Online-Journals.org

iJOE International Journal of Online and Biomedical Engineering

iJOE | eISSN: 2626-8493 | Vol. 19 No. 16 (2023) | OPEN ACCESS

https://doi.org/10.3991/ijoe.v19i16.45035

PAPER

Big Data Analytics for Early Detection and Prevention of Age-Related Diseases in Elderly Healthcare

Mohd Amran Mohd Daril[1](✉), Shazia Qayuum[2], Alhamzah F. Abbas[3], Nguyen Thuy Van[4]

[1]Quality Engineering Research Cluster, Quality Engineering Section, Malaysian Institute of Industrial Technology, Universiti Kuala Lumpur, Johor Bahru, Johor, Malaysia

[2]Institute of Applied Psychology, University of Punjab, Lahore, Pakistan

[3]Faculty of Management, Universiti Teknologi Malaysia, Johor Bahru, Malaysia

[4]Journal of Science Department, Phu Yen University, Tuy Hoa City, Vietnam

mamran@unikl.edu.my

ABSTRACT

The exponential growth of the elderly population poses considerable obstacles to healthcare systems on a global scale, hence requiring the implementation of inventive strategies to identify and mitigate age-related illnesses at an early stage. The primary objective of this study is to explore the use of big data analytics to improve healthcare practices. Specifically, the emphasis is on identifying possible risk factors and developing proactive treatments for senior citizens. The research technique used in this study is based on the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) declaration of 2020. This approach is utilised to ensure a thorough and transparent review of the relevant literature. Moreover, the use of Rstudio software is prevalent in the field of data processing, statistical analysis, and visualisation. By conducting a comprehensive examination of academic databases and medical literature, this study undertakes an analysis of a collection of pertinent papers to explore the significance of big data analytics in the early diagnosis and prevention of diseases in senior populations. The studies that have been chosen include a wide range of healthcare fields, such as cardiology, neurology, cancer, and geriatrics. This selection aims to provide a thorough comprehension of existing practises and identify any possible areas that may need more attention. The results of this study emphasise the significant impact that big data analytics may have on healthcare for the elderly. Using extensive and varied datasets, sophisticated analytical methodologies such as machine learning algorithms and data mining allow the detection of nuanced patterns and correlations that might function as precursors for age-related ailments.

KEYWORDS

Age-Related Diseases, Elderly Health, big data analytics disease prevention, PRISMA2020, RStudio software

1 INTRODUCTION

In recent decades, the integration of healthcare and technology has initiated a paradigm shift characterised by evidence-based observations and inventive

Mohd Daril, M.A., Qayuum, S., Abbas, A.F., Thuy Van, N. (2023). Big Data Analytics for Early Detection and Prevention of Age-Related Diseases in Elderly Healthcare. *International Journal of Online and Biomedical Engineering (iJOE)*, 19(16), pp. 20–37. https://doi.org/10.3991/ijoe.v19i16.45035

Article submitted 2023-08-16. Revision uploaded 2023-09-23. Final acceptance 2023-09-27.

strategies to tackle intricate medical issues [1]. In addition, within the realm of issues faced by society, the prompt identification and proactive mitigation of age-related ailments within the older demographic emerge as important endeavours that are crucial for safeguarding the overall welfare of our ageing populace [2]. According to [3], the incorporation of big data analytics in the field of geriatric healthcare offers a potential opportunity to transform the comprehension, diagnosis, and eventual mitigation of these ailments. However, the existing corpus of scholarly research provides substantial evidence supporting the pressing need for sophisticated approaches in addressing the healthcare requirements of the aged population, considering the escalating impact of age-related illnesses on healthcare systems around the globe [4]. Also, several studies have emphasised the potential of big data analytics in effectively using the vast amount of health-related data and deriving significant insights that might inform personalised healthcare treatments [5]. The amalgamation of technological advancements and medical knowledge presents the potential for proactive healthcare, whereby illnesses may be detected in their nascent phases, allowing for prompt therapies that hinder or alleviate their development [6].

According to [7], the increasing incidence of age-related illnesses is a matter of growing worldwide concern, mostly attributed to the demographic shift towards an ageing population. The ageing population is experiencing growth due to the rise in life expectancy, which serves as evidence of progress in healthcare and living standards [8]. Nevertheless, this change in demographics has concurrently resulted in an increased need for healthcare services and a necessity for enhanced approaches to avoid, identify at an early stage, and treat these ailments more efficiently [9]. In addition, the identification of age-related disorders at an early stage poses a significant difficulty within the realm of aged healthcare [10]. Numerous instances of these disorders manifest inconspicuously, exhibiting minor symptoms that might potentially be misconstrued as typical signs of ageing or less serious medical ailments [11]. As a result, individuals often remain undetected until their condition has evolved to later stages, therefore diminishing the efficacy of therapy and perhaps imposing constraints on the patient's longevity or overall well-being [12].

However, the use of big data analytics into the field of senior healthcare has been a notable and encouraging development in recent times [13]. Using data derived from diverse sources such as electronic health records, wearable devices, and genetic information, healthcare practitioners may acquire more profound understandings pertaining to the risk factors, early indicators, and viable therapies associated with age-related ailments [14]. The use of data-driven methodologies facilitates the creation of personalised healthcare strategies that are specifically designed to cater to the distinctive health profile of everyone [15].

Furthermore, the use of big data analytics has the potential to enhance predictive modelling capabilities within the healthcare sector [16]. This enables healthcare professionals to anticipate the beginning or development of diseases by using a patient's health history and other pertinent data [17]. The use of a proactive healthcare strategy enables patients and healthcare practitioners to engage in preventative measures, so possibly preventing the onset of age-related illnesses or reducing their severity [18].

Research objective is to provide a foundation for well-informed plans that facilitate progress in the early diagnosis and prevention of diseases in the aged healthcare sector. The objective of this study is to provide healthcare practitioners, academics, policymakers, and stakeholders with a comprehensive comprehension of the fundamental themes and patterns that are influencing this nascent discipline. By using a methodology grounded in empirical evidence, our objective is to cultivate cooperation and ingenuity, capitalising on the powerful amalgamation of extensive data analysis and specialised knowledge in geriatric healthcare. This endeavour aims to provide a framework for the advancement of early illness identification and prevention, paving the way for future developments in this field.

2 MATERIALS AND METHODS

The current study employed the PRISMA statement 2020 to include and exclude records from Scopus databases to incorporate high-quality materials The data were screened using the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) methodology, as recommended by [19], [20] and shown in Figure 1. For improved reporting of records and pertinent information related to the literature, the present research employed the PRISMA statement 2020. For our review of the literature, we utilised the search phrases “Big Data Analytics” AND “Elderly Healthcare” AND “Diseases”. Initially, 97 records were collected. The present evaluation included works from the fields of engineering, business management, computer science, health sciences, and transdisciplinary studies. Therefore, just 81 documents remain in the results. Additionally, we limited our selection for the present research to articles, reviews, and book chapters, bringing the total number of records down to 73. Additionally, to mimic the study’s scope for significant literature results, only published and English-language papers were considered. A total of 66 records remained after this phase. The next step was to eliminate duplicate document information that was superfluous or missing. For each discovered category, a thorough selection was made to examine related content. Only 62 papers remained, which allowed for an easy synthesis. The PRISMA statement selection and rejection method employed in the present research is shown in Figure 1.

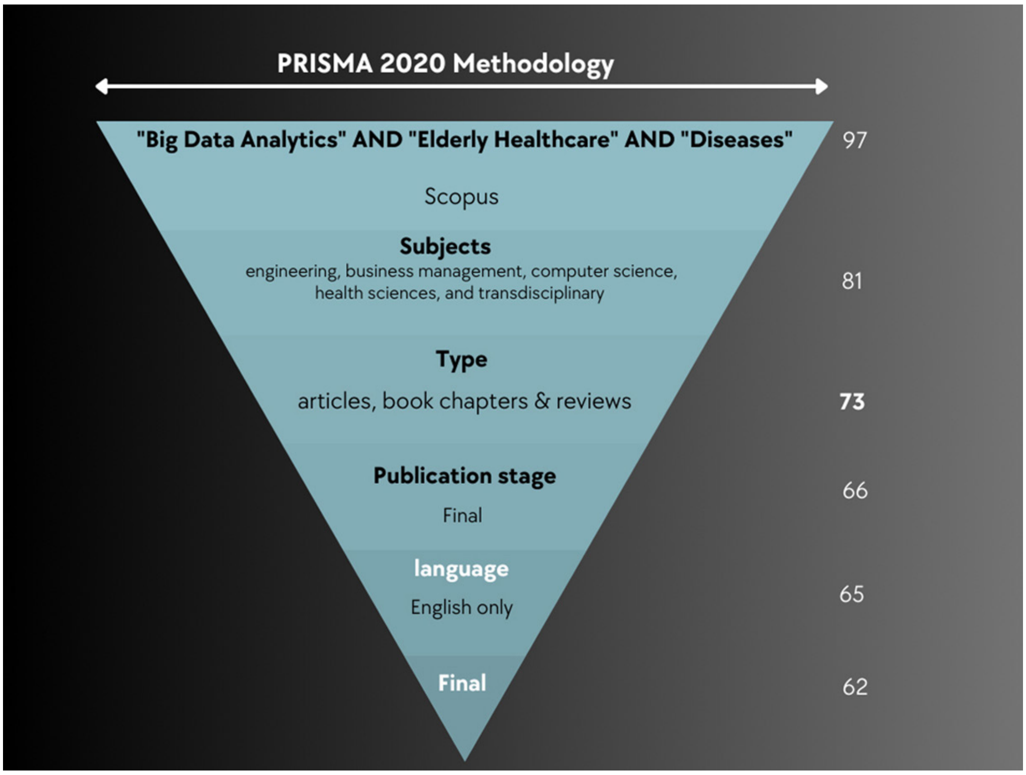

Fig. 1. The PRISMA methodology

Table 1 shows insights into many facets of the dataset's properties and content by outlining pertinent information about the dataset under review. The data's time frame is from 2013 to 2023, covering a significant amount of time and perhaps giving a thorough picture of changes through time. 52 references from journals, books, and other relevant sources make up the sources of the data, suggesting a wide variety of references that contributed to the dataset. There are 62 papers in all, representing an amazing yearly growth rate of 11.61%, which indicates a consistent intake of new documents every year.

The dataset's average document age is 3.98 years, which may indicate that relatively recent elements make up the majority of the data. Each paper obtains 24 citations on average, which reflects the importance and relevance of the texts among academics. The dataset's material was derived from a wide range of sources, as seen by the considerable number of references (2137 in total) that further substantiate these citations. By examining the text of the papers, it is possible to find 814 occurrences of Keywords Plus (ID) and 199 instances of Author's Keywords (DE), which provide important details on the subjects and themes covered by the dataset. 251 writers contributed to the dataset, and 5 of the authors wrote all the documents by themselves. This distribution shows that the dataset contains a mixture of single-author and group projects. The cooperation patterns show a significant preference for collaborative research with 5 single-authored papers and an average of 4.23 co-authors per document. International co-authorships make up a significant portion of these co-authorships (27.342%), highlighting the dataset's cross-national nature of cooperation. The dataset consists of a variety of document categories, with articles being the most frequent (29), followed by book chapters (7), conference papers (21), conference reviews (3), and reviews (2). This variety of document formats points to an inclusive collection that includes many intellectual outputs.

Table 1. Main information about the data

Description	Results
MAIN INFORMATION ABOUT DATA	
Timespan	2013:2023
Sources (Journals, Books, etc)	52
Documents	62
Annual Growth Rate %	11.61
Document Average Age	3.98
Average citations per doc	24
References	2137
DOCUMENT CONTENTS	
Keywords Plus (ID)	814
Author's Keywords (DE)	199
AUTHORS	
Authors	251
Authors of single-authored docs	5

(Continued)

Table 1. Main information about the data *(Continued)*

Description	Results
AUTHORS COLLABORATION	
Single-authored docs	5
Co-Authors per Doc	4.23
International co-authorships %	27.42
DOCUMENT TYPES	
Article	29
book chapter	7
conference paper	21
conference review	3
Review	2

In addition, Figure 2 provides a comprehensive analysis of the yearly article output within the dataset, illustrating the dispersion of articles throughout several years. The provided data presents the progression of article generation from 2013 to 2023, facilitating the examination of trends and patterns in scholarly creation. In the year 2013, a solitary paper was generated, suggesting the prospective commencement of data gathering or the concentration on a certain study domain. In the following year, namely 2014, no papers were seen, maybe attributable to fluctuations in research effort or limitations in data accessibility during that specific timeframe.

In the year 2015, there was a modest increase in research production, as shown by the publication of three publications, suggesting a potential revival in scholarly activity. The following years saw a sustained expansion, exemplified by a significant rise in 2016 with the publication of six pieces. The observed pattern remained consistent throughout the year 2017, with a total of eight scholarly publications published, suggesting a continuous increase in academic contributions. In the ensuing years, namely 2018 and 2019, there was a steady level of productivity seen, with 6 and 12 publications published, respectively. This time may indicate a phase characterized by continuous involvement in research and growth.

Furthermore, in the year 2020, a total of seven publications were published, suggesting that research production remained consistent despite possible interruptions arising from external circumstances, such as the worldwide COVID-19 pandemic. The ensuing years, namely 2021 and 2022, exhibited a continuation of the tendency, with 8 publications published each year. This observation suggests a sustained and constant level of academic engagement. In the present year, 2023, there has been a decrease in article production to a count of 3. This decline might possibly be attributed to several causes that would have influenced research output, including ongoing world events or changes in research goals. Table 2 presents a concise overview of the yearly article output within the dataset, effectively illustrating the fluctuating patterns of research effort across the years. The fluctuations in the number of articles published in various years may provide insights about shifting research patterns, the development of new disciplines, or the effect of external factors on academic output.

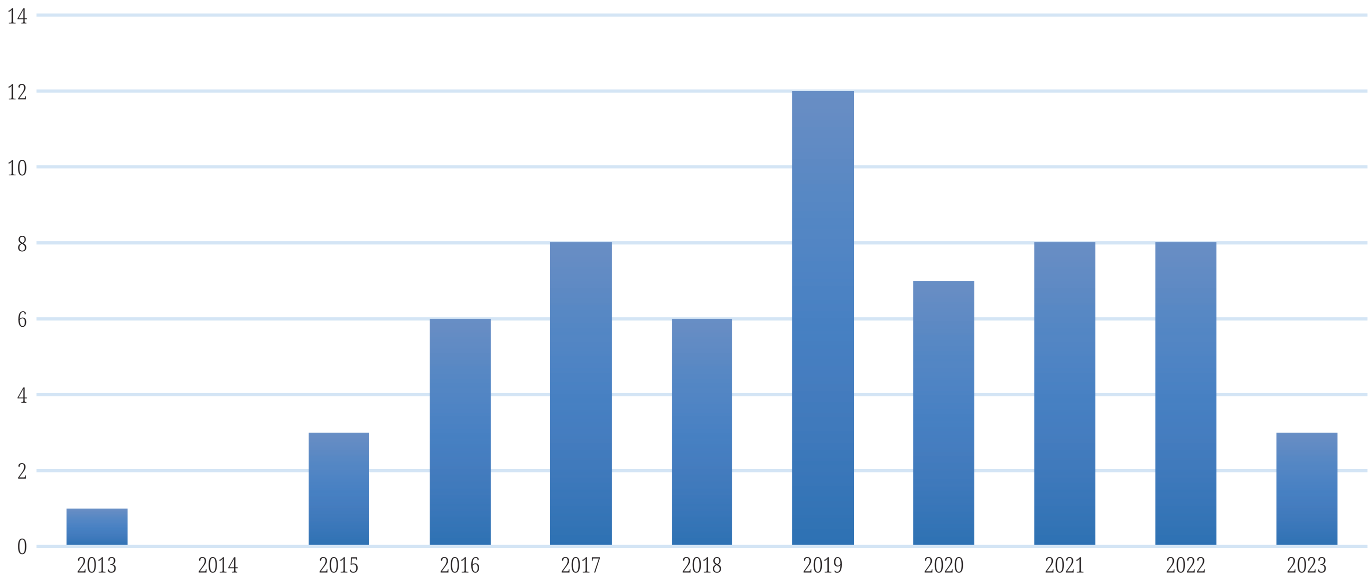

Fig. 2. Annual production of articles (2013–2023)

Furthermore, Table 2 under "Sources of Articles" shows a comprehensive compilation of sources, along with the respective count of articles coming from each source. The provided data provides valuable insights on the distribution of papers throughout various academic journals and conference proceedings, hence illuminating the variety and popularity of platforms within the discipline.

The sources included in the table exhibit a variety of scholarly journals, conference proceedings, and other publications through which the papers included in the dataset have been disseminated. Significantly, the datasets of "Advances in Intelligent Systems and Computing" and "International Journal of Environmental Research and Public Health" are distinguished by the presence of three articles each, underscoring their noteworthy contributions. It is probable that these sites mostly concentrate on subjects pertaining to intelligent systems and environmental research, correspondingly.

Multiple sources are closely associated with two articles apiece, indicating an equitable representation from diverse platforms. Some examples of academic journals and conference proceedings in the field of computer science are "Communications in Computer and Information Science," "IEEE Access," "Lecture Notes in Computer Science," "Lecture Notes in Networks and Systems," among others. The observation that different sources have generated a similar quantity of publications indicates a broad dissemination of research endeavours across several disciplines and publishing platforms.

At a lower position in the list, each unique source contributes of one article to the dataset. The presence of diverse sources in the collection indicates that it encompasses a wide range of sectors, including emergency medicine, IoT-enabled healthcare systems, air quality and atmosphere research, ophthalmology, applied computing in medicine and health, and big data. The display of the table has significant value for scholars who want to comprehend the scope of articles within this field. The analysis discloses the sources that have shown higher productivity in terms of article production and offers an overview of the wide range of subjects addressed by these sources. This information has the potential to assist academics in discovering prominent journals, conferences, and publications that may be further examined and analysed.

Table 2. Sources of articles

Sources	Articles
Advances In Intelligent Systems and Computing	3
International Journal of Environmental Research and Public Health	3
Communications In Computer and Information Science	2
IEEE Access	2
Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)	2
Lecture Notes in Networks and Systems	2
Lecture Notes of The Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering, LNICST	2
Proceedings – 2019 22nd International Conference on Control Systems and Computer Science, CSCS 2019	2
Academic Emergency Medicine	1
Ai-Enabled Iot for Smart Health Care Systems	1
Air Quality, Atmosphere and Health	1
American Journal of Ophthalmology	1
Applied Computing in Medicine and Health	1
Big Data	1

3 DESCRIPTIVE STATISTICS

3.1 Results

The descriptive-analytic study in this bibliometric inquiry used the Bibliometrix RStudio-4.2.1-win programme. The R package known as Bibliometrics is seeing a growing prevalence in scholarly publications. Bibliometrics enables users of the R programming language to import a database of bibliographic information from Scopus. In accordance with a rational bibliometric methodology, a specialised tool known as Bibliometric was developed, using statistical computing techniques and the R programming language. System R is often acknowledged as the prevailing standard platform for the development of statistical algorithms. It is extensively used as a dynamic software tool for the purpose of data analysis and visualisation, as stated in the research article titled "View of Mapping Research on Using Biblioshiny" [21]. The analysis of the data takes the form of the publishing year, author, publisher, organisation, nation, and keywords utilised in the research paper.

Table 3 of "Annual Citations per Year in Academic Writing" presents a complete analysis of the mean yearly citations for each year within the specified period. The provided data provides valuable insights into the academic impact and exposure of works over a period, together with the frequency of citations and the durability of their influence. The column labelled "Mean TC per Art" denotes the mean value of total citations per article during the given year. This statistical measure represents the mean number of citations obtained by each article published in that particular year. The column labelled "N" represents the quantity of papers published within a certain year, so indicating the magnitude of academic production.

The column labelled "Mean TC per Year" represents the mean number of citations per year for all papers published within a given year. This measure offers a more comprehensive perspective on the citation effect of the full year's production. Furthermore, the column labelled "Citable Years" displays the duration for which publications from a certain year may be cited. This observation demonstrates the enduring significance of the published research over its lifespan. Upon analysis of the data, several distinct patterns become apparent. In the year 2013, despite the limited number of articles published, one paper garnered a noteworthy average of 21 citations, suggesting the potential for significant influence within the academic community. In the year 2015, there was a decrease in the average number of citations per article, which reached a value of 14. However, despite this decline, the overall number of citations per year remained constant at 1.56, considering the presence of three articles. In the year 2016, there was a noteworthy increase in the average number of citations per article, reaching a value of 65.17. This observation indicates that a certain set of publications published during that year garnered considerable attention. The pattern persisted throughout the year 2018, with an average of 61 citations per article.

On the other hand, it is noteworthy that the years 2017 and 2022 saw a decrease in the average number of citations per article, suggesting a possible decline in the significance of the published works. The dataset "Citable Years" exhibits diverse degrees of lifespan with respect to the influence of citations. As an example, scholarly publications published in the year 2016 exhibited an average duration of cutability amounting to 8.15 years, but articles published in 2019 showed an average cutability period of 3.52 years. In the year 2023, a total of three articles were written. However, it is noteworthy that these publications have not yet received any citations, as seen by the mean citations per article being 0 and the lack of a "Citable Years" number.

Table 3. Annual citations per year

Year	Mean TC per Art	N	Mean TC per Year	Citable Years
2013	21	1.00	1.91	11
2015	14	3.00	1.56	9
2016	65.17	6.00	8.15	8
2017	9.38	8.00	1.34	7
2018	61	6.00	10.17	6
2019	17.58	12.00	3.52	5
2020	16	7.00	4.00	4
2021	32.38	8.00	10.79	3
2022	1.38	8.00	0.69	2
2023	0	3.00	0.00	1

In addition, Table 3 entitled Most Cited Documents presents a comprehensive collection of documents, accompanied by their corresponding DOIs, publication years, worldwide citation counts, and normalised global citations. The provided data provides valuable insights on the influence and acknowledgment that certain texts have received within the academic community, as well as their significance across diverse disciplines. The entries within the document exhibit a diverse selection of research articles and conference proceedings that have received significant recognition within the scholarly community. Every item in the dataset contains the title of the publication, its DOI (Digital Object Identifier), the year of publication, the total number of citations it has received globally, and a normalised score representing its worldwide citation impact. The normalised global citation score is a metric that considers the citation count of a document in relation to its publication year, therefore offering a quantification of its long-term influence.

Several papers have garnered significant attention and recognition, as seen by their large citation counts. An example of a very influential publication within its area is the research article authored by JIANG P, which was published in the IEEE Systems Journal in 2016. This study has garnered a total of 86 citations, highlighting its significant impact in the academic community. The scholarly article authored by Hossain MS, published in the IEEE Internet of Things Journal, has garnered a notable number of 187 citations, indicating its substantial influence within the respective domain. Certain scholarly works, such as those authored by Rastogi R, Al Nahian MJ, and Hussain I, have garnered significant citation counts and shown relatively high normalised citation scores. This indicates their ongoing relevance and impact within the academic community over an extended period. Conversely, several texts have garnered a restricted number of citations, maybe attributable to their more specialised or specialist subject matter. The research conducted by GAO Y in the "Advances in Intelligent Systems and Computing" series in 2020 did not earn any citations at the time of data collection (see Table 4).

Table 4. Most cited documents

Document	DOI	Global Citations	Normalized Global Citations
Jiang et al., 2016 [22]	10.1109/JSYST.2014.2308324	86	1.32
Rastogi et al., 2022 [23]	10.1007/978-3-030-77528-5_9	3	2.18
Kumar & Bagavathi, 2016 [24]	10.1109/ICCCI.2016.7479952	2	0.03
Gao & Li, 2020 [25]	10.1007/978-981-15-2568-1_7	0	0.00
Danial-Saad et al., 2022 [26]	10.1007/978-3-030-10752-9_5	3	0.17
Hossain & Muhammad, 2018 [27]	10.1109/JIOT.2017.2772959	187	3.07
Li et al., 2022 [28]	10.1007/s11136-019-02132-w	34	1.93
Bajenaru & Custura, 2019 [29]	10.1109/CSCS.2019.00116	2	0.11
Anya & Tawfik, 2016 [30]	10.1016/B978-0-12-803468-2.00005-9	3	0.21
Tsoi et al., 2017 [31]	10.1109/CCBD.2016.068	2	0.21
Rocha et al., 2019 [32]	10.1007/978-3-030-01746-0_33	0	0.00
Willetts et al., 2022 [33]	10.1016/B978-0-323-85173-2.00001-1	2	1.45
Qaffas et al., 2021 [34]	10.1089/tmj.2019.0289	15	0.46
Rezaee et al., 2022 [35]	10.1142/S0218126622400059	2	1.45

(Continued)

Table 4. Most cited documents *(Continued)*

Document	DOI	Global Citations	Normalized Global Citations
Baxter et al., 2021 [36]	10.1016/j.ajo.2021.01.008	17	0.53
Kang et al., 2019 [37]	10.1007/s10072-019-3730-1	23	1.31
Yu et al., 2019 [38]	10.1109/PlatCon.2019.8668961	8	0.45
High et al., 2019 [39]	10.1111/jgs.15975	13	0.74
Yacchirema et al., 2018 [40]	10.1016/j.pmcj.2018.07.007	37	0.61
Goel & Vishnoi, 2023 [41]	10.25103/jestr.162.02	0	
Roshni Thanka et al., 2022 [42]	10.1007/978-981-19-2177-3_55	0	0.00
Hussain & Park, 2021 [43]	10.1109/ACCESS.2021.3109806	54	1.67
Al Nahian et al., 2020 [44]	10.1007/978-3-030-59277-6_25	35	2.19

Furthermore, Figure 3 depicts the use of Bradford's Law in the context of academic writing, especially with the allocation of journals throughout several zones according to their cumulative frequency of publications. The Bradford's Law, formulated by Samuel C. Bradford, posits that within some academic disciplines, there exists a discernible trend of concentration in scholarly writing, wherein a limited number of journals are responsible for a significant proportion of published papers. In addition, the presented image illustrates the distribution of journals categorised by their cumulative frequency (cum Freq) and zones. The column labelled "SO" indicates the source or journal from which the information is derived, whereas the column labelled "Rank" indicates the publication's rank in terms of frequency. The column labelled "Freq" denotes the quantity of articles that have been published in the respective journal, while the column labelled "cum Freq" signifies the cumulative frequency up to the given moment. The column labelled "Zone" classifies journals into distinct zones based on the principles of Bradford's Law.

The top-ranking journals in "Zone 1" are shown in descending order based on their cumulative frequency, indicating their high level of productivity. These academic journals together provide a substantial contribution to the body of published publications. Two notable contributors to academic literature within the dataset are "Advances in Intelligent Systems and Computing" and "International Journal of Environmental Research and Public Health." As the transition to "Zone 2" occurs, there is a noticeable deceleration in the pace at which the cumulative frequency rises. The frequency of contributions from journals within this zone is somewhat lower when compared to those inside "Zone 1." Some examples of academic publications in the field of health care and related topics include "AI-Enabled Internet of Things (IoT) for Smart Health Care Systems," "Air Quality, Atmosphere, and Health," and "The American Journal of Ophthalmology."

In the third zone, the cumulative frequency curve exhibits a gradual decrease in slope since the inclusion of each subsequent journal contributes only a marginal increase to the total frequency. The publishing rates of these journals, while they provide valuable contributions and remain relevant, are comparatively lower when compared to those in "Zone 1" and "Zone 2." The examples under this category span a wide array of disciplines, including the "Journal of the American

Geriatrics Society," the "Journal of the American Medical Informatics Association," and "Telemedicine and E-Health." The use of Bradford's Law in this Figure exemplifies the notion of concentration and dispersion of scholarly output across different zones, facilitating researchers and stakeholders in comprehending the extent to which individual journals contribute to most of the academic writing in the dataset.

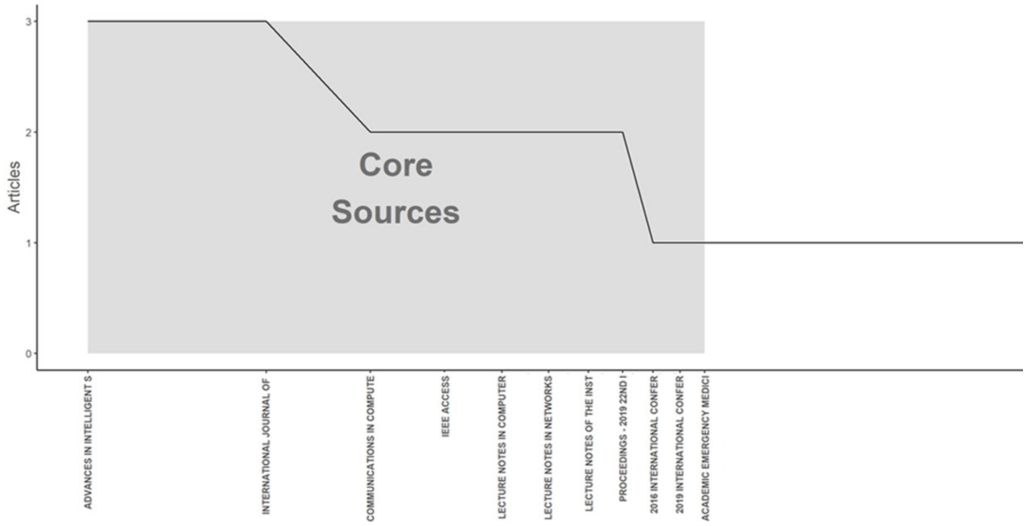

Fig. 3. Core sources by Bradford law

Moreover, Table 5 provided data provides valuable insights into the patterns of repeating themes, subjects, and vocabulary that have substantial significance within academic discourse. Table 5 presents a comprehensive compilation of terms, along by their respective frequencies throughout the dataset. The prevalence of these phrases illuminates the primary areas of emphasis and patterns within the discipline. The word "big data" is of particular significance, since it appears 36 times, suggesting a considerable level of interest in studies pertaining to data. Likewise, it is noteworthy that phrases such as "aged" (occurring 32 times), "female" (occurring 30 times), and "male" (occurring 28 times) indicate a discernible emphasis on demographic attributes within the context of diverse research endeavours. In addition, the frequency of the terms "data analytics" (22 instances) and "data handling" (8 instances) highlights the significance of proficient methodologies for managing and analysing data. The frequency of terms such as "adult" (20 instances) and "middle aged" (19 instances) in the literature indicates a deliberate focus on studying certain age cohorts.

The phrase "health care" (occurring 18 times) corresponds to the importance of research connected to healthcare. The frequency of the term "very elderly" (16 instances) and "aged 80 and over" (12 instances) indicates a notable emphasis on the demographic of older individuals. The frequent use of phrases such as "internet of things" (12 instances) and "machine learning" (11 instances) indicates the incorporation of contemporary technology into scholarly discussions. The Table 5 also includes phrases pertaining to research methods and design, such as "controlled study" (8 instances) and "major clinical study" (10 instances). Furthermore, the use of location-specific terminology such as "Taiwan" (occurring 9 times) offers valuable insights on the geographical setting of the study.

Table 5. Most frequent words

Words	Occurrences
big data	36
Aged	32
Female	30
Male	28
data analytics	22
Adult	20
Human	19
Humans	19
middle aged	19
health care	18
very elderly	16
Article	13
aged 80 and over	12
internet of things	12
machine learning	11
electronic health record	10
major clinical study	10
advanced analytics	9
controlled study	8
data handling	8
Adolescent	7
data mining	7
Diagnosis	7
electronic health records	7

Moreover, Table 6 provides a quantitative evaluation of the associations, significance, and impact of terms within the dataset. This analysis emphasises fundamental ideas and their interrelationships, offering scholars important perspectives on the fundamental framework and recurring motifs in scholarly literature. The column labelled "Node" displays the specific words that have undergone analysis. The keywords encapsulate fundamental ideas and recurring themes contained in the dataset. The column labelled "Cluster" is used to assign each phrase to a distinct cluster, denoting the thematic group to which it is associated.

The metric known as "Betweenness" quantifies the significance of a word in its role as a mediator or link between other terms in each network. Greater betweenness values indicate that the phrase assumes a central and influential position in facilitating connections among different notions. As an example, the metric of betweenness centrality assigns a value of 476.62 to the concept of "big data," indicating its notable role in facilitating connections among many subject matters. The measure known as "Closeness" is used to quantify the degree of centrality shown by a word inside a given network. A higher value of proximity indicates a stronger degree of association between the given word and other terms. Inside the given context, the term "big data" has a closeness score of 0.0185, indicating its significant centrality inside the network.

The statistic known as "PageRank" quantifies the significance of a phrase by considering its connectedness to other terms that have significant influence. A greater PageRank rating indicates that the phrase is associated with significant topics. As an example, the word "big data" has a comparatively elevated PageRank score of 0.0653, indicating its correlation with prominent terms. Additional words, such as "data analytics," "health care," and "diagnosis," also exhibit significant values across these measures, suggesting their significance in facilitating, consolidating, and impacting the network of terms.

Table 6. Co-word network analysis

Node	Betweenness	Closeness	PageRank
big data	476.619806086328	0.0185185185185185	0.065293336474265
data analytics	25.4867294032396	0.0117647058823529	0.0299444787238548
health care	63.7595302839547	0.0140845070422535	0.0308382725802729
internet of things	1.32145069805017	0.010752688172043	0.0154565364177037
advanced analytics	0.313200568990043	0.0104166666666667	0.0106764448399333
data handling	0.162708259482453	0.0104166666666667	0.0105649076450452
data mining	0.241662832989468	0.0104166666666667	0.00690682934253596
diagnosis	4.64517330688718	0.0117647058823529	0.0142690648555363
wearable sensors	0.469679454216444	0.0105263157894737	0.0106229284351244
decision making	0	0.0101010101010101	0.00627779568238119

3.2 Conclusion

Within the domain of healthcare, the integration of technology and data-driven insights has instigated significant advancements, notably in the arena of early identification and prevention of age-related disorders among the senior demographic [45]. The thorough examination of the presented data tables and figures highlights the importance of this endeavour, showcasing a complex and varied environment characterised by a range of trends, influential research sources, and essential terminology [46]. In addition, the primary aim of this study, under the framework of "Big Data Analytics for Early Detection and Prevention of Age-Related Diseases in Elderly Healthcare," is to conduct a thorough analysis and get a full understanding of the dynamic landscape of academic pursuits in this multidisciplinary domain. This purpose involves the examination of chronological patterns, prominent sources, important documents, common terminology, and complex interrelationships within the field.

The period from 2013 to 2023 demonstrates a clear temporal progression of papers in this field, which serves as a reflection of the dynamic character of scientific endeavours. The consistent yearly growth rate of 11.61% serves as a testament to the enduring commitment towards the exploration of inventive approaches in the field of geriatric healthcare. The expansion of research activities is accompanied by a corresponding increase in the variety of document kinds, such as articles, conference papers, and reviews. This diversity serves as evidence of the interdisciplinary nature of the study being conducted. The research also emphasises the significance of certain sources, shedding attention on major platforms that have made substantial contributions to the conversation. Publications such as "Advances in Intelligent Systems and Computing" and the "International Journal of Environmental Research and Public Health" have emerged as significant contributors in defining the trajectory of progress within this discipline. It is worth noting that cooperation plays a significant role in this context, with writers establishing collaborations at both local and international levels, as seen by the data on co-authorship.

The relocation towards the core study topics has conformed to the guidelines outlined by Bradford's Law. Journals are categorised into zones, which allow for the identification of concentrated output from prominent sources, as well as a wider spread of contributions from other authors. This research sheds light on the dynamics of scholarly output and identifies platforms that possess the capacity to transform the academic environment. Furthermore, by highlighting the most often referenced publications, this exhibition presents the fundamental ideas that have influenced the discourse around the early identification of diseases [47]. The publications under consideration have provided fundamental insights in several domains, such as big data analytics, Internet of Things (IoT) applications in healthcare, and machine learning, as shown by the substantial number of citations they have received. The presence of these citations serves as a measure of the influence and significance of these works, reflecting their wide recognition and relevance throughout the academic community.

Moreover, the vocabulary forms the foundation of this study endeavour. The discipline is characterised by the prevalence of terms such as "big data," "health care," and "data analytics," which provide insight into its primary issues and areas of focus [48]. The comprehensive approach to preventing age-related diseases is underscored by the focus on technology, data management, and healthcare practises [49]. The co-word network analysis offers a structural framework that elucidates the interconnections among pivotal phrases. This mapping aims to identify significant phrases such as "big data," which serve as crucial connectors between various ideas and have substantial influence in the spread of information [50]. The intricate network of links highlighted in this instance serves as a prime illustration of the multidisciplinary character of early illness detection research, underscoring its need for data-driven insights and collaborative endeavours [51].

In summary, the amalgamation of these evidence-based observations elucidates the profound metamorphosis taking place in the domain of big data analytics for the timely identification and mitigation of age-related ailments in geriatric healthcare [52]. The story of innovation is constructed via the constant growth, collaborative spirit, influential sources, foundational texts, and interrelated language, which together aim to bridge the gap between technology and healthcare to increase the well-being of the senior population [53], [54]. The use of a multidimensional approach has the capacity to fundamentally transform the trajectory of geriatric healthcare by virtue of its rigorous research methodologies and dynamic collaborative efforts.

4 REFERENCES

[1] D. J. Stein *et al.*, "Psychiatric diagnosis and treatment in the 21st century: Paradigm shifts versus incremental integration," *World Psychiatry*, vol. 21, no. 3, pp. 393–414, 2022. https://doi.org/10.1002/wps.20998

[2] E. D. Romanopoulou *et al.*, "Technology enhanced health and social care for vulnerable people during the COVID-19 outbreak," *Front Hum Neurosci*, vol. 15, p. 721065, 2021. https://doi.org/10.3389/fnhum.2021.721065

[3] J. B. Awotunde, A. E. Adeniyi, R. O. Ogundokun, G. J. Ajamu, and P. O. Adebayo, "MIoT-based big data analytics architecture, opportunities and challenges for enhanced telemedicine systems," *Studies in Fuzziness and Soft Computing*, vol. 410, pp. 199–220, 2021. https://doi.org/10.1007/978-3-030-70111-6_10

[4] V. Binnie, R. Le Brocque, M. Jessup, and A. N. B. Johnston, "Adult frequent presentation to emergency departments and adverse childhood experiences: A scoping review," *Australas Emerg Care*, vol. 24, no. 4, pp. 264–279, 2021. https://doi.org/10.1016/j.auec.2020.11.002

[5] N. Mehta and A. Pandit, "Concurrence of big data analytics and healthcare: A systematic review," *Int J Med Inform*, vol. 114, pp. 57–65, 2018. https://doi.org/10.1016/j.ijmedinf.2018.03.013

[6] V. Chamola, V. Hassija, V. Gupta, and M. Guizani, "A comprehensive review of the COVID-19 pandemic and the role of IoT, drones, AI, blockchain, and 5G in managing its impact," *IEEE Access*, vol. 8, pp. 90225–90265, 2020. https://doi.org/10.1109/ACCESS.2020.2992341

[7] H. Yin *et al.*, "Population ageing and deaths attributable to ambient PM2·5 pollution: A global analysis of economic cost," *Lancet Planet Health*, vol. 5, no. 6, pp. e356–e367, 2021. https://doi.org/10.1016/S2542-5196(21)00131-5

[8] M. Lopreite and Z. Zhu, "The effects of ageing population on health expenditure and economic growth in China: A Bayesian-VAR approach," *Soc Sci Med*, vol. 265, p. 113513, 2020. https://doi.org/10.1016/j.socscimed.2020.113513

[9] Y. K. Dwivedi *et al.*, "Metaverse beyond the hype: Multidisciplinary perspectives on emerging challenges, opportunities, and agenda for research, practice and policy," *Int J Inf Manage*, vol. 66, p. 102542, 2022. https://doi.org/10.1016/j.ijinfomgt.2022.102542

[10] E. F. Fang *et al.*, "A research agenda for ageing in China in the 21st century (2nd edition): Focusing on basic and translational research, long-term care, policy and social networks," *Ageing Res Rev*, vol. 64, p. 101174, 2020. https://doi.org/10.1016/j.arr.2020.101174

[11] O. Yilmaz, K. Pinto, and V. Deshpande, "Navigating the challenges associated with a diagnosis of autoimmune pancreatitis and IgG4-related sclerosing cholangitis," *Arch Pathol Lab Med*, vol. 147, no. 3, pp. 283–293, 2023. https://doi.org/10.5858/arpa.2021-0549-RA

[12] J. Rychik *et al.*, "Evaluation and management of the child and adult with fontan circulation: A scientific statement from the American heart association," *Circulation*, vol. 140, no. 6, pp. E234–E284, 2019. https://doi.org/10.1161/CIR.0000000000000696

[13] A. Ferri *et al.*, "The HIBAD experience: Using digital health technologies in the GDPR era," *Health Policy Technol*, p. 100788, 2023. https://doi.org/10.1016/j.hlpt.2023.100788

[14] N. Manikandan, S. N. Tadiboina, M. S. Khan, V. P. R, R. Singh, and K. K. Gupta, "Automation of smart home for the wellbeing of elders using empirical big data analysis," in *2023 3rd International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE)*, 2023, pp. 1164–1168. https://doi.org/10.1109/ICACITE57410.2023.10183103

[15] M. M. Rathore, S. A. Shah, D. Shukla, E. Bentafat, and S. Bakiras, "The role of AI, machine learning, and big data in digital twinning: A systematic literature review, challenges, and opportunities," *IEEE Access*, vol. 9, pp. 32030–32052, 2021. https://doi.org/10.1109/ACCESS.2021.3060863

[16] I. E. Agbehadji, B. O. Awuzie, A. B. Ngowi, and R. C. Millham, "Review of big data analytics, artificial intelligence and nature-inspired computing models towards accurate detection of COVID-19 pandemic cases and contact tracing," *International Journal of Environmental Research and Public Health,* vol. 17, no. 15, p. 5330, 2020. https://doi.org/10.3390/ijerph17155330

[17] M. R. King, "The future of AI in medicine: A perspective from a chatbot," *Ann Biomed Eng*, vol. 51, no. 2, pp. 291–295, 2023. https://doi.org/10.1007/s10439-022-03121-w
[18] J. Khalaily, "Preventing the need for nursing home admittance," 2023, Accessed: Sep. 01, 2023. [Online]. Available: https://seaopenresearch.eu/Journals/spas/
[19] D. Moher *et al.*, "Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement," *PLoS Medicine*, vol. 6, no. 7, p. e1000097, 2009. https://doi.org/10.1371/journal.pmed.1000097
[20] I. Mustapha, N. Khan, and M. I. Qureshi, "Is technology affecting the way our minds operate? Digital psychology of users in the era of digitalization," *Advanced Structured Materials*, vol. 174, pp. 71–92, 2022. https://doi.org/10.1007/978-3-031-01488-8_8
[21] B. Buyamin *et al.*, "The influence of war and global economy on article publication (Bibliometric analysis using Biblioshiny-R)," 2023. https://doi.org/10.21203/rs.3.rs-2680363/v1
[22] P. Jiang, J. Winkley, C. Zhao, R. Munnoch, G. Min, and L. T. Yang, "An intelligent information forwarder for healthcare big data systems with distributed wearable sensors," *IEEE Syst J*, vol. 10, no. 3, pp. 1147–1159, 2016. https://doi.org/10.1109/JSYST.2014.2308324
[23] Rastogi, D. K. Chaturvedi, S. Sagar, N. Tandon, and M. Rastogi, "Risk stratification for subjects suffering from lung carcinoma: Healthcare 4.0 approach with medical diagnosis using computational intelligence," *EAI/Springer Innovations in Communication and Computing*, pp. 171–193, 2022. https://doi.org/10.1007/978-3-030-77528-5_9
[24] K. Kumar and M. R. Bagavathi, "Thriving the ills of elderly people using big data analytics," in *2016 International Conference on Computer Communication and Informatics, ICCCI 2016*, 2016. https://doi.org/10.1109/ICCCI.2016.7479952
[25] Y. Gao and R. Li, "Development of the model of combining medical care with old-age care driven by 'Internet +,' " in *Advances in Intelligent Systems and Computing*, vol. 1117 AISC, 2020, pp. 43–49. https://doi.org/10.1007/978-981-15-2568-1_7
[26] A. Danial-Saad, M. Corzani, C. Tacconi, and L. Chiari, "Usability of a touchscreen assessment tool (TATOO) prototype for clinicians and typically developing children," 2022. https://doi.org/10.1080/17483107.2022.2137250
[27] M. S. Hossain and G. Muhammad, "Emotion-aware connected healthcare big data towards 5G," *IEEE Internet Things J*, vol. 5, no. 4, pp. 2399–2406, 2018. https://doi.org/10.1109/JIOT.2017.2772959
[28] Y. Li *et al.*, "Global, regional, and national disease burden estimates of acute lower respiratory infections due to respiratory syncytial virus in children younger than 5 years in 2019: A systematic analysis," *The Lancet*, vol. 399, no. 10340, pp. 2047–2064, 2022. https://doi.org/10.1016/S0140-6736(22)00478-0
[29] O. L. Bajenaru and A. M. Custura, "Enhanced framework for an elderly-centred platform: Big data in monitoring the health status," in *Proceedings – 2019 22nd International Conference on Control Systems and Computer Science, CSCS 2019*, pp. 643–648, 2019. https://doi.org/10.1109/CSCS.2019.00116
[30] O. Anya and H. Tawfik, "Leveraging big data analytics for personalized elderly care: Opportunities and challenges," *Applied Computing in Medicine and Health*, pp. 99–124, 2016. https://doi.org/10.1016/B978-0-12-803468-2.00005-9
[31] K. K. F. Tsoi *et al.*, "Blood pressure monitoring on the cloud system in elderly community centres: A data capturing platform for application research in public health," *Proceedings – 2016 7th International Conference on Cloud Computing and Big Data, CCBD 2016*, pp. 312–315, 2017. https://doi.org/10.1109/CCBD.2016.068
[32] R. Rocha, D. Carneiro, A. P. Pinheiro, and P. Novais, "A customizable game-inspired application for memory stimulation," *Advances in Intelligent Systems and Computing*, vol. 806, pp. 282–289, 2019. https://doi.org/10.1007/978-3-030-01746-0_33

[33] M. Willetts, A. S. Atkins, and C. Stanier, "Big data, big data analytics application to smart home technologies and services for geriatric rehabilitation," *Smart Home Technologies and Services for Geriatric Rehabilitation*, pp. 205–230, 2022. https://doi.org/10.1016/B978-0-323-85173-2.00001-1

[34] A. A. Qaffas, R. Hoque, and N. Almazmomi, "The Internet of Things and big data analytics for chronic disease monitoring in Saudi Arabia," *Telemed J E Health*, vol. 27, no. 1, pp. 74–81, 2021. https://doi.org/10.1089/tmj.2019.0289

[35] K. Rezaee, M. R. Khosravi, N. Neshat, and M. K. Moghimi, "Deep transfer learning-based fall detection approach using IoMT-enabled thermal imaging-assisted pervasive surveillance and big health data," vol. 31, no. 12, 2022. https://doi.org/10.1142/S0218126622400059

[36] S. L. Baxter *et al.*, "Predictive analytics for glaucoma using data from the all of us research program," *Am J Ophthalmol*, vol. 227, pp. 74–86, 2021. https://doi.org/10.1016/j.ajo.2021.01.008

[37] S. Y. Kang, S. W. Seo, and J. Y. Kim, "Comprehensive risk factor evaluation of postoperative delirium following major surgery: Clinical data warehouse analysis," *Neurol Sci*, vol. 40, no. 4, pp. 793–800, 2019. https://doi.org/10.1007/s10072-019-3730-1

[38] J. Yu *et al.*, "Semantic analysis of NIH stroke scale using machine learning techniques," in *2019 International Conference on Platform Technology and Service, PlatCon 2019 – Proceedings*, 2019. https://doi.org/10.1109/PlatCon.2019.8668961

[39] K. P. High *et al.*, "Use of functional assessment to define therapeutic goals and treatment," *J Am Geriatr Soc*, vol. 67, no. 9, pp. 1782–1790, 2019. https://doi.org/10.1111/jgs.15975

[40] D. Yacchirema, D. Sarabia-Jácome, C. E. Palau, and M. Esteve, "System for monitoring and supporting the treatment of sleep apnea using IoT and big data," *Pervasive Mob Comput*, vol. 50, pp. 25–40, 2018. https://doi.org/10.1016/j.pmcj.2018.07.007

[41] R. K. Goel and S. Vishnoi, "Strengthening and sustaining health-related outcomes through digital health interventions," *Journal of Engineering Science and Technology Review*, vol. 12, no. 2, pp. 10–17, 2023. https://doi.org/10.25103/jestr.162.02

[42] M. Roshni Thanka, S. Juliet, E. Bijolin Edwin, and R. Raahul John, "Deep learning-based big data analytics model for activity monitoring of elderly people," *Lecture Notes in Electrical Engineering*, vol. 905, pp. 589–599, 2022. https://doi.org/10.1007/978-981-19-2177-3_55

[43] I. Hussain and S. J. Park, "Big-ECG: Cardiographic predictive cyber-physical system for stroke management," *IEEE Access*, vol. 9, pp. 123146–123164, 2021. https://doi.org/10.1109/ACCESS.2021.3109806

[44] M. J. Al Nahian, T. Ghosh, M. N. Uddin, M. M. Islam, M. Mahmud, and M. S. Kaiser, "Towards artificial intelligence driven emotion aware fall monitoring framework suitable for elderly people with neurological disorder," *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 12241 LNAI, pp. 275–286, 2020. https://doi.org/10.1007/978-3-030-59277-6_25

[45] M. R. Lima *et al.*, "Conversational affective social robots for ageing and dementia support," *IEEE Trans Cogn Dev Syst*, vol. 14, no. 4, pp. 1378–1397, 2022. https://doi.org/10.1109/TCDS.2021.3115228

[46] N. Khan, M. I. Qureshi, I. Mustapha, S. Irum, and R. N. Arshad, "A systematic literature review paper on online medical mobile applications in Malaysia," *International Journal of Online and Biomedical Engineering*, vol. 16, no. 1, pp. 63–82, 2020. https://doi.org/10.3991/ijoe.v16i01.12263

[47] H. Sikandar, Y. Vaicondam, N. Khan, M. I. Qureshi, and A. Ullah, "Scientific mapping of industry 4.0 research: A bibliometric analysis," *International Journal of Interactive Mobile Technologies*, vol. 15, no. 18, pp. 129–147, 2021. https://doi.org/10.3991/ijim.v15i18.25535

[48] M. Hassan *et al.*, "Innovations in genomics and big data analytics for personalized medicine and health care: A review," *International Journal of Molecular Sciences,* vol. 23, no. 9, p. 4645, 2022. https://doi.org/10.3390/ijms23094645

[49] Z. Wei *et al.*, "Insights into the application of machine learning in industrial risk assessment: A bibliometric mapping analysis," *Sustainability,* vol. 15, no. 8, p. 6965, 2023. https://doi.org/10.3390/su15086965

[50] K. Ragazou, I. Passas, A. Garefalakis, E. Galariotis, and C. Zopounidis, "Big data analytics applications in information management driving operational efficiencies and decision-making: Mapping the field of knowledge with bibliometric analysis using R," *Big Data and Cognitive Computing,* vol. 7, no. 1, p. 13, 2023. https://doi.org/10.3390/bdcc7010013

[51] M. Mortati, S. Magistretti, C. Cautela, and C. Dell'Era, "Data in design: How big data and thick data inform design thinking projects," *Technovation*, vol. 122, p. 102688, 2023. https://doi.org/10.1016/j.technovation.2022.102688

[52] J. G. Siri and I. Geddes, "Mainstreaming health in urban design and planning: Advances in theory and practice," vol. 6, no. 5, pp. 853–857, 2022. https://doi.org/10.1080/23748834.2022.2148844

[53] L. Rykkje, M. B. Søvik, L. Ross, W. McSherry, P. Cone, and T. Giske, "Educational interventions and strategies for spiritual care in nursing and healthcare students and staff: A scoping review," *J Clin Nurs*, vol. 31, no. 11–12, pp. 1440–1464, 2022. https://doi.org/10.1111/jocn.16067

[54] A. Watson *et al.*, "Indigenous strength: Braiding culture, ceremony and community as a response to the COVID-19 pandemic," *Int J Indig Health*, vol. 17, no. 1, 2022. https://doi.org/10.32799/ijih.v17i1.36716

5 AUTHORS

Mohd Amran Mohd Daril, Quality Engineering Research Cluster, Quality Engineering Section, Universiti Kuala Lumpur, Malaysian Institute of Industrial Technology, Johor Bahru, Johor, Malaysia (E-mail: mamran@unikl.edu.my).

Shazia Qayuum, Institute of Applied Psychology, University of Punjab, Lahore, Pakistan (E-mail: shazia_agha@hotmail.com).

Alhamzah F. Abbas, Faculty of Management, Universiti Teknologi Malaysia, Johor Bahru 81310, Malaysia (E-mail: Alhamza.fadil@gmail.com; ORCID: http://orcid.org/0000-0002-7508-9340).

Nguyen Thuy Van, Journal of Science Department, Phu Yen University, Tuy Hoa City, Vietnam (E-mail: nguyenthuyvan@pyu.edu.vn).