Leveraging Artificial Intelligence and Big Data Analytics for Improved Management and Research of Multimorbidity

George Hernandez [1]
[1] Universidad Centroccidental Lisandro Alvarado, Barquisimeto, Estado Lara

Keywords

Multimorbidity
Artificial Intelligence
Machine Learning
Demographic Ageing
Chronic Illnesses

ABSTRACT

Multimorbidity denotes the simultaneous presence of two or more chronic illnesses inside an individual. Consequently, individuals with multimorbidity possess diverse and specific treatment requirements. In reality, fulfilling these objectives is challenging due to the fact that the organisational procedures of contemporary healthcare systems are mostly designed for a singular condition. A transformative shift in the problem-solving methodology for medical research and therapy is essential to enhance clinical decision-making and patient care in multimorbidity. Alongside the conventional reductionist methodology, we advocate for participatory research bolstered by artificial intelligence (AI) and sophisticated big data analytics. This research methodology, when used with data typically gathered in healthcare environments, offers a cohesive framework for investigating multimorbidity-related research problems. This may include, for instance, prediction, correlation, and classification challenges derived from various interaction aspects. To actualise the concept of this paradigm change in multimorbidity research, it is essential to optimise, standardise, and, most critically, integrate electronic health data into a unified national and worldwide research infrastructure. Ultimately, there is a need for the integration and use of effective AI methodologies, especially deep learning, into clinical practice directly into the workflows of healthcare professionals.

Introduction

Industrialised nations globally are confronting a rising prevalence of chronic illnesses, such as type 2 diabetes, cardiovascular disorders, neurological conditions, and different malignancies. This adverse trend stems from an ageing demographic and the ubiquity of contemporary lifestyles, including the intake of industrially processed foods, predominantly sedentary occupations, and escalating chronic psychological stress, all of which are recognised to expedite ageing and the onset of age-related ailments [1,2].

Chronic illnesses seldom manifest as a singular condition in an individual; rather, two or more diseases coexist, a phenomenon known as multimorbidity [3]. Patients with multimorbidity provide a challenge for both policymakers and healthcare providers due to their intricate care requirements, necessitating the involvement of many healthcare professionals and services to effectively manage their care [4]. General practitioners (GPs) have the challenging responsibility of synthesising various advice and prescriptions from numerous providers [5].

Furthermore, existing clinical recommendations are disease-focused, hence confounding decision-making for these patients [6]. Recommendations for the management of individual disorders the prognosis for these people may be ambiguous due to the common exclusion of individuals with multimorbidity from research studies. The provision of treatment and patient self-management may be hindered by complex pharmaceutical regimens and an overload of information [5]. Recommendations provided to a patient for several individual disorders may be inherently contradictory and result in detrimental effects rather than benefits [7].

Multimorbidity significantly adversely affects patient outcomes, and not all patients with multimorbidity have the same risk for unfavourable outcomes. It has been shown to be contingent upon the quantity of comorbidities, as well as specific combinations of illnesses present in an individual [8]. Certain illness combinations arise randomly, since certain diseases, like hypertension, are prevalent in the community, while others tend to cluster. Disease clustering often relies on shared pathophysiology, as shown by the frequent co-occurrence of cardio-metabolic and vascular illnesses, but in many instances, the aetiologies are less discernible. Nonetheless, traditional approaches for assessing multimorbidity used in epidemiological surveys, which rely on enumerating disorders, are insufficient for accurately identifying pre-existing problems. [11,12].

Adequate information is lacking about patient-centered methods for addressing multimorbidity issues, particularly in forecasting specific outcomes and formulating personalised therapies [13,14]. This arises from the absence of a methodological framework that effectively addresses the complexity of multimorbidity. To elucidate our definition of complexity, we provide the following instances.

In an older population (>60), an increase in chronic illnesses correlates with a rise in the incidence of mental disorders, including anxiety and depression [15,16]. Somatic and mental illnesses are not entirely separate, since they share overlapping processes, including connections between mental disorders and cardio-metabolic and chronic pain syndromes [17–19]. The importance of these results is evident in the documented negative impact of mental illnesses on the progression of chronic somatic diseases and health outcomes, which may warrant the proactive identification of these disorders in older, multimorbid patients within primary care environments [20–22].

A further aspect of ageing and multimorbidity is the existence of health disorders outside conventional diagnostic classifications that may adversely impact the quality of life and functional capacities of elderly individuals [23,24]. Among these conditions, which encompass disorders such as ambulation difficulties, sensory impairments, compromised equilibrium, vertigo, fall susceptibility, incontinence, persistent pain, delirium, cognitive decline, and frailty, researchers have focused significantly on their demonstrable effects on critical adverse outcomes, including dementia, disabilities, and mortality [23,25]. Frailty is characterised by diminished homeostatic reserves across several physiological systems, manifesting as signs of atrophy, reduced speed, and weakness, and may be seen as the ultimate common route in the progression of multimorbidity. Cognitive impairment is a common condition among the aged, and its advancement to dementia is exacerbated in those with cardiovascular problems, particularly when accompanied by melancholy. The detrimental impact on health is more pronounced when frailty coexists with cognitive impairment [28].

The aforementioned instances demonstrate the intricate links of physical ailments and psychological, cognitive, and functional deficits in older individuals. The complexity increases further when treatment effects are included into the analysis of multimorbidity. Pharmacologic therapy, although advantageous, may potentially be detrimental due to unexpected drug-disease interactions, particularly when administered for numerous purposes. Many symptoms and functional impairments in the elderly are indeed attributable to pharmaceutical interventions [24].

Comprehending this complexity is difficult, since the examination of individuals with multimorbidity must extend beyond clearly defined groups, such as illness classifications. Due to the inadequacy of standard statistical approaches in stratifying these patients, a more complete study framework is necessary [14].

A potential remedy seems to reside in the AI methodologies of machine learning (ML) and Big Data (BD) technologies, which have yielded successful outcomes in addressing intricate challenges across several domains of human endeavour, including industry, banking, and marketing [29,30].

The purpose of this review is to summarize the limitations of current approaches to data analysis and to present the potential and shortcomings of alternative methods in multimorbidity research. The new publication by Hassaine et al. offers a comprehensive examination of methodological advancements in detecting multimorbidity-related patterns concealed within data sequences in electronic health records (eHRs) and techniques for monitoring the temporal progression of these patterns [31]. Conversely, our summary illustrates the perception of multimorbidity research from the perspective of medical experts.

This paper seeks to enhance the comprehension of medical laypersons regarding fundamental ML/BD research methodologies and offers strategies to surmount obstacles to the wider application of these techniques in multimorbidity research, ultimately aiming to elevate the quality of medical practice. The authors contemplate these matters informed by their professional experiences and underscore the necessity for enhanced collaboration between medical professionals and data scientists during the intricate problem-solving process, encompassing problem definition, data and method selection, and interpretation of research findings.

The Moment for a Paradigm Shift in Multimorbidity Research

In the traditional research methodology of the medical domain, the range of research enquiries is confined to those for which responses can be derived using established statistical techniques. Data analysis is guided by a clearly articulated hypothesis that is to be validated or refuted (a hypothesis-driven methodology). This strategy requires a well documented knowledge base and a data collection gathered in accordance with stringent criteria [32].

Multiple regression models (MLR) in classical statistics serve as fundamental predictive tools, characterised by a specified and static model structure throughout the modelling process [30]. The MLR models rely on the assumptions of independence among input variables, linearity between dependent and independent variables, normality of residuals (indicating a balanced data distribution), and the absence of endogenous (confounding) factors [33]. The stringent regulations governing the models limit the range of research enquiries and the categories of data permissible for study. These models are unsuitable for issues that do not conform to linear models or include a substantial number of variables [34]. Classical statistical approaches alone are unable to handle concerns of multimorbidity, where components are connected within a complicated network. The emergent qualities of components, rather than just their quantity and structure, may be a crucial driver of the result.

Assume we want the theoretical paradigm shift in multimorbidity research to achieve practical application. In this instance, it is essential to alter the methodology for addressing issues in medical research [35].

Reductionism posits that cause-effect interactions in the actual world may be articulated via a finite set of logical principles and fixed mathematical models [36]. This notion posits that to comprehend the system, it must be deconstructed into its components and then examined. Scientific thinking depends on robust logic and explicit assumptions, therefore eliminating inconsistencies or ambiguities. Biological systems function as complicated systems. Within a multifaceted system,

Its properties arise from the interactions among its components. The unique phenomena resulting from these interactions include spontaneous order (self-organization) and non-linearity interrelationships (alterations in one entity do not consistently correlate with changes in another entity), redundancy (the presence of multiple complementary pathways), feedback loops (a sequence of cause-and-effect that precludes definitive conclusions regarding causal relationships), and a significant degree of adaptability (functionality). It is essential to recognise that several distinct modalities influence a choice [38].

Research in molecular science and epidemiological findings is increasingly supporting a comprehensive perspective on ageing, chronic illness progression, and multimorbidity. This process may be shown by a spectrum of trajectories that vary in the dynamics of chronic illnesses and the accumulation of functional impairment over time [39]. The individual's place in the trajectory is shaped by the interaction of internal and external variables, including genetic, environmental, social, and lifestyle elements, as well as the dynamics of their evolution over time.

This perspective aligns with the concept of complexity in biological systems, wherein ageing is characterised as a gradual deterioration of the various communication pathways that interlink organs, control systems, and regulatory loops, facilitating the flow of information among them [40]. The interruption in these communication channels correlates with a reduction in the body's functional capacities and the manifestation of phenotypes, shown by the emergence of age-related illnesses, impairments, and frailty [36,40]. Despite their heterogeneity, elderly individuals exhibit considerable similarities owing to common diseases.

The use of scientific reasoning within the framework of complex systems is anticipated to enhance our capacity to draw conclusions about phenomena related to multimorbidity, notwithstanding the insufficient understanding of the interrelationships among the system's components. This mode of reasoning considers several factors when drawing conclusions and use the concepts of "chance/probability" instead of "causality/determination" [36]. In the pursuit of solutions to complicated issues, many theoretical components, mixed techniques, and multidisciplinary approaches may be integrated. The selection of methodologies is a significant aspect of the researcher's responsibilities and is contingent upon their knowledge and intuition, which is, to a certain degree, subjective.

Challenges and Approaches of Machine Learning and Big Data in Chronic Disease and Multimorbidity Research

The advent of innovative technologies in medicine and healthcare over recent decades, including digital imaging techniques and molecular biology diagnostics, along with the implementation of patient registries and electronic health records in numerous European countries and beyond, has resulted in a significant increase in the volume and complexity of data in both medical research and clinical practice. The traditional research methodology is no longer enough to address the issues of data analysis. The methodologies and techniques from machine learning and big data in artificial intelligence approaches have been evolving to provide alternative solutions [29,30] (Table 1).

The ML/BD analytical methodology facilitates the disruption of the paradigm shift in medical research and healthcare towards precision medicine [45].

The conventional expression of the association rule is IF X THEN Y, where X and Y are subsets of a complete set of objects [47]. ARM is a favoured mining technique because to its clearly interpretable outcomes presented as a series of rules. TARs indicate that a collection of things often co-occurs with another collection of goods within the same transactions throughout a designated time period [48]. Logistic Regression (LR) is a technique for modelling the likelihood of a binary outcome. It seeks to identify the most suitable model that elucidates the link between the dichotomous variable (dependent variable—comprising data encoded as 1/0, TRUE/FALSE).

yes/no, etc., and a series of independent variables.

Naive Bayes necessitates a minimal quantity of training data to ascertain the parameters essential for classification [50]. It utilises the probability of each characteristic associated with each class to build a forecast. Naive Bayes streamlines probability calculations by supposing that the likelihood of any attribute corresponding to a certain class value is independent of all other characteristics.

To generate a forecast, it computes the probabilities of the instance belonging to each class and picks the class with the greatest probability.

Decision Trees (DTs) are structured as flowchart-like trees, whereby each non-terminal node signifies a test on an attribute, each branch denotes a test result, and leaf nodes indicate target classes or class distributions, as seen in Figure 1. Each decision tree method employs distinct splitting criteria, such as information gain or Gini index, to generate branches and leaves [51]. The new record traverses the branches according to specified requirements and concludes at the leaf node, beyond which no more branching may occur. This leaf node assigns a target class to the new record with appropriate accuracy or other criteria.

Conclusion

Multimorbidity represents one of the most significant challenges facing modern healthcare systems due to the complex interactions between multiple chronic diseases, functional impairments, and psychological conditions. Traditional healthcare models and research approaches, which are primarily focused on single diseases, are often insufficient for addressing the diverse and interconnected needs of patients with multimorbidity. As populations continue to age and the prevalence of chronic diseases increases, there is a growing need for more comprehensive and patient-centered approaches to diagnosis, treatment, and care management.

Artificial intelligence, particularly machine learning and big data analytics, offers promising opportunities to better understand the complexity of multimorbidity. These technologies can analyze large and diverse healthcare datasets, identify hidden disease patterns, and support predictive modeling for improved clinical decision-making. AI-driven approaches can also contribute to personalized treatment strategies, early risk detection, and more efficient healthcare delivery.

However, the successful integration of AI into multimorbidity research and clinical practice requires several important developments. These include the standardization and integration of electronic health records, the development of robust data infrastructures, and stronger collaboration between medical professionals and data scientists. Ethical considerations, data privacy, and transparency in AI-based decision systems must also be carefully addressed to ensure patient safety and trust.

In conclusion, adopting AI-supported research frameworks and embracing interdisciplinary collaboration can significantly improve the understanding and management of multimorbidity. By integrating advanced analytical technologies with clinical expertise, healthcare systems can move toward more effective, personalized, and sustainable care for patients with complex health conditions.

References

[1] Banik, S. and S. Dandyala. (2019) Automated vs. Manual Testing: Balancing Efficiency and Effectiveness in Quality Assurance. International Journal of Machine Learning Research in Cybersecurity and Artificial Intelligence. 10(1): 100-119.

[2] Chennareddy, R. K., & Sethuraman, P. (2023). Enterprise and RAN-Aware Data and Analytics Platforms for Mission-Critical and Low-Latency Digital Services. *International Journal of Emerging Trends in Computer Science and Information Technology, 4*(4), 184-192.

[3] Sethuraman, P. (2023). Implicit Channel Inference Techniques for Pilotless OFDM Reception in Next-Generation Wireless Systems. *International Journal of Emerging Research in Engineering and Technology, 4*(1), 143-152.

[4] Chennareddy, R. K., & Sethuraman, P. (2024b). Data and Analytics Workflows for Decision Systems Enabled by Learning-Based RAN Intelligence across Distributed

Computing Environments. *International Journal of Emerging Trends in Computer Science and Information Technology, 5*(2), 149-158.

[5] Sethuraman, P., & Chennareddy, R. K. (2023a). AI-Based Fraud Detection and Prevention at the Radio Access Network: Architectures and Mechanisms for Financial Wireless Service. *International Journal of Artificial Intelligence, Data Science, and Machine Learning, 4*(4), 132-141.

[6] Jaladi, D. S., & Vutla, S. (2023b). Revolutionizing Diagnostic Imaging: The Role of Artificial Intelligence in Modern Radiology. *The Metascience, 1*(1), 284-305.

[7] Chennareddy, R. K., & Sethuraman, P. (2024c). Decision-Centric Architectures for Intelligent and Networked Wireless Computing Environments Operating at Scale and Uncertainty. *International Journal of Emerging Trends in Computer Science and Information Technology, 5*(3), 150-160.

[8] Gudepu, B. K., Jaladi, D. S., & Gellago, O. (2023). How Data Catalogs are Transforming Enterprise Data Governance: A Systematic Literature Review. *The Metascience, 1*(1), 249-264.

[9] Chennareddy, R. K. (2023). Enterprise-Scale AI and Analytics Strategy for End-to-End Business Transformation across Global Organizations. *International Journal of AI, BigData, Computational and Management Studies, 4*(3), 134-145.

[10] Sethuraman, P., & Chennareddy, R. K. (2024). RAN-AI Architectures Supporting Personalized Customer Interaction and Virtual Assistance in Banking Services. *American International Journal of Computer Science and Technology, 6*(6), 57-66.

[11] Jaladi, D. S., & Vutla, S. (2023a). Brainy: An Intelligent Machine Learning Framework. *International Journal of Acta Informatica, 2*(1), 219-229.

[12] Chennareddy, R. K., & Sethuraman, P. (2024a). AI-Enabled Data-Driven Decision Frameworks for Enterprise Platforms and Tactical Defense Wireless Networks. *American International Journal of Computer Science and Technology, 6*(4), 39-49

[13] Sethuraman, P., & Chennareddy, R. K. (2023b). System-Level Design and Orchestration of Large-Scale Cellular Access Networks for Regulatory-Compliant Financial Services. *International Journal of Emerging Research in Engineering and Technology, 4*(3), 140-150.

[14] Jaladi, D. S., & Vutla, S. (2024a). Machine Learning Techniques for Analyzing Large-Scale Patient Databases. *International Journal of Modern Computing, 7*(1), 181-198.

[15] Chennareddy, R. K., & Sethuraman, P. (2025). Translating Artificial Intelligence into Scalable Healthcare Delivery through Adaptive Decision Capabilities and Wireless-Aware System Intelligence. *International Journal of AI, BigData, Computational and Management Studies, 6*(3), 89-96.

[16] Jaladi, D. S., & Vutla, S. (2024b). The Role of Artificial Intelligence in Modern Medicine. *The Metascience, 2*(4), 96-106.

[17] Sethuraman, P., & Chennareddy, R. K. (2022a). Intelligent Vehicular Traffic Flow Prediction Using Learning-Based Spatio-Temporal Models for Data-Driven Wireless Transportation and Urban Analytics Systems. *International Journal of Emerging Trends in Computer Science and Information Technology, 3*(2), 111-121.

[18] Gudepu, B. K., & Jaladi, D. S. (2022a). Data Discovery and Security: Protecting Sensitive Information. *International Journal of Acta Informatica*, *1*(1), 176-187.

[19] Gudepu, B. K., & Jaladi, D. S. (2022b). Why Real-Time Data Discovery is a Game Changer for Enterprises. *International Journal of Acta Informatica, 1*(1), 164-175.

[20] Jaladi, D. S., & Vutla, S. (2022a). Artificial Intelligence's Influence on Design: A New Era of Creative Collaboration. *International Journal of Acta Informatica, 1*(1), 188-198.

[21] Jaladi, D. S., & Vutla, S. (2022b). Medical Decision-Making with the Help of Quantum Computing and Machine Learning: An In-Depth Analysis. *International Journal of Acta Informatica, 1*(1), 199-215.

[22] Sethuraman, P., & Chennareddy, R. K. (2022b). Machine Learning Assisted Design of Wireless Access Systems for Reliable and Low-Latency Financial and Smart Commerce Services. *International Journal of Artificial Intelligence, Data Science, and Machine Learning, 3*(4), 133-142.

[23] Chennareddy, R. K. (2021). Designing Data and Analytics Ecosystems for High Volume Transaction Processing Applications. *International Journal of AI, BigData, Computational and Management Studies, 2*(2), 95-106.

[24] Gudepu, B. K., & Jaladi, D. S. (2021). GDPR Compliance Challenges and How to Overcome Them. *International Journal of Modern Computing*, *4*(1), 61-71.

[25] Jaladi, D. S., & Vutla, S. (2021a). Exploring the Current Landscape and Applications of Artificial Intelligence in Healthcare. *The Computertech*, 28-38.

[26] Gudepu, B. K., & Jaladi, D. S. (2018a). The Role of Data Profiling in Improving Data Quality. *The Computertech*, 21-26.

[27] Goriparthi, R.G. (2022) Interpretable Machine Learning Models for Healthcare Diagnostics: Addressing the Black-Box Problem. *Revista de Inteligencia Artificial en Medicina. 13*(1): 508-534.

[28] Jaladi, D. S., & Vutla, S. (2021b). Quantum AI: Accomplishments and Obstacles in the Convergence of Quantum Computing and Artificial Intelligence. *International Journal of Modern Computing, 4*(1), 86-95.

[29] Jaladi, D. S., & Vutla, S. (2017). Harnessing the Potential of Artificial Intelligence and Big Data in Healthcare. *The Computertech*, 31-39.

[30] Gudepu, B. K., & Jaladi, D. S. (2018b). The Role of Data Quality Scorecards in Measuring Business Success. *The Computertech*, 29-36.

[31] Jaladi, D. S., & Vutla, S. (2018a). An Analysis of Big Data Analytics in Relation to Artificial Intelligence and Business Intelligence. *The Computertech*, 37-46.

[32] Holzinger, A., Biemann, C., Pattichis, C. S., & Kell, D. B. (2017). What do we need to build explainable AI systems for the medical domain?. *arXiv preprint arXiv:1712.09923*..

[33] Jaladi, D. S., & Vutla, S. (2018b). The Use of AI and Big Data in Health Care. *The Computertech*, 45-53.

[34] Wang, F., & Preininger, A. (2019). AI in health: state of the art, challenges, and future directions. *Yearbook of medical informatics*, *28*(01), 016-026.

[35] Sethuraman, P. (2022). Latency-Aware Scheduling and Resource Control Algorithms for Emergency and Public Safety Wireless Networks. *International Journal of Emerging Research in Engineering and Technology, 3*(4), 133-140.

[36] Jaladi, D. S., & Vutla, S. (2019a). Deploying Breiman's Random Forest Algorithm in Machine Learning. *The Computertech*, 45-57.

[37] Chennareddy, R. K. (2020). Engineering Intelligence Systems Using Big Data and Cloud Architectures for Modern Data Intensive Applications. *International Journal of AI, BigData, Computational and Management Studies, 1*(2), 41-50.

[38] Jaladi, D. S., & Vutla, S. (2019b). Revolutionizing Healthcare Through Quantum Computing: Insights and Future Directions. *International Journal of Modern Computing, 2*(1), 60-83.

[39] Goriparthi, R.G. (2022) Interpretable Machine Learning Models for Healthcare Diagnostics: Addressing the Black-Box Problem. *Revista de Inteligencia Artificial en Medicina. 13*(1): 508-534.

[40] Jaladi, D. S., & Vutla, S. (2020a). Leveraging Data Mining to Innovate Agricultural Applications. *International Journal of Modern Computing, 3*(1), 34-46.

[41] Kim, D. W., Jang, H. Y., Kim, K. W., Shin, Y., & Park, S. H. (2019). Design characteristics of studies reporting the performance of artificial intelligence algorithms for diagnostic analysis of medical images: results from recently published papers. *Korean journal of radiology*, *20*(3), 405-410.

[42] Jaladi, D. S., & Vutla, S. (2020b). Machine Learning Demystified: Concepts, Algorithms, and Use Cases. *The Computertech*, 1-12.

[43] Sethuraman, P. (2025). A Safety-Constrained Reinforcement Learning Framework for Scheduling with Latency-Tail Guarantees in Industrial URLLC. *International Journal of Artificial Intelligence, Data Science, and Machine Learning, 6*(3), 147-159.