Center for Management Science Research
ISSN: 3006-5291 3006-5283 Volume 3, Issue 8, 2025

BUSINESS INTELLIGENCE AND PREDICTIVE ANALYTICS FOR IMPROVING PUBLIC HEALTHCARE OPERATIONS: A SYSTEMATIC LITERATURE REVIEW

Jahanzaib Khan

Masters of Science in Project Management, University of the Cumberlands

jjkhann2@gmail.com

DOI: http://doi.org/10.5281/zenodo.21485110

Keywords
business intelligence; predictive analytics; public healthcare operations; big data analytics; hospital management; machine learning; healthcare decision support

Article History
Received: 06 October 2025
Accepted: 04 December 2025
Published: 17 December 2025

Copyright @Author
Corresponding Author: *
Jahanzaib Khan

Abstract

Public healthcare systems worldwide face mounting pressure to deliver safe, timely, and cost-efficient services under conditions of constrained budgets, workforce shortages, and rising patient demand. Business intelligence (BI) and predictive analytics have emerged as central enablers of data-driven decision-making in this environment, transforming fragmented operational, clinical, and financial data into actionable insight. This paper presents a systematic narrative review of the peer-reviewed literature published between 2015 and 2025 on the application of BI and predictive analytics to public healthcare operations. Following a structured search of Google Scholar and allied academic databases, forty studies were selected and synthesized around four themes: (1) the conceptual architecture of healthcare BI systems, including data warehousing, extract-transform-load pipelines, and visualization layers; (2) predictive analytics techniques, ranging from classical statistical forecasting to ensemble machine learning and deep learning; (3) operational application domains, namely patient flow and bed management, workforce and staffing optimization, supply chain and inventory management, disease surveillance, financial and fraud analytics, and clinical performance dashboards; and (4) the organizational, technical, and ethical barriers that constrain adoption, particularly in resource-limited and public-sector settings. The review finds consistent evidence that BI-enabled predictive analytics improves forecasting accuracy for admissions and resource demand, reduces preventable readmissions, shortens waiting times, and strengthens managerial decision-making, while also revealing persistent challenges around data quality, interoperability, governance, workforce analytics literacy, and equitable access to these technologies in low- and middle-income health systems. The paper concludes with an integrated conceptual framework linking data infrastructure, analytic capability, and organizational readiness, and proposes an agenda for future research emphasizing explainable artificial intelligence, real-time analytics, and governance frameworks tailored to public healthcare operations.

1. INTRODUCTION

Public healthcare systems find themselves in an operationally very challenging position: they have to ensure that access to care is equitable, keep costs under a limited public budget, and ensure the quality of care in an environment where demand is volatile and often unpredictable. Unplanned surges in EDs, limited beds in Ins wards,

prioritizing stock-outs and waste at pharmacies and central stores, and chronic staff shortages among workforce planners with fluctuating acuity of patients are all common problems in hospitals.Unplanned surges in EDs, limited beds in Ins wards, balancing stock-outs and waste at pharmacies and central stores, and dealing with chronic staff shortages, while having fluctuating acuity of patients, are all common issues in hospitals. Over the last decade, these operation pressures have been exacerbated by the ageing of populations, increasing burden of chronic diseases and growing demands for quality of public health services, despite the limited financial and human resources available in many public health systems (Cozzoli et al., 2022; Kamble et al., 2019).

Previously, healthcare operations management depended on retrospective reporting, manual audit and on the tacit knowledge of clinical and administrative personnel. Staffing rosters, allocation of beds, or procurement decisions were often made in response to a crisis, rather than in anticipation. With the advent of electronic health records (EHRs), hospital information systems, laboratory information systems, and Internet of Things (IoT) equipped medical devices, there is an unprecedented volume, velocity, and variety of healthcare data (Belle et al., 2015; Dash et al., 2019). In response to this data deluge, business intelligence (BI) and predictive analytics have become the primary technologies that enable the healthcare manager to gain insight from this data by turning data into a structured, actionable, and proactive approach to management, instead of a reactive one (Wang, Kung, & Byrd, 2018).

Business intelligence is a generic term for technologies, applications and practices that are designed to collect, integrate, analyse and present business information in the best way for decision making (Ain, Vaia, DeLone, & Waheed, 2019; Sun, Sun, & Strang, 2018). In the healthcare industry, BI systems can be used to collect data from clinical, financial, and administrative sources and store it in data warehouses, where it is then analyzed to create data dashboards and key performance indicators (KPIs) that can be used to track patient outcomes, financial results, and staff performance. Predictive Analytics goes a step further by using statistical modeling and machine learning algorithms on historical and real-time data to predict future events like patient admission, readmission risk, disease outbreaks, and resource needs (Wang, Kung, Wang, & Cegielski, 2018; Chen, Hao, Hwang, Wang, & Wang, 2017). BI and predictive analytics are a continuum of hindsight (what happened), insight (why it happened), and foresight (what is likely to happen and what should do about it).

These technologies can be used outside of the private or academic medical centers. Structural challenges are being solved through the application of BI and predictive analytics in public healthcare systems, national health services, and government-run hospital networks, respectively, addressing the surges in demand in certain seasons, the allocation of scarce specialist staff, tracking disease outbreaks across regions, and ensuring transparency and accountability for public spending on health (Kumar et al., 2025; Muhunzi, Kitambala, & Mashauri, 2024). The COVID-19 pandemic served as a stark reminder of the power and powerlessness of predictive analytics in public health infrastructure when health care systems around the world scrambled to create forecasting models for bed and ventilator usage in the light of overwhelming uncertainty and limited data (Ghobadi et al., 2024).

Given the current trend, the academic literature addressing BI and predictive analytics in healthcare is still scattered across fields, such as health informatics, operations research, information systems and public health. Existing reviews have focused on either big data analytics in general (Khanra, Dhir, Islam, & Mäntymäki, 2020; Raja, Mukherjee, & Sarkar, 2020), or particular clinical prediction tasks such as patient 'readmission risk' (Artetxe, Beristain, & Graña, 2018; Chen, Madanian, Airehrour, & Cerny, 2022), or on a specific organizational management question (Cozzoli et al., 2022; Singh, Agrawal, Sahu, & Kazancoglu, 2023), but not on the integration of how these technologies change the non-clinical managerial and logistical aspects of care that shape how public health services provide care: patient flows, staffing, supply chain, disease

surveillance, financial and performance management.

To fill this gap, this review compiles literature published during the last ten years (from 2015 to 2025) from peer-reviewed sources that discuss how BI and predictive analytics are applied in public healthcare settings. The three-fold objective of the review is: to map the conceptual architecture and analytic techniques that underlie current healthcare BI systems; to synthesize the empirical evidence on the operational domains in which these technologies have been introduced and the outcomes they have produced; and to identify the organizational, technical and ethical barriers that still limit the adoption of these systems, especially in public and resource-constrained health systems. A comprehensive framework and an agenda for future research of potential value to health services researchers and practicing health care managers concludes the review..

2. Review Methodology

This study has a systematic narrative review design, a structured search and screening process as in systematic reviews and a thematic cross-disciplinary synthesis as in narrative reviews. This hybrid method was chosen because the topic has been addressed in a body of work that is spread across a number of different literatures (information systems, operations research, health informatics, and public health policy) and are inconsistently indexed across the different bibliographic databases.

2.1 Search Strategy

Structured searches were done, mainly on Google Scholar and then on specific search engines like PubMed/MEDLINE, ScienceDirect, IEEE Xplore, SpringerLink, Wiley Online Library and ACM Digital Library, and literature was captured that was indexed in information-systems and health-sciences outlets. The search included terms from three conceptual clusters with the use of Boolean operators: (1) technology terms: business intelligence, predictive analytics, big data analytics, machine learning, data warehousing, dashboards; (2) setting terms: healthcare operations, hospital management, public health, public hospital, health system; and (3) outcome terms: patient flow, bed management, staffing, supply chain, readmission, resource allocation, disease surveillance, performance management. The search was limited to English-language, peer-reviewed journal articles, conference proceedings and systematic reviews from January 2015 to mid-2025.

2.2 Inclusion and Exclusion Criteria

Studies were considered for inclusion if they (a) focused on use, design, evaluation, or adoption of a business intelligence/predictive analytics technology in an operational context within healthcare or public health; (b) were published between 2015 and 2025 in a peer-reviewed journal; a peer-participation edited volume or an indexed conference proceeding; and (c) provided empirical results, a systematic or narrative synthesis of prior empirical results or a well-developed conceptual or architectural framework. Studies that included only clinical diagnostic or therapeutic algorithms without an operational or managerial aspect, were commercial or vendor white papers that were not peer reviewed, and were conference versions of later journal publications were excluded.

2.3 Screening and Selection

Titles and abstracts from the search were initially assessed for topical relevance and then full text was judged for its relevance to the inclusion criteria and shortlisted titles. These reference lists of key systematic reviews found in the search were scanned by hand to identify further primary studies, a snowballing method which uncovered a number of frequently cited foundational papers that were not identified in the initial keyword search (e.g., Khanra et al., 2020; Cozzoli et al., 2022; Borges do Nascimento et al., 2021). The result of this process was a final set of 40 studies, covering systematic and narrative reviews of healthcare BI and big data (n = 15), empirical studies of predictive models for various operational tasks like readmission, patient flow and workforce forecasting (n = 12), studies of supply chain and disease-surveillance analytics (n = 6), and studies on adoption barriers, governance and organizational readiness (n = 7).

2.4 Data Extraction and Synthesis

Data for each study included were the year of publication, methodological approach, healthcare operational domain addressed, analytic technique used, and the main findings. Thematic synthesis approach was used due to the heterogenicity of the study designs from large N systematic review to single site predictive modeling studies. Results were inductively categorized into four thematic areas that have been followed in the rest of this paper: conceptual foundations, operational applications, reported benefits, and adoption barriers.

2.5 Limitations of the Review Methodology

Like all narrative synthesis, the current review is subject to both selection and publication bias, with reporting of successful analytics implementation efforts being more likely to be published than reporting of failed or abandoned efforts. Additionally, because of the use of English language indexed academic literature, grey literature, government evaluation reports, and non-English language public health systems literature from Africa, Latin America, and parts of Asia may be under-represented, which is discussed as a substantive finding in Section 6.

3. Conceptual Foundations of Business Intelligence and Predictive Analytics in Healthcare

3.1 Defining Business Intelligence and Predictive Analytics

The most common definition of business intelligence is the set of technologies, architectures and practices that facilitate the collection, integration, analysis and presentation of business information for supporting decision making based on facts (Ain et al., 2019). In healthcare, BI systems are most commonly used in three different levels of analysis: descriptive analytics, which involves summarizing the data and answering what has happened questions (e.g., average length of stay in the last month); diagnostic analytics, which seeks answers to why it happened questions (e.g., which departments or shifts have a higher length of stay average than others); and predictive analytics, which provides answers to what is likely to happen next questions (e.g., what is expected to happen in the coming week) (Wang, Kung, & Byrd, 2018). The fourth tier is prescriptive analytics which recommends specific actions to get to a desired outcome, like an optimal staff roster or reorder quantity showing that the prescriptive tier is now becoming mainstream within healthcare BI platforms, in addition to predictive models (Sun et al., 2018).

Healthcare predictive analytics utilizes a broad range of methodologies. Time-series techniques like autoregressive integrated moving average (ARIMA) models and exponential smoothing are common techniques that have been used for forecasting patient volumes and inventory demand, and are widely used because of their simplicity of interpretation and limited amount of data needed. When dealing with classification problems, like readmission risk or no-show prediction, machine learning techniques such as logistic regression, decision trees, random forests, gradient boosting machines (like XGBoost), and support vector machines are used extensively and generally much more effective than traditional statistics-based methods when provided with a large number of sufficient features to achieve good performance (Chen et al., 2022; Artetxe et al., 2018). As a result of their ability to model complex temporal dependencies, deep learning architectures like recurrent neural networks, as well as long short-term memory (LSTM) networks, are gaining popularity in the analysis of sequential clinical and operational data, such as vital-sign trajectories, or day-by-day time series of admission times (Chen et al., 2017).

3.2 The Architecture of Healthcare BI Systems

In the healthcare field, the general architecture of healthcare BI has four layers that are interdependent (Cozzoli et al., 2022; Dash et al., 2019). Data source layer is made up of electronic health records, hospital information systems, lab and radiology information systems, pharmacy and supply chain systems, HR and payroll systems, and more recently medical devices and wearable sensors that communicate with IoT. The data integration layer uses extract-transform-load (ETL) techniques to cope with the structural and

semantic differences between all these source systems and to transform the data from multiple, inconsistent sources into a single schema. The data storage layer usually includes an enterprise data warehouse or, increasingly, a cloud-based data lake that can store structured and unstructured data at scale optimized for performance of analytical queries instead of transactional processing. Last but not least, the analytics and presentation layer refers to the statistical and machine learning models used for forecasting, as well as the dashboards, scorecards and visualization tools that bring insights to the attention of clinical and administrative decision-makers (Wang et al., 2018).

One theme common in the literature is the complexity of the analytics layer is often surpassed by the shortcomings of the data integration and data governance layers. Gathering value from the healthcare big data is hindered by its fragmented, siloized, and inconsistently managed data infrastructure, rather than by the capabilities of analytics tools themselves, as Cozzoli et al. (2022) point out in their synthesis of the literature on healthcare BI (Khanra et al., 2020; Kamble et al., 2019; Galetsi, Katsaliaki, & Kumar, 2019).

3.3 From Business Intelligence to Predictive and Prescriptive Analytics

The evolution from BI toward predictive and prescriptive analytics reflects a broader shift in health Such a transition from BI to predictive and prescriptive analytics represents a general trend in the evolution of health information systems, from health management information systems, which are mainly designed for regulatory reporting, to decision-support systems that are integrated into the operational workflows (Panch, Szolovits, & Atun, 2018; Duan, Edwards, & Dwivedi, 2019). Duan et al. (2019) describe this progression as a maturity ladder, starting with descriptive BI (ability to report reliably on past performance), to diagnosis BI (ability to diagnose drivers of the performance), to predictive BI (ability to forecast future states based on predictive models), and to prescriptive analytics (ability to embed the prediction into automated or semi-automated decision rules, e.g., dynamic staffing algorithms or automated reorder rules in supply chain systems). However, empirical evidence indicates that most public health institutions continue to be entrenched in the descriptive and diagnostic levels, with the use of predictive analytics being limited, and prescriptive analytics being relatively infrequent, except in large academic health-care organizations and in well-funded national health systems (Singh et al., 2023; Popovič et al., 2018).

4. Applications in Public Healthcare Operations

The literature we reviewed is focused on six main areas of operation where BI and predictive analytics have been used in the public and hospital-based health care fields: patient flow and bed control; workforce and staffing optimization; supply chain and inventory management; disease monitoring and public health planning; financial management and fraud detection; and clinical performance dashboards and decision support. They are explained in turn below.

4.1 Patient Flow and Bed Management

Patient flow the flow of patients through admission, treatment and discharge is one of the most widely studied examples of predictable analytics in healthcare business, as inefficiencies at each step of the process feed into other issues, such as an overcrowded ED, cancelled elective surgeries and prolonged LOS. Using the electronic health record data from more than 63,000 admissions at a large academic hospital, Bertsimas et al. (2022) created an interpretable patient representation to predict short-term discharges, identify long-stay patients, forecast discharge destination, and forecast flows to and from intensive care units, with prediction accuracy exceeding 80 percent. Importantly, the authors built these predictions into a real-time dashboard for hospital staff to make decisions about bed placement on a daily basis, showing how predictive analytics can be taken from a research topic to an operational decision-making tool.

Such methods have been adapted to the environment of the ED, where machine learning algorithms that are trained on past arrival trends, seasonal and calendar effects, and current patient counts use real-time data to predict potential

surges in patient numbers, allowing for proactive staffing and bed allocation at the ED. In response to the COVID-19 pandemic, hospital systems quickly began using hybrid forecasting models, which integrated autoregressive time-series models, machine learning, and epidemiological compartmental models, to forecast short-term need for intensive care unit and ward beds with errors of up to just four to nine percent at one- and two-week time horizons, aiding real-time capacity planning decisions (Ghobadi et al., 2024). Ghobadi and colleagues' interactive decision-support dashboard, which was implemented by a large health system in the U.S. during the peak of the pandemic, is illustrative of the combination of predictive modelling, optimization, and participatory design, which enables hospital leaders to test different capacity-management scenarios interactively, rather than just relying on static predictions.

There are many parallel, large-scale studies on hospital readmission prediction because it is a huge cost in public healthcare systems and prevents unnecessary deaths and other quality of care issues. Artetxe, Beristain and Graña (2018) performed an early systematic review of predictive models for hospital readmission risk, identifying the transition from traditional logistic-regression based risk scores to machine learning approaches that could take advantage of the high dimensional electronic health record data. In recent systematic reviews, machine learning models, specifically gradient-boosted tree ensembles and random forest models, have been shown to outperform conventional regression-based scoring tools, like the LACE index, with reported improvements of the area under the receiver operating characteristic curve (AUC) from about 0.66 for conventional scores to greater than 0.80 for machine models (Chen et al., 2022; Huang et al., 2021; Mahmoudi et al., 2020). Yet, there are questions about the transferability of these models to other institutions, class imbalance in readmission rates, and the possibility that models based on past data could perpetuate existing inequities in access to care.

4.2 Workforce and Staffing Optimization

Nurse/physician scheduling is one of the most operationally impactful uses of predictive analytics in public healthcare because of staff shortages, staff turnover, and varying acuity of patients and the nature of the care they are seeking. Patient census and acuity forecast models include time-series models, machine learning regressors and discrete-event simulation, and these predictions are then translated to staffing needs that proactively anticipate rather than respond to gaps in staffing after the fact. Finally, reviews of nurse-demand forecasting methods all stress that time-series models are especially useful for quantifying recurring seasonal and cyclical patterns in nurse demand; machine learning models are more effective when large amounts of granular historical nurse demand data are available; and simulation-based approaches have the advantage of flexibility in testing alternative scenarios when there is high uncertainty in nurse demand.

More sophisticated workforce analytics look at demand forecasting combined with optimization-based scheduling and the evaluation of performance as one decision support architecture. These usually include three interconnected modules: a demand-forecasting module that relies on ensemble machine learning models like long short-term memory networks, gradient boosting, and random forests to predict admission and staffing levels; a scheduling-optimization module that produces 'fair' and constraint-satisfying schedules according to legal, contractual, and skill-mix requirements; and a performance-evaluation module that integrates structured operational indicators and natural-language-processing analysis of staff feedback. These integrated systems are claimed to eliminate the need for expensive agency or temporary staff, help prevent employee attrition with early warning of those at risk and minimise waiting times in the Emergency Department through adequate staff cover during predicted surge times.

4.3 Supply Chain and Inventory Management

Healthcare supply chains are susceptible to stock-outs of essential medicines and consumables, wastage due to overstocking and expiry,

fragmented procurement processes, and inefficiencies. Predictive analytics in healthcare supply chain management is mainly being used with demand forecasting models such as ARIMA, exponential smoothing, regression, and, more recently, neural network algorithms, which predict demand for consumables and pharmaceuticals based on historical usage, seasonality and, increasingly, real-time sensor data from the Internet of Things (IoT) smart storage system. Abu Zwaida et al. (2021) show how to apply inventory-optimization models particularly developed for the prevention of drug stocks-outs in hospital supply chains, and how to combine predictive demand models with the optimization of reorder points to balance the risk of stock-outs with the stock carrying cost.

The COVID-19 pandemic revealed the significant weaknesses in the healthcare supply chain which led to an increased focus on the application of predictive and prescriptive analytics in improving supply chain resilience in highly volatile demand and disrupted global supply chains (Sharma, Gupta, & Jha, 2020). Although not specific to healthcare, broader supply-chain analytics research also offers methodological grounding: systematic reviews of the use of AI in supply-chain management show that supply-chain machine-learning demand sensing tools, which replace the traditional statistical forecasting tools, can have substantial reported impacts on the forecasting accuracy, reductions in inventory holding costs, and avoidance of stock-outs (Toorajipour et al., 2021).

4.4 Disease Surveillance and Public Health Planning

In addition to hospital use, predictive analytics is now used in public health surveillance, adding to traditional epidemiological monitoring, by operating machine learning models that combine a variety of structured and unstructured data sources, such as lab reports, syndromic surveillance data, environmental and climate data, mobility data, and social media signals, to help detect and anticipate outbreaks of infectious diseases earlier than conventional case-reporting systems. Outbreak models are still mostly based on regression, Bayesian inference, and time-series methods, but the use of machine learning classifiers like random forests and support vector machines, and deep learning architectures like LSTM networks, to detect complex, non-linear patterns in surveillance data and to drive near-real-time early-warning systems by public health agencies and platforms, including those held by national centers for disease control, is growing.

The direct application of these predictive surveillance tools is through anticipatory resource allocation, with predictive estimates of the severity of outbreaks and geographic spread informing public health operations before people start getting sick (or dying) in numbers.The direct application of these predictive surveillance tools is through anticipatory resource allocation – for example, knowing in advance of an outbreak how severe it will be and how far it will travel, allows health authorities to make sure vaccines, personal protective equipment, and surge staffing are in place before people begin to get sick (or die) in large numbers. Meanwhile, reviewers warn of the high data quality/ completeness requirements for surveillance-analytics systems, sharp performance declines in predicting performance outside the population and setting where they were trained and developed, and the need for intentional governance attention to ethical and equity concerns such as the risk of predictive resource allocation systematically discriminating against under-surveilled or marginalised populations (Borges do Nascimento et al., 2021).

4.5 Financial Management, Cost Control, and Fraud Detection

Financial and revenue-cycle analytics is a significant application area for BI for public healthcare payers and providers, who are constantly being challenged by fiscal constraints. Descriptive and predictive BI dashboards help to monitor expenditure per patient, track reimbursement claims and alert revenue leakage due to billing mistakes, under-coding, or no charges at all. Claims data is also analyzed using predictive models to identify patterns that are associated with CWA within a claims dataset, such as billing frequency anomalies or unreasonable

service mixes to help insurers and/or public payers focus their audit resources more effectively than on random samples alone. In addition, big data analytics of financial and administrative data has been leveraged to support value-based care efforts; for example, by assigning costs to care and quality and outcome measures to identify high-cost, low-value care pathways for management (Wang & Hajli, 2017; Popovič et al., 2018).

4.6 Clinical Performance Dashboards and Decision Support

In all the domains listed above, dashboards are the key point of interaction where BI and predictive analytics outputs are delivered to frontline managers and clinicians. Clinical and operational dashboards use KPIs (occupancy rates, average length of stay, readmissions, staffing ratios, financial indicators and more) to create interactive visual displays that can be used for routine monitoring and ad hoc investigation. Systematic reviews of clinical dashboards as audit-and-feedback or decision-support tools generally show positive but varied effects on process measures like guideline-concordant medication use and appropriate test ordering, and recommend that the positive effects are likely to be influenced by the analytics used, the design of the clinical dashboard, and the existence of accompanying organizational change-management processes, not exclusively by the analytics themselves.

Predictive Model Outputs can be incorporated directly into the interface of more advanced decision support dashboards, so that the user can interact with the outputs of the predictive models and explore alternate scenarios instead of viewing static projections. The COVID-19 hospital capacity dashboard, developed with hospital administrators, combining real-time data, predictive models, and optimization routines in an integrated interface that was used daily by hospital leadership during the COVID-19 pandemic, is a prime example of this participatory, decision-oriented design: as an example, it was technically possible and very useful that predictive analytics were tightly coupled with operational decision-making, thereby demonstrating the technical feasibility and the organizational value of such an integrated user interface (Ghobadi et al., 2024).

5. Synthesis of Reported Benefits and Performance Outcomes

There are several common themes in the benefits reported across the reviewed studies, although there is high variation in methodology, context, and measures of benefit. First, forecasting accuracy: when applied to tasks of patient flow, readmission, and disease surveillance, models beat traditional rule-based or purely retrospective reporting approaches (consistently by wide margins), and, in the case of well-specified forecasting tasks, machine learning ensembles often outperform conventional risk-scoring approaches by several percentage points, with classification accuracy or AUC values in the 0.80-0.90 range (Chen et al., 2022; Bertsimas et al., 2022). Second, operational efficiency, such as reported reductions in average length of stay, improved bed-turnover times, reduced waiting time in emergency departments, and decreased need for costly agency and temporary staffing, due to earlier and more accurate prediction of need (Ghobadi et al., 2024).

Third, cost containment: BI-driven supply chain and financial analytics are reported to help reduce inventory carrying costs, minimise stock-outs of life-saving medicines, and identify billing errors and fraudulent claims, which help to use limited public health budgets more efficiently (Abu Zwaida et al., 2021; Toorajipour et al., 2021). Fourth, quality and safety: dashboard-based decision-support tools are linked to increased compliance with clinical guidelines and more prudent use of diagnostic tests and drugs, but the impact is highly variable by study and context, and seems to be heavily influenced by implementation quality, more so than analytic sophistication.

Finally, and somewhat tentatively, a number of studies have reported organizational and strategic benefits that are more difficult to quantify but nevertheless often referred to in qualitative and case-study evidence: a better sense of managerial visibility of the performance of the system, increased accountability and transparency in reporting to the public, and improved capacity to

do scenario planning in the face of uncertainty, as experienced by the pandemic-era capacity-management dashboards in Section 4.1 (Ghobadi et al., 2024; Cozzoli et al., 2022). Combined, the conclusive evidence provided in this research suggests that BI and predictive analytics can make a significant contribution to the performance of public healthcare systems, and that the impact of these is heavily dependent on the quality of the underlying data infrastructure, and the extent to which the results of the analytics are integrated into genuine business processes, rather than merely used for reporting purposes.

6. Challenges and Barriers to Adoption

Despite the proven advantages, the literature reviewed highlights significant and persistent challenges to the successful implementation of BI and predictive analytics in public health systems. The barriers can be divided into four major groups: data quality and interoperability; governance, privacy and security; organizational and cultural readiness; and structural limitations in public-sector and developing country health systems.

6.1 Data Quality, Fragmentation, and Interoperability

Poor quality, incompleteness, and fragmentation of underlying health care data are most commonly mentioned barriers across the studies reviewed. The implementation of EHR, laboratory information systems, pharmacy information systems, and administrative databases are usually done by separate vendors, with inconsistent data models, coding systems, and update frequencies, resulting in data silos that need to be reconciled before meaningful insights can be gained from the data (Khanra et al., 2020; Kamble et al., 2019). The lack of a well-designed, structured and standardized data source was a consistently reported methodological limitation in primary studies and is a key determinant of reported predictive accuracy, with an impact on the generalizability of models between institutions in systematic reviews of big data analytics in healthcare (Borges do Nascimento et al., 2021). In some jurisdictions, standards like HL7 and FHIR have enhanced data exchange, although not all jurisdictions have good uptake, especially public hospital networks that are using information systems in their legacy state without adequate budgets for integration.

6.2 Data Governance, Privacy, and Security

Health information is sensitive, and has more stringent governance requirements than other domains of BI application. Existing research highlights the importance of regulatory and ethical safeguards for healthcare BI and predictive analytics efforts, such as protecting patient privacy, ensuring role-based access control, maintaining audits and monitoring for biases in machine learning models (Janssen, van der Voort, & Wahyudi, 2017; Mooney & Pejaver, 2018). Research that unpacks the obstacles to data-sharing in LMICs offers further insights into the governance complexity of data-sharing: complex laws and policies on data-sharing, weak institutional capacity to negotiate data-sharing agreements between public and private providers, and the lack of willingness among hospitals to make their data available without guidance from appropriate legislation.

6.3 Organizational and Cultural Barriers

One theme that has emerged consistently in the adoption research is that having technological capability is not enough to achieve value from BI and predictive analytics; having organizational readiness (which includes leadership commitment, analytics literacy among clinical and administrative staff, and a culture of making decisions using analytics) is also critical (Ain et al., 2019; Duan et al., 2019). According to a review of technology-organization-environment determinants of big data adoption in healthcare settings, organizational factors, such as perceived top management support and staff readiness, have a similar influence on the intention to adopt as purely technological factors such as system compatibility. In addition to workforce-related challenges, public healthcare systems can be further constrained by high staff turnover, insufficient protected time for employees to learn to use new analytic tools, and, in some instances, a lack of trust in algorithmic recommendations by

workers that are seen as being opaque or that invade their clinical judgment.

6.4 Structural Constraints in Public-Sector and Developing-Country Health Systems

There are a number of reviewed studies that specifically focus on the unique challenges encountered by public healthcare systems in developing countries, where infrastructure challenges amplify the generic challenges mentioned above. Muhunzi, Kitambala, and Mashauri (2024) report that major barriers to widespread adoption of big data analytics in HCPs in developing countries include the lack of trustworthy ICT infrastructure, limited human resources in data science and health informatics, and inadequate support for big data enterprise-scale data warehousing and data analytics. Cultural, regulatory and operational challenges, such as the need for a certain level of AI literacy – the practical understanding of how to interpret and utilize AI outputs – are recognized in related studies of health and public services as a unique and under-researched barrier to the operational benefits of predictive analytics, beyond the issues of data availability and technical infrastructure (Kumar et al., 2025).

The inequities in the structure of the field are particularly relevant in the context of health systems, as the evidence-base for the use of BI and predictive analytics in healthcare operations is mostly developed within well resourced academic medical centers and high-income health systems, which may limit the ability of these technologies to positively impact operational efficiency and quality of care for all health systems (Borges do Nascimento et al., 2021; Muhunzi et al., 2024).

7. An Integrated Framework for BI-Enabled Public Healthcare Operations

The paper synthesizes the conceptual and empirical literature reviewed above and proposes a framework that integrates three interdependent layers needed for BI and predictive analytics to lead to better public healthcare operations (data infrastructure layer, analytic capability layer, organizational readiness layer, and a governance layer that is shared and cascades between the three).

The data infrastructure layer includes the interoperable data sources, pipelines, and warehousing that are discussed in Section 3.2 and is a necessary but not sufficient pillar of analytics maturity; without reliable, integrated data there is no way even sophisticated predictive models can generate reliable forecasts (Khanra et al., 2020). The analytic capability layer includes the statistical and machine learning tools, the dashboarding tools (described in Section 3.3), and the emerging prescriptive optimization routines (described in Section 4), and defines the level of sophistication of insight that can be derived from the underlying data. Organizational readiness layer represents leadership sponsorship, staff analytics literacy, and workflow integration, and has been the most commonly reported in the literature as the binding constraint on realized value (Ain et al., 2019; Duan et al., 2019).

The governance layer is the cross-cutting layer that should be integrated from the very beginning of any BI or predictive analytics project, not after it has been implemented, and includes data privacy and security, algorithmic fairness and bias monitoring, and regulatory compliance (Janssen et al., 2017; Kumar et al., 2025). The framework suggests that public healthcare organizations seeking to improve operations through BI and predictive analytics should assess and invest across all four layers concurrently: infrastructure investment without organizational change-management is unlikely to translate into use; analytic sophistication without adequate governance risks eroding public trust; and organizational enthusiasm without a reliable data foundation will generate unreliable and ultimately discredited forecasts.

8. Future Research Directions

The literature reviewed suggests a number of areas where future research in the field of BI and predictive analytics can be focused on public healthcare operations.

Explainable and interpretable analytics. Future research should focus on interpretable methods in the development of machine learning models and

explainability techniques that serve as tools for clinical and administrative decision-makers to understand, believe, and appropriately challenge algorithmic recommendations, given that interpretable models can be as accurate as black-box approaches for predicting patient flow in early research (Bertsimas et al., 2022). Real-time, streaming analytics. The majority of the literature reviewed is based on batch-processed and retrospective data, with future work recommended to further develop real-time and near-real-time analytics architectures which use IoT-enabled sensors and streaming data pipelining to enable truly proactive operational management and not just analysis on a lag. Equity-focused evaluation. Future research should explicitly examine the equities implications of deployment of predictive analytics in health systems, considering whether the algorithms for resource allocation generate a bias against under-surveilled or marginalised populations, and extend the evidence base beyond well-resourced health systems (Muhunzi et al., 2024).

Governance models specific to public health. More work needs to be done to create practical, usable governance structures for data sharing, algorithmic accountability, and privacy protection – ones that are appropriate for the institutional and legal context of public health systems – and not simply copied from the private sector or from the clinical governance of AI systems (Kumar et al., 2025). Longitudinal & multi-site assessment. More longitudinal, multi-site evaluations are required to determine if the operational benefits reported on in the evidence base are lasting and if they can be translated across various institutional and national settings. Prescriptive Analytics and human decision making integration. Future research should explore how prescriptive analytics automated or semi-automated recommendations for staffing, procurement, or capacity allocation can be integrated with (not replaced by) professional judgment and the organizational and workflow conditions that facilitate this integration.

9. Conclusion

This review aims to provide a synthesised overview of the state of the art of applying business intelligence and predictive analytics to public healthcare operations, based on the study of 40 peer-reviewed studies published over the last 10 years (2015 to 2025). The evidence shows that these technologies have the potential to deliver significant and proven benefits in the areas of demand forecasting, patient flow and workforce management, supply chain resiliency, disease surveillance timeliness, and transparency of financial and performance management. The benefits of these are also not guaranteed in the literature, critically relying on the quality and interoperability of the underlying infrastructure, governance mechanisms to ensure privacy and address algorithmic bias, and organizational capacity to apply the results of the analytics to modify practice. For public healthcare systems, especially those with limited resources, the lessons are that the investment in the predictive analytics capability has to be coupled with parallel investments in data infrastructure, staffing analytics literacy, and governance structures, not just seen as a technological procurement item. The integrated framework suggested in this review with its underlying governance layer, data infrastructure and analytic capability provides a disciplined platform for investing in planning. With the continued challenges of increasing demand, staffing issues, and limited budgets in the public healthcare sector, the role of business intelligence and predictive analytics will likely stay at the core of improving the efficiency, responsiveness, and equitability of health care operations if the value and utility of these tools are embraced deliberately, transparently and with persistent focus on the issues that fit into the organizational and ethical frameworks that will make them effective.

REFERENCES

Abu Zwaida, T., Pham, C., & Beauregard, Y. (2021). Optimization of inventory management to prevent drug shortages in the hospital supply chain. Algorithms, 14(10), 27.

Ain, N., Vaia, G., DeLone, W. H., & Waheed, M. (2019). Two decades of research on business intelligence system adoption, utilization and success – A systematic literature review. Decision Support Systems, 125, 113113.

Alonso, S. G., de la Torre-Díez, I., Hamrioui, S., López-Coronado, M., Barreno, D. C., Nozaleda, L. M., & Franco, M. (2018). Data mining algorithms and techniques in mental health: A systematic review. Journal of Medical Systems, 42(9), 161.

Alotaibi, S., & Mehmood, R. (2018). Big data enabled healthcare supply chain management: Opportunities and challenges. In Proceedings of the International Conference on Smart Cities, Infrastructure, Technologies and Applications (pp. 207–215). Springer.

Artetxe, A., Beristain, A., & Graña, M. (2018). Predictive models for hospital readmission risk: A systematic review of methods. Computer Methods and Programs in Biomedicine, 164, 49–64.

Batarseh, F. A., Ghassib, I., Chong, D., & Su, P. H. (2020). Preventive healthcare policies in the US: Solutions for disease management using big data analytics. Journal of Big Data, 7(1), 38.

Belle, A., Thiagarajan, R., Soroushmehr, S. M. R., Navidi, F., Beard, D. A., & Najarian, K. (2015). Big data analytics in healthcare. BioMed Research International, 2015, 370194.

Bertsimas, D., Pauphilet, J., Stevens, J., & Tandon, M. (2022). Predicting inpatient flow at a major hospital using interpretable analytics. Manufacturing & Service Operations Management, 24(6), 2809–2824.

Borges do Nascimento, I. J., Marcolino, M. S., Abdulazeem, H. M., Weerasekara, I., Azzopardi-Muscat, N., Gonçalves, M. A., & Novillo-Ortiz, D. (2021). Impact of big data analytics on people's health: Overview of systematic reviews and recommendations for future studies. Journal of Medical Internet Research, 23(4), e27275.

Chen, M., Hao, Y., Hwang, K., Wang, L., & Wang, L. (2017). Disease prediction by machine learning over big data from healthcare communities. IEEE Access, 5, 8869–8879.

Chen, T., Madanian, S., Airehrour, D., & Cerny, T. (2022). Machine learning methods for hospital readmission prediction: Systematic analysis of literature. Journal of Reliable Intelligent Environments, 8, 49–66.

Cozzoli, N., Salvatore, F. P., Faccilongo, N., & Milone, M. (2022). How can big data analytics be used for healthcare organization management? Literary framework and future research from a systematic review. BMC Health Services Research, 22(1), 809.

Dash, S., Shakyawar, S. K., Sharma, M., & Kaushik, S. (2019). Big data in healthcare: Management, analysis and future prospects. Journal of Big Data, 6(1), 54.

Duan, Y., Edwards, J. S., & Dwivedi, Y. K. (2019). Artificial intelligence for decision making in the era of Big Data – evolution, challenges and research agenda. International Journal of Information Management, 48, 63–71.

Dwivedi, Y. K., Hughes, L., Ismagilova, E., Aarts, G., Coombs, C., Crick, T., Duan, Y., Dwivedi, R., Edwards, J., Eirug, A., Galanos, V., Ilavarasan, P. V., Janssen, M., Jones, P., Kar, A. K., Kizgin, H., Kronemann, B., Lal, B., Lucini, B., ... Williams, M. D. (2021). Artificial Intelligence (AI): Multidisciplinary perspectives on emerging challenges, opportunities, and agenda for research, practice and policy. International Journal of Information Management, 57, 101994.

Galetsi, P., Katsaliaki, K., & Kumar, S. (2019). Values, challenges and future directions of big data analytics in healthcare: A systematic review. Social Science & Medicine, 241, 112533.

Ghobadi, K., et al. (2024). An interactive decision-support dashboard for optimal hospital capacity management [Preprint]. arXiv:2403.15634.

Huang, Y., Talwar, A., Chatterjee, S., & Aparasu, R. R. (2021). Application of machine learning in predicting hospital readmissions: A scoping review of the literature. BMC Medical Research Methodology, 21(1), 96.

Imran, S., Mahmood, T., Morshed, A., & Sellis, T. (2021). Big data analytics in healthcare A systematic literature review and roadmap for practical implementation. IEEE/CAA Journal of Automatica Sinica, 8(1), 1–22.

Janssen, M., van der Voort, H., & Wahyudi, A. (2017). Factors influencing big data decision-making quality. Journal of Business Research, 70, 338–345.

Kamble, S. S., Gunasekaran, A., Goswami, M., & Manda, J. (2019). A systematic perspective on the applications of big data analytics in healthcare management. International Journal of Healthcare Management, 12(3), 226–240.

Khanra, S., Dhir, A., Islam, A. K. M. N., & Mäntymäki, M. (2020). Big data analytics in healthcare: A systematic literature review. Enterprise Information Systems, 14(7), 878–912.

Kruse, C. S., Goswamy, R., Raval, Y., & Marawi, S. (2016). Challenges and opportunities of big data in health care: A systematic review. JMIR Medical Informatics, 4(4), e38.

Kumar, R., Singh, A., Kassar, A. S. A., Humaida, M. I., Joshi, S., & Sharma, M. (2025). Adoption challenges to artificial intelligence literacy in public healthcare: An evidence-based study in Saudi Arabia. Frontiers in Public Health, 13, 1558772.

Mahmoudi, E., Kamdar, N., Kim, N., Gonzales, G., Singh, K., & Waljee, A. K. (2020). Use of electronic medical records in development and validation of risk prediction models of hospital readmission: Systematic review. BMJ, 369, m958.

Mooney, S. J., & Pejaver, V. (2018). Big data in public health: Terminology, machine learning, and privacy. Annual Review of Public Health, 39, 95–112.

Muhunzi, D., Kitambala, L., & Mashauri, H. L. (2024). Big data analytics in the healthcare sector: Opportunities and challenges in developing countries. A literature review. Health Information Management Journal. Advance online publication.

Obermeyer, Z., & Emanuel, E. J. (2016). Predicting the future big data, machine learning, and clinical medicine. New England Journal of Medicine, 375(13), 1216–1219.

Panch, T., Szolovits, P., & Atun, R. (2018). Artificial intelligence, machine learning and health systems. Journal of Global Health, 8(2), 020303.

Popovič, A., Hackney, R., Tassabehji, R., & Castelli, M. (2018). The impact of big data analytics on firms' high value business performance. Information Systems Frontiers, 20(2), 209–222.

Raja, R., Mukherjee, I., & Sarkar, B. K. (2020). A systematic review of healthcare big data. Scientific Programming, 2020, 5471849.

Ristevski, B., & Chen, M. (2018). Big data analytics in medicine and healthcare. Journal of Integrative Bioinformatics, 15(3), 20170030.

Sharma, A., Gupta, P., & Jha, R. (2020). COVID-19: Impact on health supply chain and lessons to be learnt. Journal of Health Management, 22(2), 248–261.

Singh, R. K., Agrawal, S., Sahu, A., & Kazancoglu, Y. (2023). Strategic issues of big data analytics applications for managing health-care sector: A systematic literature review and future research agenda. The TQM Journal, 35(2), 262–281.

Sun, Z., Sun, L., & Strang, K. (2018). Big data analytics services for enhancing business intelligence. Journal of Computer Information Systems, 58(2), 162–169.

Toorajipour, R., Sohrabpour, V., Nazarpour, A., Oghazi, P., & Fischl, M. (2021). Artificial intelligence in supply chain management: A systematic literature review. Journal of Business Research, 122, 502–517.

Wamba, S. F., Gunasekaran, A., Akter, S., Ren, S. J. F., Dubey, R., & Childe, S. J. (2017). Big data analytics and firm performance: Effects of dynamic capabilities. Journal of Business Research, 70, 356–365.

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

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

Wang, Y., Kung, L., Wang, W. Y. C., & Cegielski, C. G. (2018). An integrated big data analytics-enabled transformation model: Application to health care. Information & Management, 55(1), 64–79.

IEER
Institute for Excellence in Education & Research