

Integration of Big Data in Patient Safety Predictive Systems: A Systematic Literature Review

Erwan Ahmad¹, Dwi Nopriyanto², Pratiwi Christa Simamarta³, and Lailya Khusna⁴

^{1,2,3,4}Faculty of Medicine, Universitas Mulawarman, Samarinda, Indonesia

*Corresponding author: erwanahmad@fk.unmul.ac.id

ABSTRACT

Big data plays a transformative role in enhancing patient safety through the development of predictive systems based on real-time analytics and artificial intelligence (AI). This study aims to map the integration of big data technology into patient safety systems, identify thematic trends, and highlight innovations in clinical and nursing practices. A total of 50 articles published between 2014–2024 were analyzed using the PRISMA approach. The findings were classified into three main categories: Big Data in Patient Safety, Patient Risk Prediction, and AI Implementation in Healthcare. The findings indicate that the successful integration of big data in patient safety is heavily dependent on data validity and technological readiness at the point of care. To address the challenge of global disparities, future research should focus on developing predictive systems that are adaptive to local conditions in developing countries as well as strengthening interprofessional collaboration to ensure that the resulting solutions are practical, contextual, and sustainable. Several significant implementations include machine learning models to predict acute kidney injury (AKI), fall risk, ICU needs, and drug-drug interactions, with Python being the most dominant software. AI integration is also evident in nursing tasks such as risk assessment, care planning, and perioperative monitoring via wearable devices.

Keywords: Patient Safety, Big Data, Predictive Systems

1. INTRODUCTION

Patient safety is a top priority in healthcare systems and an important indicator in assessing the quality of services in various healthcare facilities. The World Health Organization (WHO) reports that unsafe care causes more than 3 million deaths globally each year and approximately 50% of these incidents are actually preventable. In low and middle income countries (LMICs), there are 134 million adverse events annually, resulting in approximately 2,6 million hospital deaths (WHO, 2024). These findings underscore that patient safety is not merely a clinical issue but also a public health concern with systemic and economic implications.

In Indonesia, the patient safety system is regulated by the Ministry of Health emphasizing the importance of identifying, managing risks, reporting and learning from incidents by establishing *Sistem Pelaporan dan Pembelajaran Keselamatan Pasien Nasional* (SP2KPN) as a data center for patient safety incidents (Permenkes No. 11 of 2017). The Ministry of Health's evaluation report shows that the quality and risk management indicator value at community health centers is the lowest at 35,86%, far below other indicators (Kementerian Kesehatan RI, 2022). This indicates that risk detection and prevention systems remain a weak point in efforts to improve patient safety in primary care.

In this context big data technology becomes highly relevant in healthcare referring to large and complex datasets from sources such as *electronic health records* (EHR), medical sensors, laboratories, social media, and genomic data (Badawy R, et al, 2024). According to Mazumder (2024) this technology enables the processing and analysis of large amounts of data in *real-time* to generate insights useful for clinical decision-making.

Big data has also been used to build machine learning and analytics-based prediction systems capable of identifying high-risk patients for adverse events. Research by Kessler (2020) found that using only administrative data available while patients were still hospitalized could predict the risk of post-discharge suicide in psychiatric hospitals. Other studies have shown that data from *Electronic Health Records* (EHR) can be used to predict Major Adverse Cardiac Events (Huang and Dong, 2018). Finally, research by Ramírez et al. (2019) shows that data from wearable devices, medical sensors, and Internet of Things (IoT) platforms can be utilized to support early detection and real-time analytics-based clinical decision-making.

This study comprehensively examines these phenomena by adopting a systematic literature review approach based on the PRISMA protocol, enabling systematic filtering and thematic classification of literature. The study also contributes new insights by dividing the findings into three main domains and highlighting technological adaptations in healthcare settings because to date there have been few systematic studies that comprehensively map the relationship between big data, risk modeling, and AI implementation in clinical practice and nursing. Through this approach, researchers synthesize existing scientific evidence, map technological approaches, and identify challenges and opportunities in using big data to improve patient safety across various healthcare service contexts.

2. METHODS

The method used in compiling this systematic literature review follows the PRISMA statement guidelines. Articles were searched for in the Scopus database using the keywords “*big data patient safety*” and “*predictive analytics adverse events*”.

All articles in this systematic literature review employ a quantitative research design with various research methods: 1 article uses an observational cohort study, 2 articles use observational studies, 2 articles use systematic reviews, 4 articles use experiments, 5 articles use surveys, 9 articles use meta-analyses, and 27 articles use case studies. The criteria for articles included in this systematic literature review were those related to patient safety and the use of big data published in the last 10 years (2014-2024). The criteria for articles excluded from this systematic literature review were those published before 2014.

3. RESULTS AND DISCUSSIONS

The PRISMA diagram for systematic literature review (SLR) uses *Watase Uake Tools* which outlines a clear process for identifying, screening, and including studies for review. During the identification phase, 130 articles were retrieved from the Scopus database, with no duplicate articles found. However, one article was excluded using an automated tool because its publication year fell outside the specified range (2014–2024). Additionally, four

articles were excluded for failing to meet the required quality standards (Tier Q1–Q4) and three articles lacked abstracts for screening purposes.

During the screening stage, 122 articles were screened based on their abstracts. Of these, 58 articles were excluded for being irrelevant, resulting in 64 articles being sought for full-text retrieval. However, 7 articles could not be retrieved. Subsequently, 57 articles were assessed for eligibility, and 7 additional articles were excluded at this stage for various reasons. Ultimately, 50 articles were deemed relevant and included in the final review. The article selection mechanism in Figure 1 shows the PRISMA flowchart depicting the study selection process is identification of records, removal of duplicates, screening of titles/abstracts, assessment of full-text articles for eligibility, and final inclusion in the systematic review.

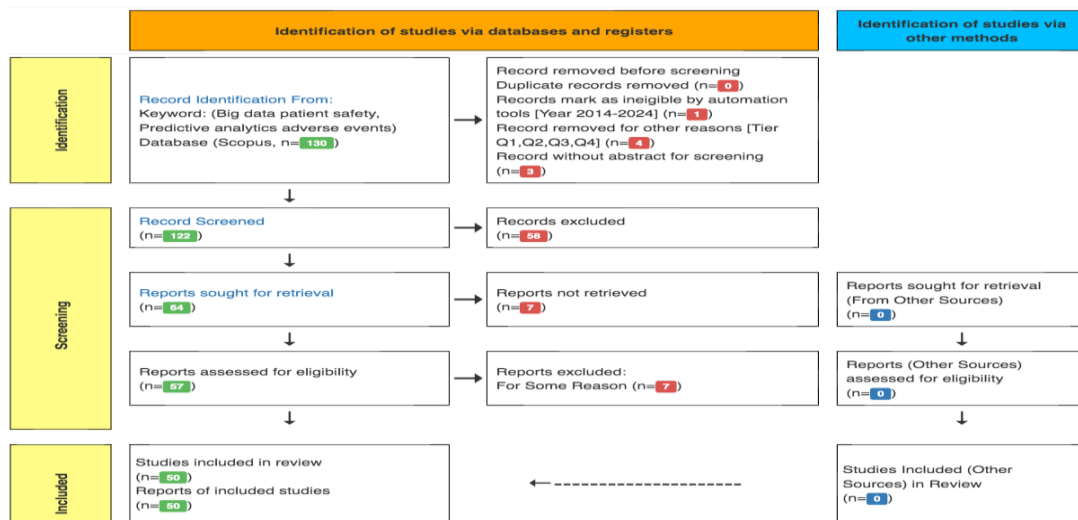


Figure 1. PRISMA Diagram

3.1 Article Year

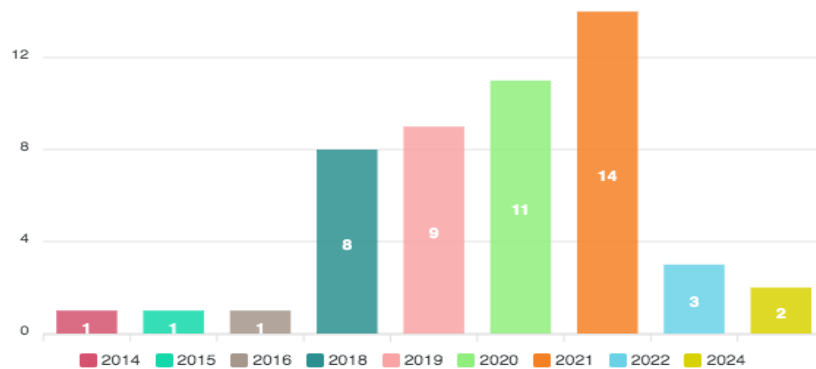


Figure 2. Year Article Classification

The distribution of articles from 2014 to 2024 shows a fluctuating trend, where publications were still low in the 2014–2016 period, reflecting the early stages of interest in this topic with only three articles (Caster et al., 2014; Sipes et al., 2015; Jelovsek, 2016). A significant increase began in 2018 with a surge in publications peaking in 2021 with 14 articles related to the growing attention to digital health solutions during the COVID-19 pandemic. One tangible implementation of this increase is the research by Soangra and Lockhart (2021) which developed a smartphone-based prediction model for cardiac surgery outcomes by measuring the patient's gait and posture before the operation. This affirms the capability of digital technology to support clinical assessment and accurately predict postoperative mortality and morbidity, reflecting the increased interest in health technology innovation during the pandemic. The number of publications decreased from 2022 to 2024 however, this trend still reflects an exploratory and transitional phase in big data-based patient safety research (Stahel et al., 2024)

3.2 Country

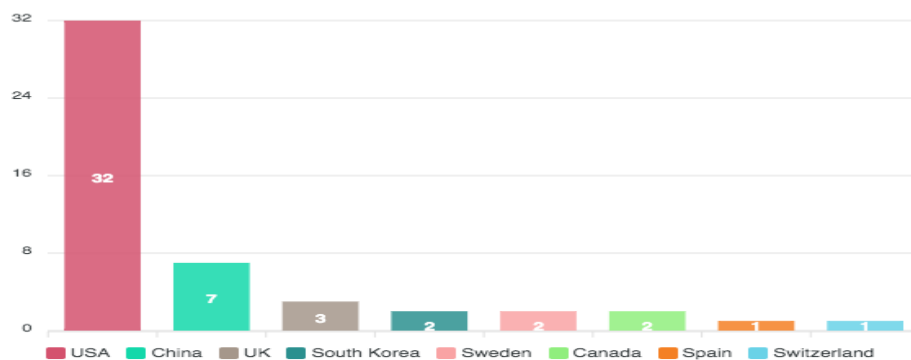


Figure 3. Country Classification

The chart above shows that the United States dominates the research with 32 articles, reflecting its leadership in health technology and data analytics such as the study by Ji et al. (2021) on improving the prediction of adverse drug effects (ADE) using an information component-pharmacology network model (IC-PNM). China follows in second place with 7 articles in line with its national policies supporting digital innovation especially in the health sector. One example is the development of a machine learning-based predictive model that has the potential to help doctors make more precise decisions regarding therapy for gynecological cancer patients (Liu et al., 2021). Other contributions from the United Kingdom, South Korea, and Sweden demonstrate the active participation of developed countries with strong healthcare and research systems (Cho et al., 2021; Rebane et al., 2020; Spinu, 2019). Furthermore, Canada, Spain, and Switzerland also play a role in applying machine learning technology to health big data for predictive analysis, thereby enhancing clinical decision support and leading to more personalized and efficient treatment (Finkelstein et al., 2021; Staartjes et al., 2020; Hassler et al., 2019).

3.3 Software

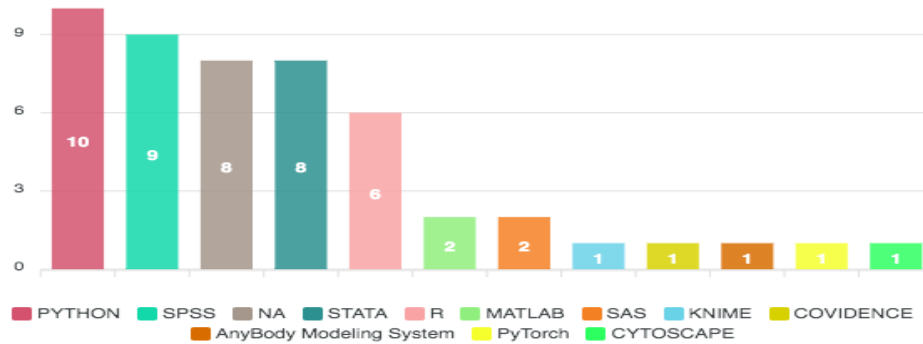


Figure 4. Software Classification

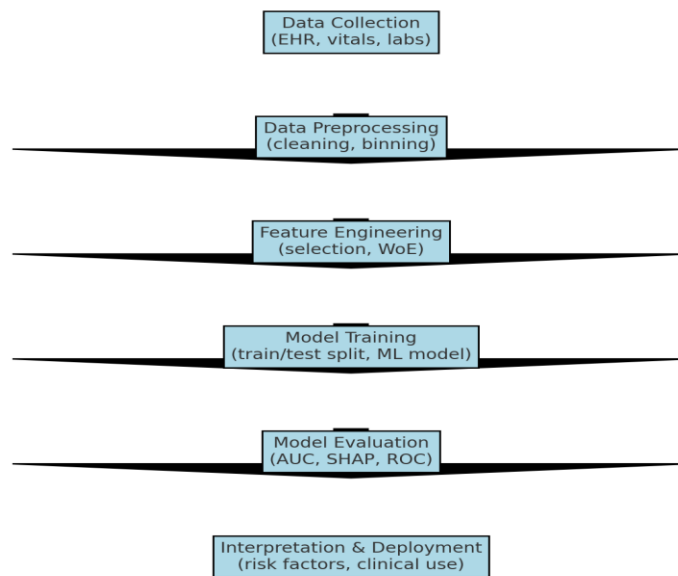


Figure 5. Flowchart Machine Learning

The software classification section shows that Python is the most dominant software in this research because its use allows for the efficient management of complex patient data, training of various predictive models, evaluation of their performance, and identification of relevant risk factors (Barker et al., 2024). Below are example Python code snippets used in the development of machine learning models for patient safety prediction.

Table 1. Python Code

Data Preparation & Binning	Model Training & Evaluation
<pre>import pandas as pd from optbinning import BinningProcess from sklearn.model_selection import train_test_split df = pd.read_csv("patient_data.csv") X = df[["age", "bmi", "asa_status", "surgery_time"]] y = df["uce"] binning_process = BinningProcess(variable_names=X.columns) X_binned = binning_process.fit_transform(X, y) X_train, X_test, y_train, y_test = train_test_split(X_binned, y, test_size=0.2, random_state=42)</pre>	<pre>from sklearn.linear_model import LogisticRegressionCV model = LogisticRegressionCV(cv=10, penalty='elasticnet', solver='saga', l1_ratios=[0.5]) model.fit(X_train, y_train) print(f"Model accuracy on test set: {model.score(X_test, y_test):.2f}")</pre>
	Model Interpretation with SHAP
	<pre>import shap explainer = shap.Explainer(model, X_train) shap_values = explainer(X_test) shap.summary_plot(shap_values, X_test)</pre>

In second place, SPSS is excellent for structured tabular data (rows and columns) commonly used in *electronic medical records* (EMRs) or patient databases (Hernandez et al., 2021). This is followed by STATA with its strengths in econometric analysis and statistical computing flexibility which can identify risk factors and support decision-making in the modern healthcare ecosystem (Strombotne et al., 2022). The next positions are held by R and MATLAB which excel in scientific and technical data visualization playing a crucial role in understanding big data patterns. Their ability to efficiently process, analyze, and model signals is a major advantage as sensor data is often in a time-series format suitable for matrix representation (Zimmet et al., 2021; Bonello et al., 2022). Several other software tools such as PyTorch, KNIME, CYTOSCAPE, SAS, and COVIDENCE are used in limited numbers for specialized functions like deep learning, network analysis, systematic reviews, and predictive data management (Petre et al., 2021; Rebane et al., 2020; Spinu et al., 2019; Davazdahemami and Delen, 2018; Stahel et al., 2024).

3.4 Big Data in Patient Safety

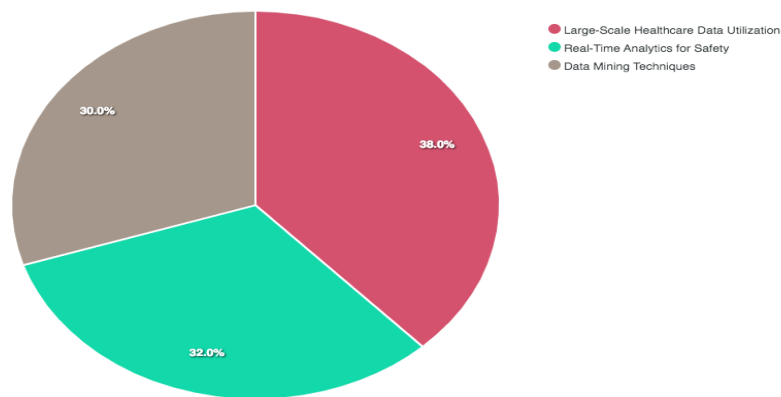


Figure 6. Category I Classification

Category I: *Big Data in Patient Safety* focuses on the utilization of large-scale health data to detect risks and improve responses to adverse medical events. This category is divided into three sub-categories.

The first sub-category *Large-Scale Healthcare Data Utilization* (38%), is a crucial foundation for developing predictive systems that enhance patient safety and healthcare efficiency. The massive volume of data from electronic health records and monitoring devices provides a unique opportunity to identify complex patterns and risk factors (Byun et al., 2021). For example in the context of major vascular surgery, research by Filiberto et al. (2021) showed that integrating intraoperative data into machine learning models significantly improves accuracy, discrimination, and precision in predicting postoperative Acute Kidney Injury (AKI). The potential of *Large-Scale Healthcare Data Utilization* extends to complex areas such as predicting *Drug-Drug Interactions* (DDI) a significant public health issue. This is demonstrated by the use of large-scale transcriptomic data from databases to develop deep learning models that predict DDI more effectively which is useful for treatment and the prevention of adverse side effects (Luo et al., 2021).

The second sub-category *Real-Time Analytics for Safety* (32%) involves using big data technology to monitor patient parameters in *real-time*, allowing for the early detection of critical conditions so that medical teams can respond more quickly to prevent patient deterioration. *Real-Time Analytics for Safety* leverages instantaneous data analysis to trigger early warnings and enable rapid medical interventions which is crucial for preventing adverse events and optimizing patient outcomes (Gillies et al., 2020). Research by Cumming et al. (2021) shows that the application of *real-time* predictive models can accurately predict critical events such as transfer to the intensive care unit (ICU), the need for mechanical ventilation, or death.

The third sub-category *Data Mining Techniques* (30%), consists of applying various methods to mine hidden patterns from large-scale historical data to identify previously undetected patient safety trends and risks. Research by Yao et al. (2019) demonstrated the application of an Artificial Neural Network (ANN)-based model to perform knowledge

mining on historical data of Traditional Chinese Medicine (TCM) prescriptions, which could predict whether a prescription had the potential to cause adverse effects. This illustrates how data mining techniques can extract knowledge from non-conventional data sources for patient safety.

3.5 Patient Risk Prediction

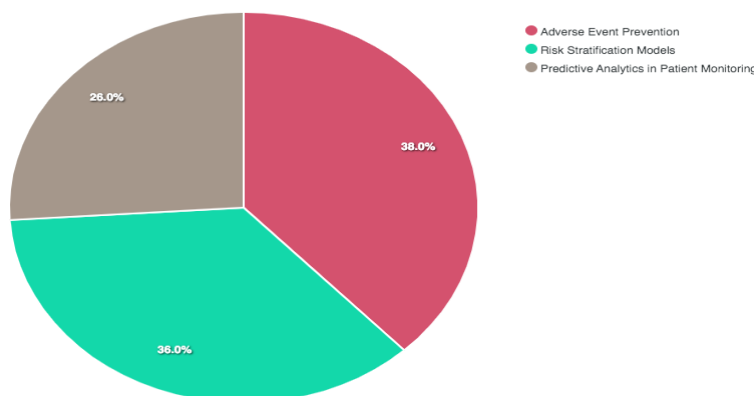


Figure 7. Category II Classification

Category II: *Patient Risk Prediction* is divided into three sub-categories which encompass the development of data-driven predictive models to anticipate adverse events, including the risk of falls, infections or postoperative complications.

The first sub-category *Adverse Event Prevention* (38%) focuses on developing prediction models to anticipate and prevent various adverse events ranging from poor patient clinical conditions to adverse drug effects (Wu et al., 2020). These models are designed to provide early warnings to healthcare professionals enabling proactive interventions before a risk materializes into an actual incident (Classen et al., 2018). One advanced approach in this area is the development of a machine learning framework to predict the risk of mortality and ICU admission/readmission by leveraging dynamic data (time series of vital signs and lab results from the first 24 hours) and static data (demographics and comorbidities), this framework can accurately identify high-risk patients even for rare events (Ibrahim et al., 2021).

The second sub-category *Risk Stratification Models* (36%) is used to group patients based on their level of vulnerability to complications. This approach ensures that care can be tailored earlier for high-risk groups and that clinical resources can be used more efficiently (Passias et al., 2020). *Risk stratification* is applied through the use of machine learning-based prediction models to evaluate the postoperative risk for patients undergoing cardiothoracic procedures. This is done by integrating key clinical variables and intraoperative parameters to classify patients into low, medium, and high-risk groups for mortality and complications (Lee et al., 2018).

The third sub-category *Predictive Analytics in Patient Monitoring* (26%) emphasizes the use of big data to continuously predict changes in a patient's clinical status. This approach transforms passive monitoring into an intelligent system that can detect risks and trigger early interventions (Harriton et al., 2021). A prime example of this application is the research by

Fruscione et al. (2019) who developed a specific predictive model to differentiate the risks for patients undergoing routine versus complex cholecystectomy with prediction accuracy reaching up to 98% by applying these predictive analytics, medical teams can adjust perioperative management strategies in *real-time* which is useful for optimizing patient clinical outcomes and significantly reducing postoperative complications.

3.6 Healthcare AI Implementation

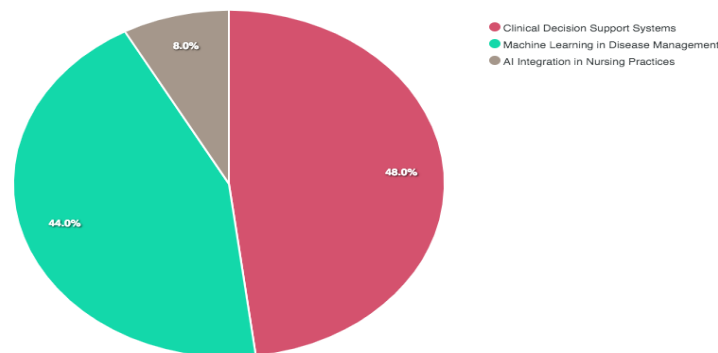


Figure 8. Category III Classification

Category III: *Healthcare AI Implementation* similar to Categories I and II consists of three sub-categories. It focuses on exploring the direct application of artificial intelligence in clinical practice, whether through decision support systems, complications or integration into nursing practices.

The first sub-category is *Clinical Decision Support Systems* (CDSS) (48%) based on big data and machine learning these systems are used to support clinical decision-making such as selecting therapies, determining drug dosages, or predicting procedural success. This improves practice consistency and reduces potentially harmful clinical variations (Kurian et al., 2020). The application of CDSS was demonstrated in a study by Sutherland (2018) who developed a predictive system based on *electronic health record* data to estimate patients' needs for *post-acute care* (PAC) services like rehabilitation. This system was proven to improve the accuracy and speed of care transition planning compared to traditional methods, thereby accelerating the discharge process and reducing the risk of delays in follow-up treatment.

The second sub-category *Machine Learning in Disease* (44%) is used to help predict the likelihood of complications following specific medical procedures. One example is in the care of aortic aneurysms where an artificial intelligence system analyzes pre-procedure patient data (such as age, blood pressure, cholesterol levels, and other lab results). The system is able to recognize patterns that may not be immediately visible to doctors and can predict the risk of complications such as vascular leakage, blockages, or even death with a high accuracy rate of over 86% (Kordzadeh et al., 2020).

Finally, the third sub-category which is important for the nursing field is *AI Integration in Nursing Practices* (8%). This involves the application of AI to support various aspects of nursing practice such as care planning, discharge planning, patient education and self-monitoring using digital devices. This expands the nurse's role in providing data-driven care

and supporting better clinical outcomes. One direct implementation example is the use of a predictive fall-risk analytics tool integrated into the *electronic health record* (EHR). Although its direct impact on fall incidence rates is complex, the tool's use has been shown to drive changes in nurse behavior such as increased risk awareness, conducting more multifactorial assessments, and implementing more targeted interventions tailored to the patient's individual risk profile (Cho et al., 2021). AI integration also occurs through accessible digital devices like smartphones. Sensor data from a smartphone can be used to assess a patient's pre-operative walking patterns and posture which is then fed into a predictive model to estimate the risk of post-operative complications. This provides nurses with a quick and objective risk assessment method that can help in more personalized and anticipatory pre-operative care planning (Soangra and Lockhart, 2021).

4. CONCLUSION

This systematic literature review confirms the transformative role of big data in enhancing patient safety through predictive analytics, risk stratification, and AI-driven clinical decision support. Key advancements include the use of large-scale data for early risk detection, development of predictive models for adverse event prevention and integration of AI in both medical and nursing practices. These innovations collectively demonstrate the potential of data-driven systems to personalize care, improve outcomes, and optimize resource use.

However, the study faces several limitations. The dominance of research from high-income countries particularly the United States raises concerns about external validity as predictive systems developed in well-resourced contexts may not be transferable to low and middle income countries (LMICs) with different healthcare infrastructures and data capacities. Methodological diversity among the included studies and the absence of qualitative insights further constrain generalizability. These limitations highlight the need for future research to adopt multi-database strategies, include more geographically diverse contexts, and integrate stakeholder perspectives particularly from LMICs to ensure equitable, ethical, and clinically relevant implementation of big data in patient safety.

REFERENCES

- Badawy, M., Ramadan, N., & Hefny, H. A. (2024). Big data analytics in healthcare: data sources, tools, challenges, and opportunities. *Journal of Electrical Systems and Information Technology*, 11(1), 63.
- Barker, A. B., Melvin, R. L., Godwin, R. C., Benz, D., & Wagener, B. M. (2024). Machine Learning Predicts Unplanned Care Escalations for Post-Anesthesia Care Unit Patients during the Perioperative Period: A Single-Center Retrospective Study. *Journal of Medical Systems*, 48(1), 69.
- Bonello, K., Emani, S., Sorensen, A., Shaw, L., Godsay, M., Delgado, M., ... & Kheir, J. N. (2022). Prediction of impending central-line-associated bloodstream infections in hospitalized cardiac patients: development and testing of a machine-learning model. *Journal of Hospital Infection*, 127, 44-50.

- Byun, J., Kim, S. Y., Kim, J. H., Kim, M. J., Yoo, C., Shim, J. H., & Lee, S. S. (2021). Prediction of transarterial chemoembolization refractoriness in patients with hepatocellular carcinoma using imaging features of gadoxetic acid-enhanced magnetic resonance imaging. *Acta Radiologica*, 62(12), 1548-1558.
- Caster, O., Juhlin, K., Watson, S., & Norén, G. N. (2014). Improved statistical signal detection in pharmacovigilance by combining multiple strength-of-evidence aspects in vigiRank: retrospective evaluation against emerging safety signals. *Drug safety*, 37, 617-628.
- Cho, I., sun Jin, I., Park, H., & Dykes, P. C. (2021). Clinical impact of an analytic tool for predicting the fall risk in inpatients: controlled interrupted time series. *JMIR Medical Informatics*, 9(11), e26456.
- Classen, D., Li, M., Miller, S., & Ladner, D. (2018). An electronic health record-based real-time analytics program for patient safety surveillance and improvement. *Health Affairs*, 37(11), 1805-1812.
- Cummings, B. C., Ansari, S., Motyka, J. R., Wang, G., Medlin Jr, R. P., Kronick, S. L., ... & Gillies, C. E. (2021). Predicting intensive care transfers and other unforeseen events: analytic model validation study and comparison to existing methods. *JMIR medical informatics*, 9(4), e25066.
- Davazdahemami, B., & Delen, D. (2018). A chronological pharmacovigilance network analytics approach for predicting adverse drug events. *Journal of the American Medical Informatics Association*, 25(10), 1311-1321.
- Filiberto, A. C., Ozrazgat-Baslanti, T., Loftus, T. J., Peng, Y. C., Datta, S., Efron, P., ... & Cooper, M. A. (2021). Optimizing predictive strategies for acute kidney injury after major vascular surgery. *Surgery*, 170(1), 298-303.
- Finkelstein, J. A., Stark, R. B., Lee, J., & Schwartz, C. E. (2021). Patient factors that matter in predicting spine surgery outcomes: a machine learning approach. *Journal of Neurosurgery: Spine*, 35(1), 127-136.
- Fruscione, M., Kirks, R. C., Cochran, A., Murphy, K., Baker, E. H., Martinie, J. B., ... & Vrochides, D. (2019). Routine versus difficult cholecystectomy: using predictive analytics to assess patient outcomes. *HPB*, 21(1), 77-86.
- Gillies, C. E., Taylor, D. F., Cummings, B. C., Ansari, S., Islim, F., Kronick, S. L., ... & Ward, K. R. (2020). Demonstrating the consequences of learning missingness patterns in early warning systems for preventative health care: a novel simulation and solution. *Journal of Biomedical Informatics*, 110, 103528.
- Hariton, E., Chi, E. A., Chi, G., Morris, J. R., Braatz, J., Rajpurkar, P., & Rosen, M. (2021). A machine learning algorithm can optimize the day of trigger to improve in vitro fertilization outcomes. *Fertility and Sterility*, 116(5), 1227-1235.
- Hassler, A. P., Menasalvas, E., García-García, F. J., Rodríguez-Mañas, L., & Holzinger, A. (2019). Importance of medical data preprocessing in predictive modeling and risk factor discovery for the frailty syndrome. *BMC medical informatics and decision making*, 19, 1-17.

- Hernandez, L., Kim, R., Tokcan, N., Derksen, H., Biesterveld, B. E., Croteau, A., ... & Gryak, J. (2021). Multimodal tensor-based method for integrative and continuous patient monitoring during postoperative cardiac care. *Artificial Intelligence in Medicine*, 113, 102032.
- Huang, Z., & Dong, W. (2018). Adversarial MACE prediction after acute coronary syndrome using electronic health records. *IEEE Journal of Biomedical and Health Informatics*, 23(5), 2117-2126.
- Ibrahim, Z. M., Bean, D., Searle, T., Qian, L., Wu, H., Shek, A., ... & Dobson, R. J. (2021). A knowledge distillation ensemble framework for predicting short-and long-term hospitalization outcomes from electronic health records data. *IEEE Journal of Biomedical and Health Informatics*, 26(1), 423-435.
- Jelovsek, J. E. (2016). Predicting urinary incontinence after surgery for pelvic organ prolapse. *Current Opinion in Obstetrics and Gynecology*, 28(5), 399-406.
- Ji, X., Wang, L., Hua, L., Wang, X., Zhang, P., Shendre, A., ... & Li, L. (2019). Improved adverse drug event prediction through information component guided pharmacological Network model (IC-pnm). *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, 18(3), 1113-1121.
- Karhade, A. V., Ogink, P. T., Thio, Q. C., Broekman, M. L., Cha, T. D., Hershman, S. H., ... & Schwab, J. H. (2019). Machine learning for prediction of sustained opioid prescription after anterior cervical discectomy and fusion. *The Spine Journal*, 19(6), 976-983.
- Kementerian Kesehatan Republik Indonesia. (2017). *Peraturan Menteri Kesehatan Republik Indonesia Nomor 11 Tahun 2017 tentang Keselamatan Pasien*. Kementerian Kesehatan RI.
- Kementerian Kesehatan Republik Indonesia. (2022). *Kebijakan mutu pelayanan kesehatan–Diseminasi INM Klinik & TPMD*. Direktorat Jenderal Pelayanan Kesehatan, Kementerian Kesehatan RI.
- Kessler, R. C., Bauer, M. S., Bishop, T. M., Demler, O. V., Dobscha, S. K., Gildea, S. M., ... & Bossarte, R. M. (2020). Using administrative data to predict suicide after psychiatric hospitalization in the veterans health administration system. *Frontiers in psychiatry*, 11, 390.
- Kordzadeh, A., Hanif, M. A., Ramirez, M. J., Railton, N., Prionidis, I., & Browne, T. (2021). Prediction, pattern recognition and modelling of complications post-endovascular infra renal aneurysm repair by artificial intelligence. *Vascular*, 29(2), 171-182.
- Kurian, S. J., Wahood, W., Yolcu, Y. U., Blaginykh, E., Goyal, A., Sebastian, A. S., & Bydon, M. (2020). Risk factors for unplanned readmissions following anterior cervical discectomy and fusion and posterior lumbar fusion procedures: comparison of two national databases. *World Neurosurgery*, 143, e613-e630.
- Lancaster, M. C., Salem Omar, A. M., Narula, S., Kulkarni, H., Narula, J., & Sengupta, P. P. (2019). Phenotypic clustering of left ventricular diastolic function parameters: patterns and prognostic relevance. *JACC: Cardiovascular Imaging*, 12(7 Part 1), 1149-1161.

- Lee, H. K., Jin, R., Feng, Y., Bain, P. A., Goffinet, J., Baker, C., & Li, J. (2018). An analytical framework for TJR readmission prediction and cost-effective intervention. *IEEE journal of biomedical and health informatics*, 23(4), 1760-1772.
- Liu, X., Xiao, Z., Song, Y., Zhang, R., Li, X., & Du, Z. (2021). A machine learning-aided framework to predict outcomes of anti-PD-1 therapy for patients with gynecological cancer on incomplete post-marketing surveillance dataset. *IEEE Access*, 9, 120464-120480.
- Luo, Q., Mo, S., Xue, Y., Zhang, X., Gu, Y., Wu, L., ... & Hu, Y. (2021). Novel deep learning-based transcriptome data analysis for drug-drug interaction prediction with an application in diabetes. *BMC bioinformatics*, 22, 1-15.
- Mazumder, M. S. A. (2024). The Transformative Impact of Big Data in Healthcare: Improving Outcomes, Safety, and Efficiencies. *Global Mainstream Journal of Business, Economics, Development & Project Management*, 3(03), 01-12.
- Milano, M. T., Grimm, J., Niemierko, A., Soltys, S. G., Moiseenko, V., Redmond, K. J., ... & Kleinberg, L. R. (2021). Single-and multifraction stereotactic radiosurgery dose/volume tolerances of the brain. *International Journal of Radiation Oncology* Biology* Physics*, 110(1), 68-86.
- Passias, P. G., Horn, S. R., Oh, C., Lafage, R., Lafage, V., Smith, J. S., ... & International Spine Study Group. (2020). Predicting the occurrence of postoperative distal junctional kyphosis in cervical deformity patients. *Neurosurgery*, 86(1), E38-E46.
- Petre, M. A., Saha, B., Kasuya, S., Englesakis, M., Gai, N., Peliowski, A., & Aoyama, K. (2021). Risk prediction models for emergence delirium in paediatric general anaesthesia: a systematic review. *BMJ open*, 11(1), e043968.
- Ramírez, M. R., Rojas, E. M., Núñez, S. O. V., & de los Angeles Quezada, M. (2019). Big data and predictive health analysis. In *Innovation in Medicine and Healthcare Systems, and Multimedia: Proceedings of KES-InMed-19 and KES-IIMSS-19 Conferences* (pp. 271-280). Springer Singapore.
- Rebane, J., Samsten, I., & Papapetrou, P. (2020). Exploiting complex medical data with interpretable deep learning for adverse drug event prediction. *Artificial Intelligence in Medicine*, 109, 101942.
- Sipes, T., Jiang, S., Moore, K., Li, N., Karimabadi, H., & Barr, J. R. (2014). Anomaly detection in healthcare: Detecting erroneous treatment plans in time series radiotherapy data. *International Journal of Semantic Computing*, 8(03), 257-278.
- Soangra, R., & Lockhart, T. (2021). Smartphone-based prediction model for postoperative cardiac surgery outcomes using preoperative gait and posture measures. *Sensors*, 21(5), 1704.
- Spinu, N., Bal-Price, A., Cronin, M. T., Enoch, S. J., Madden, J. C., & Worth, A. P. (2019). Development and analysis of an adverse outcome pathway network for human neurotoxicity. *Archives of toxicology*, 93(10), 2759-2772.

- Staartjes, V. E., Sebök, M., Blum, P. G., Serra, C., Germans, M. R., Krayenbühl, N., ... & Esposito, G. (2020). Development of machine learning-based preoperative predictive analytics for unruptured intracranial aneurysm surgery: a pilot study. *Acta neurochirurgica*, 162, 2759-2765.
- Stahel, P. F., Belk, K. W., McInnis, S. J., Holland, K., Nanz, R., Beals, J., ... & Mastriani, K. S. (2024). The Rothman Index predicts unplanned readmissions to intensive care associated with increased mortality and hospital length of stay: a propensity-matched cohort study. *Patient Safety in Surgery*, 18(1), 10.
- Strombotne, K. L., Legler, A., Minegishi, T., Trafton, J. A., Oliva, E. M., Lewis, E. T., ... & Frakt, A. B. (2023). Effect of a predictive analytics-targeted program in patients on opioids: a stepped-wedge cluster randomized controlled trial. *Journal of General Internal Medicine*, 38(2), 375-381.
- Sutherland, S. M. (2018). Electronic health record-enabled big-data approaches to nephrotoxin-associated acute kidney injury risk prediction. *Pharmacotherapy: The Journal of Human Pharmacology and Drug Therapy*, 38(8), 804-812.
- World Health Organization. (2024). *Global patient safety report 2024*. World Health Organization.
- Wu, Z., Peng, Y., Yu, Z., Li, W., Liu, G., & Tang, Y. (2020). NetInfer: a web server for prediction of targets and therapeutic and adverse effects via network-based inference methods. *Journal of Chemical Information and Modeling*, 60(8), 3687-3691.
- Yao, Y., Wang, Z., Li, L., Lu, K., Liu, R., Liu, Z., & Yan, J. (2019). An Ontology-Based Artificial Intelligence Model for Medicine Side-Effect Prediction: Taking Traditional Chinese Medicine as an Example. *Computational and mathematical methods in medicine*, 2019(1), 8617503.
- Zimmet, A. M., Sullivan, B. A., Fairchild, K. D., Moorman, J. R., Isler, J. R., Wallman-Stokes, A. W., ... & Lake, D. E. (2021). Vital sign metrics of VLBW infants in three NICUs: implications for predictive algorithms. *Pediatric research*, 90(1), 125-130.