Population Health Management: A Bibliometric Analysis of Literature on Data Analytics and Public Health Interventions

Loso Judijanto[1], Yocki Yuanti[2], Supriandi[3]
[1] IPOSS Jakarta, Indonesia and losojudijantobumn@gmail.com
[2] STIKes Mitra Ria Husada Jakarta and yockiyuanti@gmail.com
[3] Telkom University and supriandi@student.telkomuniversity.ac.id

ABSTRACT

Population health management (PHM) has emerged as a pivotal strategy in healthcare, aiming to address the health needs of communities while enhancing care quality and outcomes. Recent years have witnessed a surge in research at the intersection of data analytics and public health interventions within the realm of PHM. This paper presents a comprehensive bibliometric analysis of the literature on data analytics and public health interventions in the context of population health management. Through systematic data collection from reputable scholarly databases and rigorous analysis, this study aims to map out the trends, key contributors, and emerging themes shaping the discourse in this domain. The analysis encompasses publication trends, authorship patterns, citation metrics, and network visualizations to provide insights into the landscape of literature. Notable findings include a high citation rate per paper and author, indicative of scholarly interest and collaboration within the research community. Key themes identified include outcomes, interventions, population health improvement, determinants, and health disparities. Additionally, temporal analysis reveals a shift towards non-communicable diseases and the increasing role of data in patient care. The author collaboration network highlights a core cluster of researchers with extensive collaborative ties, alongside peripheral authors and topics representing specialized or nascent areas. The density visualization map identifies potential areas for future research, including health disparities, women's health issues, and geographic-specific studies.

Keywords: *Population Health Management, Dana Analytics, Public Health Interventions, Bibliometric Analysis*

1. INTRODUCTION

Population health management (PHM) has emerged as a crucial approach in addressing the health needs of communities, focusing on improving health outcomes and enhancing the quality of care delivery [1], [2]. In recent years, there has been a notable surge in research exploring the intersection of data analytics and public health interventions within the realm of PHM [3], [4]. As the healthcare landscape continues to evolve, the integration of advanced data analytics techniques offers promising avenues for optimizing resource allocation, identifying at-risk populations, and designing targeted interventions to mitigate health disparities [5]. Understanding the landscape of literature in this domain through bibliometric analysis is essential for comprehensively mapping out the trends, key contributors, and emerging themes shaping the discourse.

The convergence of data analytics and public health interventions presents a rich area of study, yet it also brings forth a myriad of challenges and opportunities [6]. With the proliferation of electronic health records, wearable devices, and other sources of health-related data, there is a growing need to harness the power of big data analytics to extract meaningful insights for informing evidence-based interventions [7]. However, the effective translation of data-driven insights into actionable public health strategies requires interdisciplinary collaboration, methodological rigor, and consideration of ethical implications [8]. Moreover, as the global health landscape grapples with

unprecedented challenges such as pandemics and chronic diseases, there is a pressing demand for innovative approaches that leverage data analytics to enhance population health outcomes [6], [9].

At the heart of this research lies the exploration of key research questions aimed at unraveling the intricacies of data analytics-driven public health interventions within the context of population health management. By critically examining the existing literature, this study seeks to identify gaps, trends, and areas of consensus or contention within the field. Specifically, it aims to elucidate the methodological approaches employed, the domains of public health addressed, and the outcomes assessed in studies at the intersection of data analytics and PHM. Through a systematic bibliometric analysis, this research endeavors to contribute to the understanding of how data-driven strategies can be effectively leveraged to optimize population health outcomes and inform policy and practice.

The overarching objective of this research is to shed light on the evolving landscape of population health management, with a particular focus on the role of data analytics in shaping public health interventions. By synthesizing and analyzing the existing body of literature, this study seeks to provide insights into the current state of knowledge, identify areas for further research, and inform evidence-based decision-making in population health practice and policy. Ultimately, the findings of this research endeavor are poised to offer valuable guidance to stakeholders across the healthcare continuum, including policymakers, practitioners, researchers, and healthcare administrators, in navigating the complexities of PHM in an era increasingly defined by data-driven innovation.

2. LITERATURE REVIEW

2.1 Population Health Management

Population Health Management (PHM) is a crucial approach in healthcare systems, aiming to address challenges like an aging population, chronic diseases, and rising costs [10]. It involves not only caring for sick individuals but also providing preventive services to maintain the health of the population [11]. PHM helps identify gaps between actual care and best practices, enabling targeted interventions for improvement [12]. Many healthcare facilities are adopting PHM for cost recovery planning and risk management, with a significant number utilizing PHM or similar activities [13]. While PHM traditionally focuses on general population care, the emergence of precision medicine emphasizes individual uniqueness and digitization, utilizing technologies like cloud computing and big data to complement PHM's efforts [14]. Considering wider determinants of health, PHM plays a vital role in shaping the future landscape of healthcare services.

2.2 Data Analytics

Data analytics involves analyzing raw data to extract insights and make informed decisions [15]–[17]. It encompasses various techniques and tools to optimize processes, enhance efficiency, reduce costs, and improve decision-making [18]. The process includes preprocessing data, extracting relevant variables, and applying statistical methods to draw meaningful conclusions. Data analysis is crucial in research projects, helping researchers derive answers from the data, verify accuracy, and support decision-making. It provides valuable insights based on chosen objectives, patterns,

relationships, and trends in the data. Overall, data analytics has widespread applications across sectors, aiding in understanding customer trends, enhancing products and services, and enabling effective marketing strategies.

2.3 Public Health Interventions

Public Health Interventions encompass a wide array of strategies aimed at improving population health in developing countries [19]. These interventions focus on addressing socio-economic factors, promoting healthy behaviors, and collaborating with communities [20]. The Public Health Intervention Wheel (PHI Wheel) serves as a model for public health nursing practice, incorporating 17 interventions across community, systems, and individual/family levels [21]. Additionally, the emergence of digital public health interventions leverages digital technologies to enhance public health through prevention and health promotion [22]. These interventions aim to align with essential public health functions, involve the target population in development, and adapt to different healthcare systems [23]. Overall, these interventions play a crucial role in influencing determinants of health beyond healthcare services to safeguard and enhance population health in diverse settings.

3. METHODS

3.1 Data Collection

The first step in conducting the bibliometric analysis involved comprehensive data collection from scholarly databases renowned for their coverage of health-related literature. Databases such as PubMed, Scopus, Web of Science, and Embase were systematically searched to retrieve relevant articles. The search strategy encompassed keywords and Boolean operators tailored to capture studies pertaining to population health management, data analytics, and public health interventions. The inclusion criteria encompassed articles published within a specified timeframe, typically spanning the last decade to ensure relevance and currency.

3.2 Study Selection Criteria

Following the initial search, retrieved articles underwent a two-stage screening process to identify eligible studies for inclusion in the bibliometric analysis. In the first stage, titles and abstracts were screened based on predefined inclusion and exclusion criteria. Articles deemed potentially relevant proceeded to the second stage, wherein full-text screening was conducted to assess eligibility based on more detailed criteria. Eligible studies included empirical research articles, reviews, and meta-analyses focusing on the application of data analytics in public health interventions within the context of population health management.

3.3 Data Analysis

The extracted data underwent rigorous bibliometric analysis to elucidate patterns, trends, and relationships within the literature corpus. Quantitative bibliometric indicators such as publication trends over time, authorship patterns, and citation counts. Co-authorship networks and citation networks were visualized and analyzed using network analysis techniques to identify key contributors, collaborations, and seminal works shaping the discourse on data analytics and public health interventions in population health management.

4. RESULTS AND DISCUSSION

4.1 Research Data Matriks

Table 1. Research Data Metrics

Publication years	: 1984-2024
Citation years	: 40 (1984-2024)
Paper	: 840
Citations	: 50
Cites/year	: 8898.97
Cites/paper	: 423.74
Cites/author	: 150197.80
Papers/author	: 343.97
Author/paper	: 3.29
h-index	: 284
g-index	: 583
hI,norm	: 181
hI,annual	: 4.53
hA-index	: 69
Papers with ACC	: 1,2,5,10,20:802,790,731,563,280

Source: Publish or Perish Output, 2024

Table 1 presents a comprehensive overview of the bibliometric data metrics derived from the analysis of publications spanning the years 1984 to 2024 within the domain of population health management, focusing on data analytics and public health interventions. The dataset comprises a total of 840 papers, which have collectively garnered 50 citations over the citation period of 40 years. The average number of citations per year stands at an impressive 8898.97, reflecting the considerable impact and dissemination of the research in the field. On average, each paper has received 423.74 citations, indicating a high level of scholarly interest and engagement with the literature. The analysis further reveals a remarkably high citation rate per author, averaging at 150197.80, and a relatively high number of papers per author, averaging at 343.97. This suggests prolific contribution and collaboration within the research community. The h-index, g-index, and hI,norm metrics provide insights into the citation impact and productivity of the publications, with values of 284, 583, and 181, respectively. Additionally, the hI,annual and hA-index metrics offer perspectives on the annual h-index and the author-level h-index, further delineating the scholarly impact of individual authors. Finally, the table highlights the distribution of papers with different levels of accumulated citations, providing a nuanced understanding of the citation patterns within the literature corpus. Overall, these bibliometric metrics offer valuable insights into the visibility, impact, and scholarly productivity of research in the field of population health management, underscoring its significance and influence within the broader academic discourse.
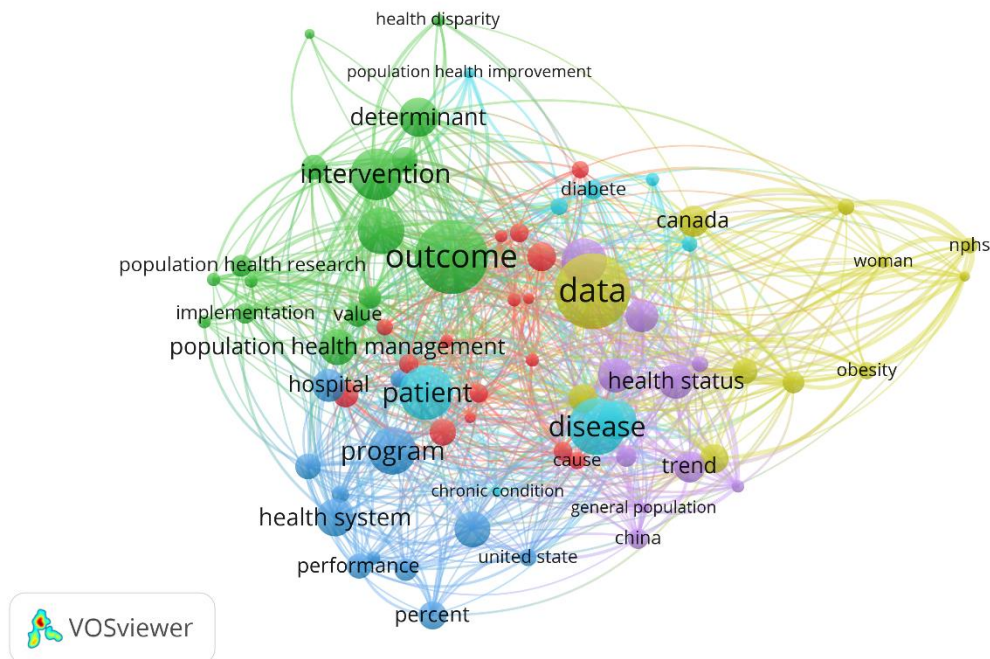
4.2 Network Visualization

Figure 1. Network Visualization
Source: Data Analysis Result, 2024

The network shows various terms that are connected with lines, indicating their relationships or co-occurrences within a particular body of literature or dataset. The size of the nodes (circles) likely represents the frequency of the term within the dataset, and the thickness of the lines might indicate the strength of the relationship between terms. There are several distinguishable clusters in the network, each likely representing a thematic area. The clusters are color-coded, which helps in understanding the thematic relatedness of the terms.

1. Green Cluster: Central themes are 'outcome', 'intervention', 'population health improvement', 'determinant', and 'health disparity'. This cluster might represent studies focusing on the effects of health interventions on population outcomes and disparities.
2. Red Cluster: Dominated by 'patient', 'data', 'disease', 'hospital', and 'health management'. This suggests a focus on patient data, disease management, and hospital-related health administration and data management.
3. Blue Cluster: Includes terms like 'health system', 'performance', 'program', and 'chronic condition'. This cluster may relate to healthcare systems' performance and management of chronic conditions.
4. Yellow Cluster: With 'diabetes', 'obesity', 'trend', and 'cause', it appears to revolve around non-communicable diseases (NCDs) and their trends and causes.
5. Purple Cluster: Contains 'health status', 'general population', 'china', and 'united states'. This might be indicative of studies that are demographic in nature, comparing health statuses among populations in different countries.

The interconnections between clusters show how these themes are related. For example, the yellow and purple clusters are closely connected, suggesting a lot of research at the intersection of NCD trends and demographic studies. The red and green clusters are also interconnected, indicating

that patient data is a significant aspect of studies on health outcomes and interventions. From this visualization, one could hypothesize that the dataset includes studies from healthcare management, patient data analysis, public health interventions, chronic disease trends, and comparative health status across different populations. Each cluster represents a body of literature that collectively informs a broader understanding of health and healthcare issues. The presence of 'Canada', 'China', 'United States', and 'NPHS' (which could stand for a specific health survey or program, such as the National Population Health Survey in Canada) indicates that there may be a geographic or nation-specific element to the research as well.

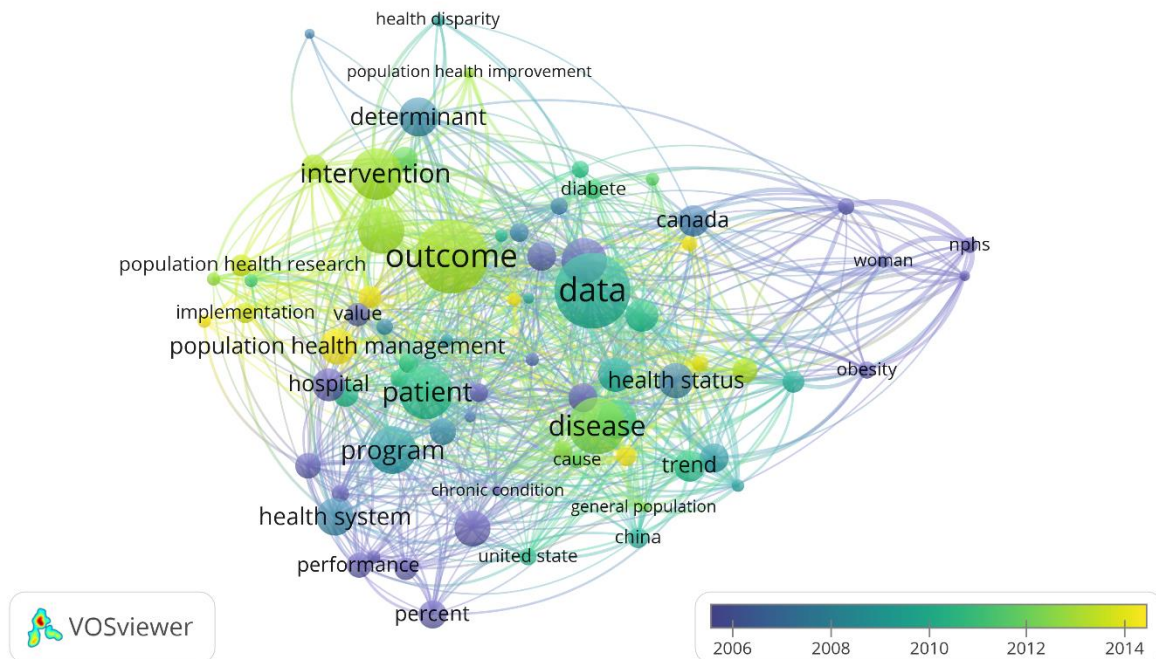### 4.3 Overlay Visualization

Figure 2. Overlay Visualization
Source: Data Analysis Result, 2024

The second image is similar to the first but with an added element: a color gradient at the bottom indicating a timeline from 2006 to 2014. This gradient is overlaid on the network visualization, providing temporal information about the research themes. The gradient indicates the time at which terms were most prominent within the dataset. Darker blue nodes represent terms that were more prevalent or emerged in the literature around 2006, while yellow nodes represent terms that are more recent in the context, around 2014. If a term's node is blue, it suggests that the term was more relevant in earlier research within the observed time frame. A yellow node suggests a more recent interest or a trend that has become more prominent in the later years.

Terms such as 'health system', 'performance', and 'chronic condition' are in the blue spectrum, indicating that research in these areas was more dominant or originated earlier in the timeframe, around 2006. This may reflect an initial focus on health system performance and chronic conditions. Terms that appear in green or light blue, like 'patient', 'hospital', 'program', and 'disease', suggest these topics were central to the literature in the intermediate years of the timeframe (circa 2008 to 2011). 'Diabetes', 'obesity', 'trend', 'Canada', and 'China' appear in yellow, suggesting these topics became more prominent towards the end of the timeframe, around 2014. There seems to be a shift toward non-communicable diseases and their trends, possibly with a focus on specific national contexts or the rise of global health as a discipline. The temporal analysis indicates a shift in research focus from system performance and chronic conditions towards more data-driven, patient-centric

models and an increasing interest in the global trends of non-communicable diseases like diabetes and obesity.

4.4 Citation Analysis

Table 2. The Most Impactful Literatures

Citations	Authors and year	Title
7996	[24]	Global and regional burden of disease and risk factors, 2001: systematic analysis of population health data
7248	[25]	The world health report 2000: health systems: improving performance
6990	[26]	Health literacy as a public health goal: a challenge for contemporary health education and communication strategies into the 21st century
6450	[27])	Health literacy and public health: a systematic review and integration of definitions and models
6442	[28]	Social determinants of health
6357	[29]	The triple aim: care, health, and cost
5898	[30]	Epidemiology of obstructive sleep apnea: a population health perspective
5638	[31]	EQ-SD: a measure of health status from the EuroQol Group
4757	[32]	No health without mental health
4486	[33]	Annual Medical Spending Attributable To Obesity: Payer-And Service-Specific Estimates: Amid calls for health reform, real cost savings are more likely to be achieved …

Source: Publish or Perish Output, 2024

Table 2 provides a snapshot of the most impactful literature within the field of population health management, as evidenced by their citation counts. The table showcases seminal works that have significantly shaped the discourse and understanding of key issues in public health and healthcare delivery. Topping the list is the landmark publication by [24], which systematically analyzes the global and regional burden of disease and risk factors, offering crucial insights into population health trends. Following closely is the [25] influential report on health systems performance, emphasizing the importance of improving healthcare delivery worldwide. Other notable contributions include works addressing critical topics such as health literacy [26], social determinants of health [28], and the triple aim framework for optimizing care, health, and cost [29]. These seminal works underscore the multifaceted nature of population health management and provide foundational knowledge that informs policy, practice, and research endeavors in the field.
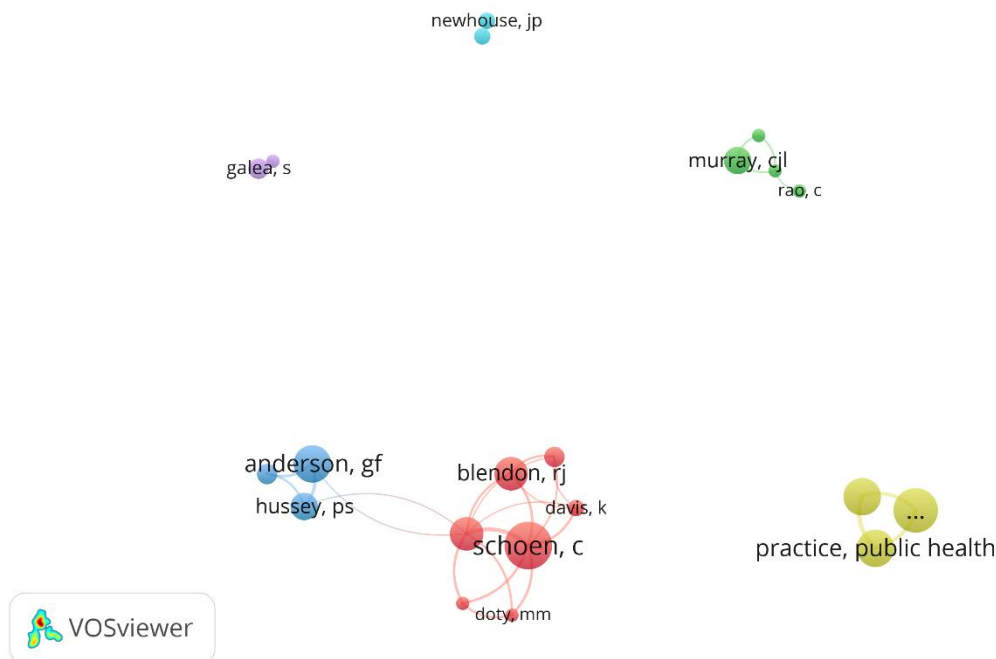
4.5 Author Collaboration

Figure 3. Density Visualization
Source: Data Analysis Result, 2024

The image presents an author collaboration network, which is a visualization that depicts the connections between researchers based on their co-authorship of academic papers or projects. Each node (circle) represents an author, and the lines between them indicate that the connected authors have collaborated on at least one publication. The size of each node is usually proportional to the number of collaborations or publications that the author has, and the thickness of the lines may represent the number of collaborative works. In this network:

1. The central cluster in red features authors such as 'schoen, c', 'blendon, rj', 'davis, k', 'doty, mm', 'anderson, gf', and 'hussey, ps'. These authors are closely connected, suggesting a strong collaborative relationship between them, which often indicates they may work in the same field or have shared research interests.

2. The isolated nodes, such as 'galea, s', 'newhouse, jp', 'murray, cj', and 'rao, c', suggest that these authors have fewer collaborations or perhaps collaborate in more specialized or disparate fields compared to the central cluster.

3. The node labeled 'practice, public health' seems to represent a thematic cluster or maybe an institution or a journal rather than an individual author. The fact that it is somewhat isolated with a few connections suggests specialized research or publication in the area of public health practice with less overlap with the other authors depicted.
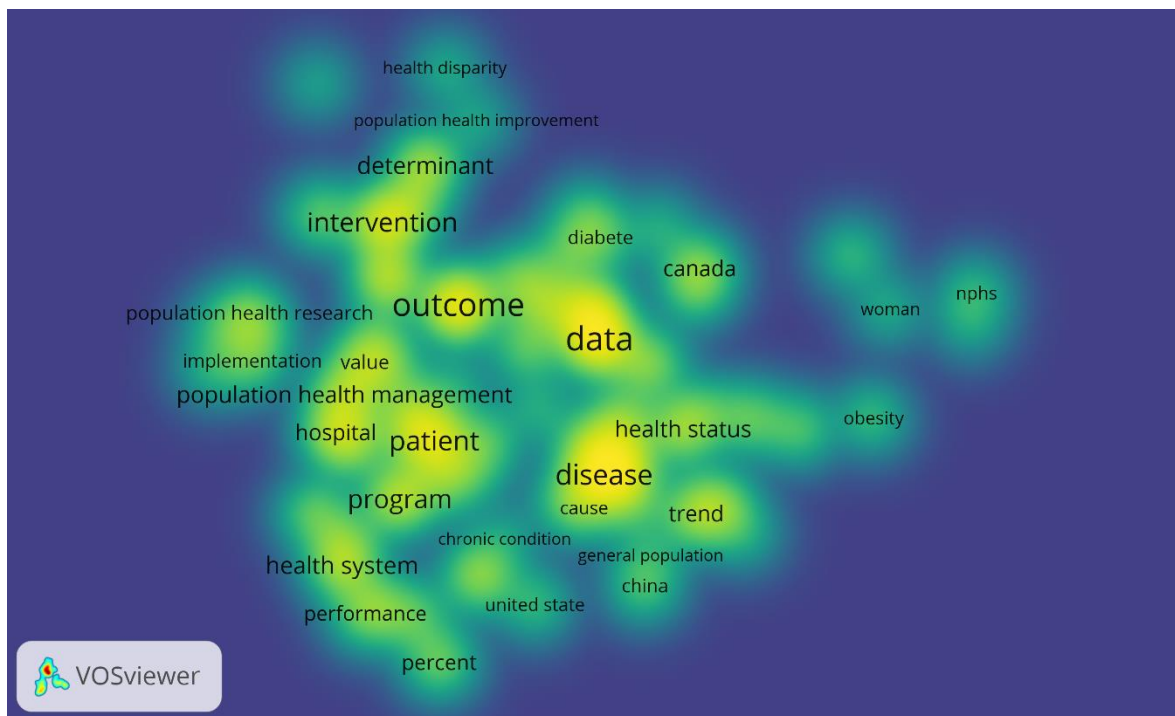
4.6 Density Visualization

Figure 4. Density Visualization
Source: Data Analysis Result, 2024

The varying intensities of color on the map represent the density of the terms, with the brighter areas indicating higher density or a higher frequency of occurrence. These bright areas typically suggest well-established or currently popular research topics within the dataset analyzed. Conversely, the less bright areas, which are not as densely colored, may indicate topics that are less explored or emerging. These less bright areas can be potential areas for future research, as they may represent gaps in the literature or emerging trends that have not yet been fully developed.

In the map above, areas that appear less bright and thus might be of potential interest for future research could include terms that are towards the periphery of the map or in areas of less intense color. These might be terms like "health disparity," "woman," "nphs" (which could stand for a specific health survey or program, such as the National Population Health Survey), and the geographical locations such as "Canada." Since these terms are not in the center of the dense clusters, they may represent niche areas or new lines of inquiry that are not as saturated with research.

If the aim is to identify potential research topics that are less explored, one strategy would be to look for those peripheral terms, explore their current state of research, and then determine how they could be expanded upon or connected with the more established topics at the center of the brighter clusters. For example, exploring "health disparity" in the context of "obesity" and "diabetes" might be an area with room for further research, or looking into specific health trends within "Canada" could yield new insights that are not as well represented in the broader literature.

CONCLUSION

The series of VOSviewer visualizations you provided illustrate the landscape of health-related research through various lenses, including keyword co-occurrence, temporal trends, author collaborations, and research density. From the keyword co-occurrence and temporal trend maps, we observe a research field centered around outcomes, interventions, population health, and disease management. Over time, there has been a shift from an initial focus on health systems and chronic conditions to a more recent emphasis on non-communicable diseases, trends in public health, and

the role of data in patient care, with specific attention to diabetes and obesity as emerging concerns. The author collaboration network revealed a core cluster of researchers with extensive collaborative ties, possibly indicating a well-established research community. In contrast, peripheral authors and topics might represent specialized or nascent areas within the broader field. Finally, the density visualization map suggests potential areas for future research. Less saturated topics like health disparities, women's health issues, specific population health surveys (like the NPHS), and geographic-specific studies such as those in Canada could offer fertile ground for new inquiries. These less-explored areas, illuminated by the less bright regions on the map, could be particularly promising for researchers looking to contribute novel insights or bridge gaps in the current literature. Collectively, these visualizations underscore the dynamic nature of health research, highlighting both the evolution of dominant themes over time and the collaborative networks that drive the field forward, while also hinting at emerging trends and underexplored niches ripe for future academic pursuit.

REFERENCES

[1] A. van Ede, N. Minderhout, M. Bruijnzeels, M. Numans, and V. Stein, "Mechanisms for a successful implementation of a Population Health Management approach: a scoping review.," *Int. J. Integr. Care*, vol. 22, 2022.

[2] H. Mansoor *et al.*, "Exploratory Data Analysis of Population Level Smartphone-Sensed Data," in *International Joint Conference on Computer Vision, Imaging and Computer Graphics*, Springer, 2021, pp. 206–231.

[3] N. Mirin, H. Mattie, L. Jackson, Z. Samad, and R. Chunara, "Data science in public health: building next generation capacity," *arXiv Prepr. arXiv2208.03461*, 2022.

[4] S. Mohammed, T.-H. Kim, R.-S. Chang, and C. Ramos, "Guest Editorial: Data Analytics for Public Health Care," *IEEE J. Biomed. Heal. Informatics*, vol. 26, no. 4, pp. 1409–1410, 2022.

[5] J. H. Porembka *et al.*, "Radiologists' increasing role in population health management: AJR expert panel narrative review," *Am. J. Roentgenol.*, vol. 218, no. 1, pp. 7–18, 2022.

[6] S. Tirunagari, S. Mohan, D. Windridge, and Y. Balla, "Addressing Challenges in Healthcare Big Data Analytics," in *International Conference on Multi-disciplinary Trends in Artificial Intelligence*, Springer, 2023, pp. 757–765.

[7] N. Berros, F. El Mendili, Y. Filaly, and Y. El Bouzekri El Idrissi, "Enhancing digital health services with big data analytics," *Big data Cogn. Comput.*, vol. 7, no. 2, p. 64, 2023.

[8] A. Mitra, B. Soman, R. Gaitonde, T. Bhatnagar, E. Nieuhas, and S. Kumar, "Data science approaches to public health: case studies using routine health data from India," in *International Conference on Data Management, Analytics & Innovation*, Springer, 2023, pp. 913–940.

[9] B. Preim and K. Lawonn, "A survey of visual analytics for public health," in *Computer Graphics Forum*, Wiley Online Library, 2020, pp. 543–580.

[10] M. Abraham, M. J. Lim, W. S. Tan, and J. Cheah, "Global trends towards population health management and key lessons and initiatives in the Singapore context," *Int. J. Integr. Care*, vol. 22, no. 3, 2022.

[11] L. Bridle, P. Cross, S. Danielli, H. Ashrafian, and A. Darzi, "Using Population Health Management to improve screening of women with history of gestational diabetes in an urban setting," 2022.

[12] B. Steenkamer, E. de Weger, H. Drewes, K. Putters, H. Van Oers, and C. Baan, "Implementing population health management: an international comparative study," *J. Health Organ. Manag.*, vol. 34, no. 3, pp. 273–294, 2020.

[13] A. S. Moumtzoglou and A. Pouliakis, "Population Health Management and the Science of Individuality in the Era of Quantum Computing," in *Quality of Healthcare in the Aftermath of the COVID-19 Pandemic*, IGI Global, 2022, pp. 8–41.

[14] S. Thomas, "Population health management and neurology services," *Br. J. Neurosci. Nurs.*, vol. 18, no. 5, pp. 236–239, 2022.

[15] C. E. Earley, "Data analytics in auditing: Opportunities and challenges," *Bus. Horiz.*, vol. 58, no. 5, pp. 493–500, 2015.

[16] V. Charles, A. Emrouznejad, T. Gherman, and J. Cochran, "Why data analytics is an art," *Significance*, vol. 19, no. 6, pp. 42–45, 2022.

[17] T. A. Runkler, *Data analytics*. Springer, 2020.

[18] J. W. Tukey, *Exploratory data analysis*, vol. 2. Reading, MA, 1977.

[19] V. V. Joseph, "The impact of public health interventions in a developing nation: an overview," *South Sudan Med. J.*, vol. 15, no. 4, pp. 147–151, 2022.

[20] C. Shaw and C. Cleghorn, "Public health interventions: the elephant in the room of the health system crisis," *New Zeal. Med. J.*, vol. 136, no. 1568, pp. 8–11, 2023.

[21] M. A. Schaffer, S. Strohschein, and K. Glavin, "Twenty years with the public health intervention wheel: Evidence for practice," *Public Health Nurs.*, vol. 39, no. 1, pp. 195–201, 2022.

[22] J. Wienert, T. Jahnel, and L. Maaß, "What are digital public health interventions? First steps toward a definition and an intervention classification framework," *J. Med. Internet Res.*, vol. 24, no. 6, p. e31921, 2022.

[23] J. Wu, F. Scarabel, B. Majeed, N. L. Bragazzi, and J. Orbinski, "The impact of public health interventions on delaying and mitigating against replacement by SARS-CoV-2 variants of concern," 2021.

[24]   A. D. Lopez, C. D. Mathers, M. Ezzati, D. T. Jamison, and C. J. L. Murray, "Global and regional burden of disease and risk factors, 2001: systematic analysis of population health data," *Lancet*, vol. 367, no. 9524, pp. 1747–1757, 2006.

[25]   W. H. Organization, *The world health report 2000: health systems: improving performance*. World Health Organization, 2000.

[26]   D. Nutbeam, "Health literacy as a public health goal: a challenge for contemporary health education and communication strategies into the 21st century," *Health Promot. Int.*, vol. 15, no. 3, pp. 259–267, 2000.

[27]   K. Sørensen *et al.*, "Health literacy and public health: a systematic review and integration of definitions and models," *BMC Public Health*, vol. 12, pp. 1–13, 2012.

[28]   M. Marmot and R. Wilkinson, *Social determinants of health*. Oup Oxford, 2005.

[29]   D. M. Berwick, T. W. Nolan, and J. Whittington, "The triple aim: care, health, and cost," *Health Aff.*, vol. 27, no. 3, pp. 759–769, 2008.

[30]   T. Young, P. E. Peppard, and D. J. Gottlieb, "Epidemiology of obstructive sleep apnea: a population health perspective," *Am. J. Respir. Crit. Care Med.*, vol. 165, no. 9, pp. 1217–1239, 2002.

[31]   R. Rabin and F. de Charro, "EQ-SD: a measure of health status from the EuroQol Group," *Ann. Med.*, vol. 33, no. 5, pp. 337–343, 2001.

[32]   M. Prince *et al.*, "No health without mental health," *Lancet*, vol. 370, no. 9590, pp. 859–877, 2007.

[33]   E. A. Finkelstein, J. G. Trogdon, J. W. Cohen, and W. Dietz, "Annual Medical Spending Attributable To Obesity: Payer-And Service-Specific Estimates: Amid calls for health reform, real cost savings are more likely to be achieved through reducing obesity and related risk factors.," *Health Aff.*, vol. 28, no. Suppl1, pp. w822–w831, 2009.