets that cannot achieve all desired objectives [90]. For example, in cases where predictive modeling is used to determine disease patterns in certain locations or among certain demographics, geographic subdivisions data and information on an individual's medical history such as admission and discharge dates are required. Thus, the healthcare system needs to establish a strategy that balances data needs and patient protection policies to optimize big data technologies fully.

6. Conclusions

Big Data analytics have significantly impacted the healthcare sector by introducing opportunities and technologies that provide practical solutions to various challenges. For instance, data technologies can generate, process, and store large volumes of data to improve medical practice. Big data frameworks integrated with health informatics include EHRs, machine learning, predictive analytics, and the internet of things (IoT). Big data has led to establishing the five pathways model of the healthcare system that involves dimensions including right living, right care, right provider, right value, and right innovation. Right living involves patients living a healthier and better life by actively participating in their treatment. With digital technologies, patients have access to medical information, which can be used to make decisions and better choices that enhance their wellbeing [35]. Right care ensures that all patients have access to effective treatment. This involves a coordinated approach where caregivers within the network have access to similar information and focus on achieving shared goals. Right provider involves assigning care delivery duties to the most talented practitioners to ensure high outcomes. Only the high-performing professional with the right knowledge and skills needed to treat a patient's specific problems should be assigned the task [35]. Right value ensures that the outcomes achieved match the costs of care; identify various measures that can be implemented to increase financial transparency and reduce costs of care [43]. Right innovation ensures that the data generated through big data is used to ensure consistent developments. Through innovations, care providers can identify new interventions and strategies for delivering care, thus improving healthcare outcomes and efficiency [44]. Thus, this research shows several practical implications that big data analytics has in healthcare related to the opportunities created through the use of modern technologies for all stakeholders, including patients and care providers. Investing in the right big data analytics tools and infrastructures can improve clinicians, providers, and policy-makers capability to plan and implement patient-centered and evidence-based interventions characterized by high-quality care, low costs, and higher outcomes. Also, this research has made a theoretical contribution by highlighting big data analytics as practical tools to support and revolutionize the health sector. However, we consider that the study has some limitations. First, for considering only the SCOPUS database for the Systematic Review of Bibliometric Literature and excluding the other indexing academic databases [28, 29, 30, 31]. Future studies may combine various databases for improved generalizability of the results. Second, the review focused on English language publications and did not include other relevant publications in other languages. Third, the review process used articles and review papers, excluding other publications such as dissertations, books, and book chapters to extend the range of the data.

Author Contributions: Conceptualization, A.R. and J.D.; methodology, A.R. and J.D.; software, A.R. and J.D.; validation, A.R. and J.D.; formal analysis, A.R. and J.D.; investigation, A.R. and J.D.; resources, A.R. and J.D.; data curation, A.R. and J.D.; writing—original draft preparation, A.R. and J.D.; writing—review and editing, A.R. and J.D.; visualization, A.R. and J.D.; supervision, A.R. and J.D.; project ad-ministration, A.R. and J.D.; funding acquisition, A.R. and J.D. All authors have read and agreed to the published version of the manuscript.

Funding: This research is supported by national funding's of FCT - Fundação para a Ciência e a Tecnologia, I.P., in the project «UIDB/04005/2020».

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: We would like to express our gratitude to the editor and the referees. They offered valuable suggestions or improvements. The authors were supported by the GOVCOPP Research Center of the University of Aveiro, and COMEGI.

Conflicts of Interest: The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

Appendix A

Table A1. Overview of document citations period ≤2014 to 2021

Documents	Dat	≤2014	2015	2016	2017	2018	2019	2020	2021	Total
State-of-the-art in integrated prognostics and health m…	2021	-	-	-	-	-	-	-	1	**1**
Autopsy, thanatopraxy, cemeteries and cremat…	2021	-	-	-	-	-	-	1	4	**5**
Big data analytics capabilities: a novel integrated fitness ...	2021	-	-	-	-	-	-	-	1	**1**
Beyond 'AI for Social Good' (A14SG): social tra…	2021	-	-	-	-	-	-	-	2	**2**
The 'thanato-resistome' - The funeral industry as a potentia ...	2020	-	-	-	-	-	-	-	4	**6**
A data-informed public health policy-makers platform	2020	-	-	-	-	-	-	1	1	**2**
Knowledge formulation in the health domain: A sem…	2020	-	-	-	-	-	-	1	2	**3**
Development of Novel Big Data Analytics Framework …	2020	-	-	-	-	-	-	-	2	**2**
Development of Big Data Analytics Platform for El…	2019	-	-	-	-	-	-	1	-	**1**
Random forest for big data classification in the internet of..	2019	-	-	-	-	-	11	21	21	**53**

 doi:10.20944/preprints202203.0407.v1

An automated detection system of drug-drug interact…	2019	-	-	-	-	-	-	1	1	**2**
lntegrating wearable technology products and big data analyt...	2019	-	-	-	-	-	1	4	5	**10**
A big data architectu refor the extraction and analysis of E ...	2019	-	-	-	-	-	1	2	1	**4**
lot-based healthcare remote monitoring platform for …	2019	-	-	-	-	-	-	2	4	**6**
Twitter Health Surveillance (THS) System	2019	-	-	-	-	-	1	-	2	**3**
Analyzing countermeasure effectiveness utilizing big …	2019	-	-	-	-	-	-	-	2	**2**
A prognostics and health management framework forwind	2019	-	-	-	-	-	-	1	1	**2**
General System Theory and lhe Use of Process Mining …	2019	-	-	-	-	-	-	1	1	**2**
Usability Evaluation of a Co-created Big Data Analytics Plat...	2019	-	-	-	-	-	-	2	-	**2**
A Hadoop/MapReduce based platform for supporting heal…	2019	-	-	-	-	-	-	1	2	**3**
lmpact ofhealth differences and longitudinal changes on dec. ..	2019	-	-	-	-	-	-	2	4	**6**
An Architecture for Performing Real Time lntegrated Heal…	2018	-	-	-	-	-	1	-	-	**1**
Re-architecting oral healthcare for the 21st century	2018	-	-	-	-	-	1	5	5	**11**
Privacy preserving data by conceptualizing smart cities usin ...	2018	-	-	-	-	-	1	2	2	*5*
Large scale integration of wireless sensor network technolog ...	2018	-	-	-	-	-	2	8	6	**16**
Formal Methods, artificial intelligence, big-data analytics, ...	2018	-	-	-	-	-	-	2	1	**3**
Ontologies in Big Health Data Analytics: Application to Rout...	2018	-	-	-	-	-	-	1	2	**3**
Terminology Coverage from Semantic Annotated Health Doc…	2018	-	-	-	-	-	-	1	-	**1**
lmproving the use ofbig data analytics within electronic he ...	2018	-	-	-	-	1	3	3	5	**12**
Mist Data: Leveraging Mist Computing for Secure and Sc…	2018	-	-	-	-	8	5	6	3	**22**
Big Data Analytics for Air Quality Monitoring ata Logistics ...	2017	-	-	-	-	1	5	-	-	**6**

Meeting Technology and Methodology into Health Big Data …	2017	-	-	-	-	1	1	1	2	**5**
Efficient Large-scale Medical Data (eHealth Big Data) …	2017	-	-	-	-	5	7	11	4	**27**
Algorithms for big data delivery overthe internet ofthings	2017	-	-	-	-	-	4	1	-	**5**
Big data analytic based personalized air quality health advi ...	2017	-	-	-	-	-	2	-	-	**2**
lntegrating Big Data analytics, virtual reality, and ARAIG t...	2017	-	-	-	-	1	3	3	1	**8**
-Omic and Electronic Health Record Big Data Analytics …	2017	-	-	-	8	22	32	50	23	**135**
Big data analytics advances in health intelligence, public h ...	2017	-	-	-	-	2	-	2	1	**5**
Patients' written reviews as a resource for public healthcar. ..	2017	-	-	-	-	1	1	-	1	**3**
Mobile health application running on public cloud …	2017	-	-	-	-	-	-	1	1	**2**
Big data analytics for continuous assessment of …	2017	-	-	-	1	-	-	-	1	**2**
Big data application: Study and archival of mental health da ...	2016	-	-	-	2	1	1	2	-	**6**
Hl-risk: A method to analyse health information risk intelli ...	2016	-	-	-	-	-	-	1	-	**1**
Smart Clothing: Connecting Human with Clouds and Big Data fo ...	2016	-	-	12	35	57	42	50	41	**237**
Proposed application ofbig data analytics in healthcare at ...	2016	-	-	-	3	5	1	-	7	**16**
Quality of information for quality of life: Healthcare big d ...	2016	-	-	-	1	-	-	-	-	**1**
Point-of-care testing in the time of PS medicine: A preface	2016	-	-	-	-	2	-	-	-	**2**
Big Social Data in Public Health: A Mixed-methods Case Study ...	2016	-	-	-	-	2	1	1	-	**4**
Using data science &: big data analytics to make healthca ...	2015	-	-	-	3	3	3	1	1	**11**
Security Solutions for Big Data Analytics in Healthcare	2015	-	-	-		4	2		2	**8**
Mobile sensing and network analytics for realizing smart aut...	2015	-	-	4	4	4	1	3	5	**21**
H-DRIVE: A Big Health Data Analytics Platform for Evidence-1. ..	2015	-	-	-	-	2	3	2	1	**8**

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 31 March 2022 doi:10.20944/preprints202203.0407.v1

Indoor Air Monitoring Platform and Personal Health Reporting ...	2015	-	-	-	-	2	3	2	2	**9**
Big Data Analytics Framework for System Health Monitoring	2015	-	1	-	3	4	7	4	1	**20**
Comparitive study on healthcare prediction systems using big ...	2015	-	-	2	1	-	2	1	2	**8**
Adoption of big data analytics in healthcare: The efficiency ...	2015	-	-	3	1	-	1	2	-	**7**
Information accountability and Health Big Data Analytics: A ...	2015	-	-	1	1	3	-	-	-	**5**
Digital fitness: Four principies for successful development ...	2015	-	-	-	-	-	1	-	-	**1**
Knowle: A semantic link network based system for organizing ...	2015	5	22	43	15	19	7	3	3	**117**
Implementation of electronic medical record in community hea ...	2014	-	-	-	2	1	2	1	1	**7**
Case studies: Big data analytics for system health monitorin ...	2014	-	1	-	-	-	-	1	-	**2**
Spatial big data analytics of influenza epidemie in Vellore, ...	2014	-	2	4	7	31	12	6	5	**67**
Big Data in Health Informatics Architecture	2014	-	1	1	-	-	-	2	-	**4**
	Total	**5**	**27**	**70**	**87**	**182**	**171**	**220**	**187**	**954**

Source: own elaboration

Appendix B

Table A1. Overview of document self-citation period *≤2014 to 2021*

Documents		**≤2014**	**2015**	**2016**	**2017**	**2018**	**2019**	**2020**	**2021**	**Total**
Autopsy, thanatopraxy, cemeteries and cremat…	2021	-	-	-	-	-	-	-	2	**2**
Online status assessment o distribution network transformer ...	2020	-	-	-	-	-	-	-	6	**6**
Knowledge formulation in the health domain: A sem…	2020	-	-	-	-	-	-	-	1	**1**
Random forest for big data classification in the internet of..	2019	-	-	-	-	-	7	1	1	**9**
Integrating wearable technology products and big data analyt...	2019	-	-	-	-	-		2	4	**6**
A big data architectu refor the extraction and analysis of E ...	2019	-	-	-	-	-	1	1	-	**2**
Iot-based healthcare remote	2019	-	-	-	-	-	-	-	1	**1**

monitoring platform for …										
Analyzing countermeasure effectiveness utilizing big …	2019	-	-	-	-	-	-	-	2	**2**
A prognostics and health management framework forwind	2019	-	-	-	-	-	-	-	1	**1**
Usability Evaluation of a Co-created Big Data Analytics Plat...	2019	-	-	-	-	-	-	2	-	**2**
lmpact ofhealth differences and longitudinal changes on dec. ..	2019	-	-	-	-	-	-	2	2	**4**
Re-architecting oral healthcare for the 21st century	2018	-	-	-	-	-	-	1	-	**1**
Large scale integration of wireless sensor network technolog ...	2018	-	-	-	-	-	-	1	-	**1**
lmproving the use ofbig data analytics within electronic he ...	2018	-	-	-	-	-	-	1	-	**1**
Mist Data: Leveraging Mist Computing for Sec…	2018	-	-	-	-	-	4	1	-	**5**
Big Data Analytics for Air Quality Monitoring ata Logistics ...	2017	-	-	-	-	1	-	-	-	**1**
Efficient Large-scale Medical Data (eHealth Big Data) …	2017	-	-	-	-	1	1	1	-	**3**
lntegrating Big Data analytics, virtual reality, and ARAIG t...	2017	-	-	-	-	1	1	2	1	**5**
-Omic and Electronic Health Record Big Data Analytics …	2017	-	-	-	-	1	1	-	-	**2**
Mobile health application running on public cloud …	2017	-	-	-	-	-	-	-	1	**1**
Big data analytics for continuous assessment of …	2017	-	-	-	1	-	-	-	1	**2**
Big data application: Study and archival of mental health da ...	2016	-	-	-	-	1	-	1	-	**2**
Smart Clothing: Connecting Human with Clouds and Big Data fo ...	2016	-	-	3	9	2	-	2	-	**16**
Point-of-care testing in the time of PS medicine: A preface	2016	-	-	-	-	1	-	-	-	**1**
Big Social Data in Public Health: A Mixed-methods Case Study ...	2016	-	-	-	-	1	-	-	-	**1**
Using data science &: big data analytics to make healthca ...	2015	-	-	-	-	2	1	-	-	**3**
Mobile sensing and network analytics for realizing smart aut...	2015	-	-	2	1	2	1	2	2	**10**

H-DRIVE: A Big Health Data Analytics Platform for Evidence-1. ..	2015	-	-	-	-	1	2	-	-	**3**
Big Data Analytics Framework for System Health Monitoring	2015	-	-	-	3	-	4	-	-	**7**
Comparitive study on healthcare prediction systems using big ...	2015	-	-	1	-	-	-	-	-	**1**
Adoption of big data analytics in healthcare: The efficiency ...	2015	-	-	1	-	-	-	-	-	**1**
Digital fitness: Four principles for successful development ...	2015	-	10	19	4	5	-	-	-	**38**
Spatial big data analytics of influenza epidemie in Vellore, ...	2014	-	1	2	4	18	3	-	-	**28**
	Total	**-**	**11**	**28**	**22**	**37**	**26**	**20**	**25**	**169**

Source: own elaboration

References

1. Kumar, S., Singh, M. Big data analytics for healthcare industry: impact, applications, and tools. Big Data Mining and Analytics, **2018**, 2(1), 48-57.
2. Dhar, V. Big Data and Predictive Analytics in Health Care. Big Data, **2014**, 2(3), 113-116.
3. Do Nascimento, I. J. B., Marcolino, M. S., Abdulazeem, H. M., Weerasekara, I., Azzopardi-Muscat, N., Gonçalves, M. A., Novillo-Ortiz, D. Impact of Big Data Analytics on People's Health: Overview of Systematic Reviews and Recommendations for Future Studies. Journal of Medical Internet Research, **2021**, 23(4), e27275.
4. Bhardwaj, N., Wodajo, B., Spano, A., Neal, S., Coustasse, A. The impact of big data on chronic disease management. The Health Care Manager, **2018**, 37(1), 90-98.
5. Ibrahim, S. A., Charlson, M. E.,Neill, D. B. Big Data Analytics and the Struggle for Equity in Health Care: The Promise and Perils. Health Equity, **2020**, 4(1), 99-101.
6. Mariano, D. C., Leite, C., Santos, L. H., Rocha, R. E., de Melo-Minardi, R. C. A guide to performing systematic literature reviews in bioinformatics. arXiv preprint, **2017**, arXiv:1707.05813.
7. Cofre-Martel, S., Droguett, E. L., Modarres, M. Big machinery data pre-processing methodology for data-driven models in prognostics and health management. Sensors, **2021**, 21(20), 6841. doi:10.3390/s21206841
8. Brannigan, R. E. What does the future hold? Emerging diagnostic and therapeutic strategies for male infertility. Fertility and Sterility, **2020**, 113(3), 467-468. doi:10.1016/j.fertnstert.2020.01.039
9. Cui, W., Bachi, K., Hurd, Y., Finkelstein, J. Using Big Data to Predict Outcomes of Opioid Treatment Programs. Studies in Health Technology and Informatics, **2020**, 272, 366-369. doi:10.3233/SHTI200571
10. Chalumporn, G., Hewett, R. Health data analytics with an opportunistic big data algorithm. Paper presented at the PervasiveHealth: Pervasive Computing Technologies for Healthcare, **2020**, doi:10.1145/3406601.3406628
11. Gandomi, A., Haider, M. Beyond the hype: Big data concepts, methods, and analytics. International journal of information management, **2015**, 35(2), 137-144.
12. Ahsan, M., Hon, S. T., Albarbar, A. Development of novel big data analytics framework for smart clothing. IEEE, **2020**, Access, 8, 146376-146394. doi:10.1109/ACCESS.2020.3015152
13. Kan, C., Chen, Y., Leonelli, F., Yang, H. Mobile sensing and network analytics for realizing smart automated systems towards health internet of things. Paper presented at the IEEE International Conference on Automation Science and Engineering, **2015**, October 1072-1077. doi:10.1109/CoASE.2015.7294241
14. Emmanuel, I., & Stanier, C. Defining big data. In Proceedings of the International Conference on Big Data and Advanced Wireless Technologies, **2016**, (pp. 1-6).
15. Cleland, B., Wallace, J., Bond, R., Muuraiskangas, S., Pajula, J., Epelde, G., Arrúe, M., Álvarez, R., Black, M., Mulvenna, M.D., Rankin, D. Usability evaluation of a co-created big data analytics platform for health policy-making. In International Conference on Human-Computer Interaction **2019**, (pp. 194-207). Springer, Cham. doi:10.1007/978-3-030-22660-2_13
16. Kaul, C., Kaul, A., Verma, S. Comparitive study on healthcare prediction systems using big data. Paper presented at the ICIIECS 2015 - 2015 IEEE International Conference on Innovations in Information, Embedded and Communication Systems, **2015**, doi:10.1109/ICIIECS.2015.7193095
17. Yeung, J., McGregor, C. Analyzing countermeasure effectiveness utilizing big data analytics for space medicine decision support: A case study. Paper presented at the Proceedings of the International Astronautical Congress, IAC, **2019**-October

18. Chen, L., Xu, J., Zhang, L., Xue, Y. Big data analytics-based personalized air quality health advisory model. Paper presented at the IEEE International Conference on Automation Science and Engineering, **2017**, August 88-93. doi:10.1109/COASE.2017.8256082
19. Talukder, A. K. Big data analytics advances in health intelligence, public health, and evidence-based precision medicine, **2017**, doi:10.1007/978-3-319-72413-3_17.
20. Gonzalez-Alonso, P., Vilar, R., Lupianez-Villanueva, F. Meeting technology and methodology into health big data analytics scenarios. Paper presented at the Proceedings - IEEE Symposium on Computer-Based Medical Systems, **2017**, June 284-285. doi:10.1109/CBMS.2017.71
21. Dantanarayana, G. G. T., Sahama, T., Wikramanayake, G. N. Quality of information for quality of life: Healthcare big data analytics. Paper presented at the 15th International Conference on Advances in ICT for Emerging Regions, ICTer **2015** - Conference Proceedings, **2016**, 281. doi:10.1109/ICTER.2015.7377714
22. Do, M. H., Söffker, D. State-of-the-art in integrated prognostics and health management control for utility-scale wind turbines. Renewable and Sustainable Energy Reviews, **2021**, 145 doi:10.1016/j.rser.2021.111102
23. Godbole, N. S., Lamb, J. Using data science big data analytics to make healthcare green. Paper presented at the 2015 12th International Conference and Expo on Emerging Technologies for a Smarter World, CEWIT, **2015**, doi:10.1109/CEWIT.2015.7338161
24. Garzón-Alfonso, C. C., Rodríguez-Martínez, M. Twitter health surveillance (THS) system. Paper presented at the Proceedings - 2018 IEEE International Conference on Big Data, Big Data 2018, **2019**, 1647-1654. doi:10.1109/BigData.2018.8622504
25. Molka-Danielsen, J., Engelseth, P., Wang, H. Large-scale integration of wireless sensor network technologies for air quality monitoring at a logistics shipping base. Journal of Industrial Information Integration, **2018**, 10, 20-28. doi:10.1016/j.jii.2018.02.001
26. Ganesh, S., Talukder, A. K. Formal methods, artificial intelligence, big-data analytics, and knowledge engineering in medical care to reduce disease burden and health disparities. In International Conference on Big Data Analytics **2018**, (pp. 307-321). Springer, Cham. doi:10.1007/978-3-030-04780-1_21
27. Grunwell, D., Sahama, T. Information accountability and health big data analytics: A consent-based model. Paper presented at the 2015 17th International Conference on E-Health Networking, Application, and Services, HealthCom **2015**, 195-199. doi:10.1109/HealthCom.2015.7454497
28. Raimundo, R. Rosário, A., Blockchain system in the Higher Education, European Journal of Investigation in Health, Psychology and Education, **2021**, 11(1), 276-293. ISSN 2254-9625, doi.org/10.3390/ejihpe1101002
29. Rosário, A., Moniz, L., Cruz, R. Data Science Applied to Marketing: A Literature Review, Journal of Information Science and Engineering, **2021**, 37(5), 1067-1081, ISSN: 1016-2364. doi.org/10.6688/JISE.202109_37(5).0006
30. Rosário. A. T., Raimundo, R. Consumer Marketing Strategy on E-commerce in the Last Decade, Journal of Theoretical and Applied Electronic Commerce Research, **2021**, 16(7), 3003–3024, ISSN: 0718-1876, doi.org/10.3390/jtaer16070164
31. Rosário. A., Vilaça, F., Raimundo, R., Cruz, R. Literature review on Health Knowledge Management in the last 10 years (2009-2019), The Electronic Journal of Knowledge Management, **2021**, 18(3), 338-355, ISSN 1479-4411. doi.org/10.34190/ejkm.18.3.2120
32. Okoli, C. A guide to conducting a standalone systematic literature review. Communications of the Association for Information Systems, **2015**, 37(43), 879-910.
33. Snyder, H. Literature review as a research methodology: An overview and guidelines. Journal of Business Research, 2019, 104, 333-339.
34. Abusharekh, A., Stewart, S. A., Hashemian, N., & Abidi, S. S. R. H-DRIVE: A big health data analytics platform for evidence-informed decision making. Paper presented at the Proceedings - 2015 IEEE International Congress on Big Data, BigData Congress, **2015**, 416-423. doi:10.1109/BigDataCongress.2015.68.
35. Groves, P., Kayyali, B., Knott, D., Van Kuiken, S. The 'big data' revolution in healthcare: Accelerating value and innovation. McKinsey Quarterly, **2013**, 2(3), 1-22.
36. Kowalski, R. Patients' written reviews as a resource for public healthcare management in England. Paper presented at the Procedia Computer Science, 2017, 113 545-550. doi:10.1016/j.procs.2017.08.275
37. Kuiler, E. W., McNeely, C. L. Knowledge formulation in the health domain: A semiotics-powered approach to data analytics and democratization. Data democracy: At the nexus of artificial intelligence, software development, and knowledge engineering **2020**, (pp. 127-146) doi:10.1016/B978-0-12-818366-3.00007-1
38. Barik, R.K., Dubey, A.C., Tripathi, A., Pratik, T., Sasane, S., Lenka, R.K., Dubey, H., Mankodiya, K. and Kumar, V. Mist data: Leveraging mist computing for secure and scalable architecture for smart and connected health. Paper presented at the Procedia Computer Science, **2018**, 125 647-653. doi:10.1016/j.procs.2017.12.083
39. Jha, A. K., Nayak, S., Veerabhadrappa, N. K. An architecture for performing real-time integrated health monitoring of aircraft systems using avionics big data. Paper presented at the 2nd International Conference on Computational Systems and Information Technology for Sustainable Solutions, CSITSS 2017, **2018**, doi:10.1109/CSITSS.2017.8447679
40. Brennan, P. F., Bakken, S. Nursing needs big data, and big data needs nursing. Journal of Nursing Scholarship, 2015, 47(5), 477-484.

41. McGregor A. M., C., Orlov, O., Baevsky, R., Chernikova, A., Rusanov, V. Big data analytics for continuous assessment of astronaut health risk and its application to human-in-the-loop (HITL) related aerospace. Paper presented at the 19th AIAA Non-Deterministic Approaches Conference, **2017**
42. Ojha, M., Mathur, K. Proposed application of big data analytics in healthcare at Maharaja Yeshwantrao Hospital. Paper presented at the 2016 3rd MEC International Conference on Big Data and Smart City, ICBDSC, **2016**, 40-46. doi:10.1109/ICBDSC.2016.7460340
43. Bouzillé, G., Morival, C., Westerlynck, R., Lemordant, P., Chazard, E., Lecorre, P., Busnel, Y. and Cuggia, M. An automated detection system of drug-drug interactions from electronic patient records using big data analytics. In MedInfo, **2019**, (pp. 45-49). doi:10.3233/SHTI190180
44. McGregor, C., Bonnis, B., Stanfield, B., Stanfield, M. Integrating big data analytics, virtual reality, and ARAIG to support resilience assessment and development in tactical training. Paper presented at the 2017 IEEE 5th International Conference on Serious Games and Applications for Health, SeGAH, **2017**, doi:10.1109/SeGAH.2017.7939256
45. Brdarić, D., Samardžić, S., Huskić, I.M., Dritsakis, G., Sessa, J., Śliwińska-Kowalska, M., Pawlaczyk-Łuszczyńska, M., Basdekis, I. and Spanoudakis, G. A data-informed public health policy-makers platform. International Journal of Environmental Research and Public Health, **2020**, 17(9) doi:10.3390/ijerph17093271
46. Karmawijaya, M. I., Nashirul Haq, I., Leksono, E., Widyotriatmo, A. Development of big data analytics platform for electric vehicle battery management system. Paper presented at the ICEVT 2019 - Proceeding: 6th International Conference on Electric Vehicular Technology **2019**, 151-155. doi:10.1109/ICEVT48285.2019.8994013
47. Palanisamy, V., Thirunavukarasu, R. Implications of big data analytics in developing healthcare frameworks–A review. Journal of King Saud University-Computer and Information Sciences, **2019**, 31(4), 415-425.
48. Buchanan, W. J., Van Deursen, N. HI-risk: A method to analyze health information risk intelligence. Paper presented at the 2016 IEEE 18th International Conference on e-Health Networking, Applications, and Services, Healthcom, **2016**, doi:10.1109/HealthCom.2016.7749536
49. Johnson, O. General system theory and the use of process mining to improve care pathways 2019, doi:10.3233/SHTI190107
50. Dhaka, P., Johari, R. Big data application: Study and archival of mental health data, using MongoDB. Paper presented at the International Conference on Electrical, Electronics, and Optimization Techniques, ICEEOT, **2016**, 3228-3232. doi:10.1109/ICEEOT.2016.7755300
51. Huber, L., Honeder, T., Hackl, W. O. FHIR analytics - pragmatic review of recent studies, 2020, doi:10.3233/SHTI200083
52. Holzmeyer, C. Beyond 'AI for social good' (AI4SG): Social transformations—not tech-fixes—for health equity. Interdisciplinary Science Reviews, **2021**, 46(1-2), 94-125. doi:10.1080/03080188.2020.1840221
53. Howcroft, J., Knoefel, F., Wallace, B., Goubran, R., Porter, M. M., Marshall, S. Impact of health differences and longitudinal changes on deceleration driving patterns in older adult drivers. Transportation Research Part F: Traffic Psychology and Behaviour, **2019**, 60, 137-146. doi:10.1016/j.trf.2018.10.006
54. Chien, T., Chenî, C., Jinî, G., Ting, H. Disease trajectory visualization system based on big data analytics. Paper presented at the ACM International Conference Proceeding Series, **2018**, 73-75. doi:10.1145/3239438.3239485
55. Onyejekwe, E. R. Big data in health informatics architecture. Paper presented at the ASONAM 2014 - Proceedings of the 2014 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining, **2014**, 728-736. doi:10.1109/ASONAM.2014.6921667
56. Khennou, F., Khamlichi, Y. I., Chaoui, N. E. H. Improving the use of big data analytics within electronic health records: A case study based OpenEHR. Paper presented at the Procedia Computer Science, **2018**, 127 60-68. doi:10.1016/j.procs.2018.01.098
57. Hansen, K. S. Socially shared health information. Paper presented at the ACM International Conference Proceeding Series, **2018**, April 152-153. doi:10.1145/3194658.3194688
58. Hansen, K. S., Mukkamala, R., Hussain, A., Grønli, T., Langberg, H., Vatrapu, R. Big social data in public health: A mixed-methods case study of sundhed. DK's Facebook strategy, engagement, and performance. Paper presented at the Procedia Computer Science, **2016**, 58 298-307. doi:10.1016/j.procs.2016.09.046
59. Ndangang, M., Grosjean, J., Lelong, R., Dahamna, B., Kergourlay, I., Griffon, N., Darmoni, S. J. Terminology coverage from semantic annotated health documents, **2018**, doi:10.3233/978-1-61499-921-8-20
60. Lakshmanaprabu, S. K., Shankar, K., Ilayaraja, M., Nasir, A. W., Vijayakumar, V., Chilamkurti, N. Random forest for big data classification in the internet of things using optimal features. International Journal of Machine Learning and Cybernetics, **2019**, 10(10), 2609-2618. doi:10.1007/s13042-018-00916-z
61. Ho, K., Hirai, H. W., Kuo, Y., Meng, H. M., Tsoi, K. K. F. Indoor air monitoring platform and personal health reporting system: Big data analytics for public health research. Paper presented at the Proceedings - 2015 IEEE International Congress on Big Data, BigData Congress **2015**, 309-312. doi:10.1109/BigDataCongress.2015.51
62. Nayak, B., Bhattacharyya, S. S., Krishnamoorthy, B. Integrating wearable technology products and big data analytics in business strategy: A study of health insurance firms. Journal of Systems and Information Technology, **2019**, 21(2), 255-275. doi:10.1108/JSIT-08-2018-0109
63. Khoury, G., Figueiredo, M. Improved decision-making, safety, and reliability with visual asset performance management. Paper presented at the Society of Petroleum Engineers - Abu Dhabi International Petroleum Exhibition and Conference 2020, ADIP, **2020**

64. Chan, N. K., Lee, A. S., Zainol, Z. Predicting employee health risks using classification ensemble model. Paper presented at the Proceedings - CAMP 2021: 2021 5th International Conference on Information Retrieval and Knowledge Management: Digital Technology for IR 4.0 and Beyond, **2021**, 52-58. doi:10.1109/CAMP51653.2021.9498106
65. Nafea, I. Mobile health application running on public cloud during hajj, 2017, doi:10.1007/978-3-319-65515-4_11
66. Li, H., Wu, J., Liu, L., Li, Q. Adoption of big data analytics in healthcare: The efficiency and privacy. Paper presented at the Pacific Asia Conference on Information Systems, PACIS 2015 – Proceedings
67. Shetty, V., Yamamoto, J., Yale, K. Re-architecting oral healthcare for the 21st century. Journal of Dentistry, **2018**, 74, S10-S14. doi:10.1016/j.jdent.2018.04.017
68. Alexandru, A., Coardos, D., Tudora, E. Iot-based healthcare remote monitoring platform for elderly with fog and cloud computing. Paper presented at the Proceedings - 2019 22nd International Conference on Control Systems and Computer Science, CSCS, **2019**, 154-161. doi:10.1109/CSCS.2019.00034
69. Plageras, A. P., Stergiou, C., Kokkonis, G., Psannis, K. E., Ishibashi, Y., Kim, B., Gupta, B. B. Efficient large-scale medical data (eHealth big data) analytics in the internet of things. Paper presented at the Proceedings - 2017 IEEE 19th Conference on Business Informatics, CBI, **2017**, 2 21-27. doi:10.1109/CBI.2017.3
70. Shetty, V., Yamaguchi, M. Point-of-care testing in the time of P5 medicine: A preface. Sensors and Materials, **2016**, 28(10), 1141-1148. doi:10.18494/SAM.2016.1275
71. Setiawan, A. W., Utami, N., Mengko, T. R., Indrayanto, A. (2014). Implementation of electronic medical records in community health center towards medical big data analytics application. Paper presented at the Proceedings of 2014 International Conference on Electrical Engineering and Computer Science, ICEECS, 2014, 30-35. doi:10.1109/ICEECS.2014.7045215
72. Wu, P., Cheng, C., Kaddi, C. D., Venugopalan, J., Hoffman, R., Wang, M. D. -Omic and electronic health record big data analytics for precision medicine. IEEE Transactions on Biomedical Engineering, **2017**, 64(2), 263-273. doi:10.1109/TBME.2016.2573285
73. Rao, S., Suma, S. N., Sunitha, M. Security solutions for big data analytics in healthcare. Paper presented at the Proceedings - 2015 2nd IEEE International Conference on Advances in Computing and Communication Engineering, ICACCE, **2015**, 510-514. doi:10.1109/ICACCE.2015.83
74. Xiao, F., Cui, Y., Wang, W., Zhang, C., He, X., Gu, J., Cao, L. Online status assessment o distribution network transformer-based o random matrix theory. Paper presented at the 2020 IEEE 4th Conference on Energy Internet and Energy System Integration: Connecting the Grids Towards a Low-Carbon High-Efficiency Energy System, EI2, **2020**, 4227-4231. doi:10.1109/EI250167.2020.9346854
75. Sheng, S., Guo, Y. A prognostics and health management framework for wind. Paper presented at the Proceedings of the ASME Turbo Expo, **2019**, 9 doi:10.1115/GT2019-91533
76. White, S. E. A review of big data in health care: challenges and opportunities. Open Access Bioinformatics, **2014**, 6, 13-18.
77. Pathak, S., Krishnaswamy, V., Sharma, M. Big data analytics capabilities: A novel integrated fitness framework based on tool-based content analysis. Enterprise Information Systems, **2021**, doi:10.1080/17517575.2021.1939427
78. Molka-Danielsen, J., Engelseth, P., Olesnanikova, V., Sarafin, P., Zalman, R. Big data analytics for air quality monitoring at a logistics shipping base via autonomous wireless sensor network technologies. Paper presented at the Proceedings - 2017 5th International Conference on Enterprise Systems: Industrial Digitalization by Enterprise Systems, ES **2017**, 38-45. doi:10.1109/ES.2017.14
79. Xu, B., Kumar, S. A. Big data analytics framework for system health monitoring. Paper presented at the Proceedings - 2015 IEEE International Congress on Big Data, BigData Congress **2015**, 401-408. doi:10.1109/BigDataCongress.2015.66
80. Xu, Z., Wei, X., Luo, X., Liu, Y., Mei, L., Hu, C., Chen, L. Knowle: A semantic link network-based system for organizing large-scale online news events. Future Generation Computer Systems, 2015, 43-44, 40-50. doi:10.1016/j.future.2014.04.002
81. Alam, M., Amjad, M. Analysis of climate change and its impact on health using big data analytics in cloud environment. IETE Journal of Research, **2021**, doi:10.1080/03772063.2021.1886605
82. Silvestri, S., Esposito, A., Gargiulo, F., Sicuranza, M., Ciampi, M., De Pietro, G. A big data architecture for the extraction and analysis of EHR data. Paper presented at the Proceedings - 2019 IEEE World Congress on Services, SERVICES, **2019**, 283-288. doi:10.1109/SERVICES.2019.00082
83. Chen, M., Ma, Y., Song, J., Lai, C., Hu, B. Smart clothing: Connecting humans with clouds and big data for sustainable health monitoring. Mobile Networks and Applications, 2016, 21(5), 825-845. doi:10.1007/s11036-016-0745-1
84. Liyanage, H., Williams, J., Byford, R., Stergioulas, L., De Lusignan, S. Ontologies in big health data analytics: Application to routine clinical data, **2018**, doi:10.3233/978-1-61499-921-8-65
85. Lopez, D., Gunasekaran, M., Murugan, B. S., Kaur, H., Abbas, K. M. Spatial big data analytics of influenza epidemic in Vellore, India. Paper presented at the Proceedings - 2014 IEEE International Conference on Big Data, IEEE Big Data, **2014**, 19-24. doi:10.1109/BigData.2014.7004422
86. Kuo, A., Chrimes, D., Qin, P., Zamani, H. A Hadoop/MapReduce based platform for supporting health big data analytics 2019, doi:10.3233/978-1-61499-951-5-229
87. Mylaraswamy, D., Xu, B., Dietrich, P., Murugan, A. Case studies: Big data analytics for system health monitoring. Paper presented at the Proceedings of the 2014 International Conference on Artificial Intelligence, ICAI 2014 - WORLDCOMP **2014**, 26-30.

88. Anjum, A., Ahmed, T., Khan, A., Ahmad, N., Ahmad, M., Asif, M., Reddy, A.G., Saba, T., and Farooq, N. Privacy-preserving data by conceptualizing smart cities using MIDR-Angelization. Sustainable Cities and Society, **2018**, 40, 326-334. doi:10.1016/j.scs.2018.04.014
89. Melville, N. P. (2015). Digital fitness: Four principles for successful development of digital initiatives. Paper presented at the Proceedings of the Annual Hawaii International Conference on System Sciences, **2015**, March 4387-4396. doi:10.1109/HICSS.2015.526
90. Plageras, A. P., Psannis, K. E. Algorithms for big data delivery over the internet of things. Paper presented at the Proceedings - 2017 IEEE 19th Conference on Business Informatics, CBI, **2017**, 1 202-206. doi:10.1109/CBI.2017.27.