Precision Medication 3 (2026) 100080

Contents lists available at ScienceDirect

Precision Medication

journal homepage: www.keaipublishing.com/en/journals/precision-medication/

Challenges in utilizing the FAERS database for adverse drug reaction data mining: A critical analysis

Xuelin Sun[a,*], Yatong Zhang[a], Dongfang Qian[b]

[a] Department of Pharmacy, Beijing Hospital, National Center of Gerontology, Institute of Geriatric Medicine, Chinese Academy of Medical Sciences, Beijing 100730, China
[b] Beijing North Medical Health Economic Research Center, Beijing 100021, China

ARTICLE INFO

Keywords:
FAERS database
Adverse drug reactions
Data mining
Artificial intelligence
Database optimization

ABSTRACT

The FAERS database is a vital tool for identifying adverse drug reactions (ADRs). However, data mining in FAERS faces significant challenges, including data quality issues (e.g., integrity, consistency, and completeness) and limitations in traditional model selection. These issues can introduce biases and affect the reliability of safety signal detection. This review critically analyzes the current state and limitations of FAERS data mining, particularly by briefly comparing it with other mainstream global databases to contextualize its unique challenges. It then proposes optimization strategies, focusing on improved data preprocessing, algorithm refinement, and the integration of emerging technologies. We emphasize the potential of Artificial Intelligence (AI) and multi-source data fusion to enhance detection sensitivity, accelerate the risk signal identification cycle, and address challenges in data-limited scenarios, such as rare diseases. We recommend promoting database standardization, strengthening validation, and formulating policy changes to fully realize FAERS's potential for precision pharmacovigilance.

Introduction

The FAERS (FDA Adverse Event Reporting System) database is a crucial tool for monitoring and identifying adverse drug reactions (ADRs). It plays an important role in drug safety, particularly in the context of the increasingly complex landscape of drug management and monitoring. The FAERS database provides the key data foundation for analyzing the relationship between drugs and reactions.[1] ADR monitoring is crucial in clinical practice, as it helps identify and reduce potential risks associated with drugs, and provides scientific evidence for the development of new drugs and the safe use of existing ones.[2]

Despite its widespread application in ADR monitoring and research, the FAERS database faces several challenges in data mining applications. Existing data mining methods suffer from issues with data completeness and accuracy, as well as limitations in model selection and algorithm suitability. These problems can lead to inaccurate data analysis results, which in turn impact the effectiveness of drug safety assessments and decision-making.[3,4]

Aim

The main objective of this paper is to analyze the current situation and limitations of FAERS data mining and to propose and explore methods to improve its performance, thus better supporting drug safety monitoring. This paper discusses various optimization methods for the FAERS database, including improving data preprocessing techniques, enhancing the precision of model algorithms, and the application of AI and multi-source data fusion technologies. The research will also explore the future development directions of the FAERS database, focusing on the promotion of database standardization and sharing, as well as how to strengthen the validation and clinical application of data mining results.

FAERS database and the context of ADR monitoring

Background and role of the FAERS database

The FAERS is an important tool for monitoring and identifying ADRs. The database contains ADRs submitted by healthcare professionals, consumers, and manufacturers worldwide, playing a crucial role in evaluating drug safety. FAERS supports post-market drug safety surveillance by collecting and analyzing ADR data to inform drug safety assessments and decisions. This process not only safeguards public health but also helps in improving the safety and effectiveness of medications.[1] The FAERS database's role in ADR monitoring can be illustrated by its data mining process, as shown in Fig. 1.

* Corresponding author.
E-mail address: sxl1220@163.com (X. Sun).

https://doi.org/10.1016/j.prmedi.2026.100080
Received 8 September 2025; Received in revised form 27 November 2025; Accepted 2 December 2025
Available online 25 March 2026
2950-5232/© 2026 Chinese Journal Practice Publishing House Co.,Ltd.. Publishing services by Elsevier B.V. on behalf of KeAi Communications Co. Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

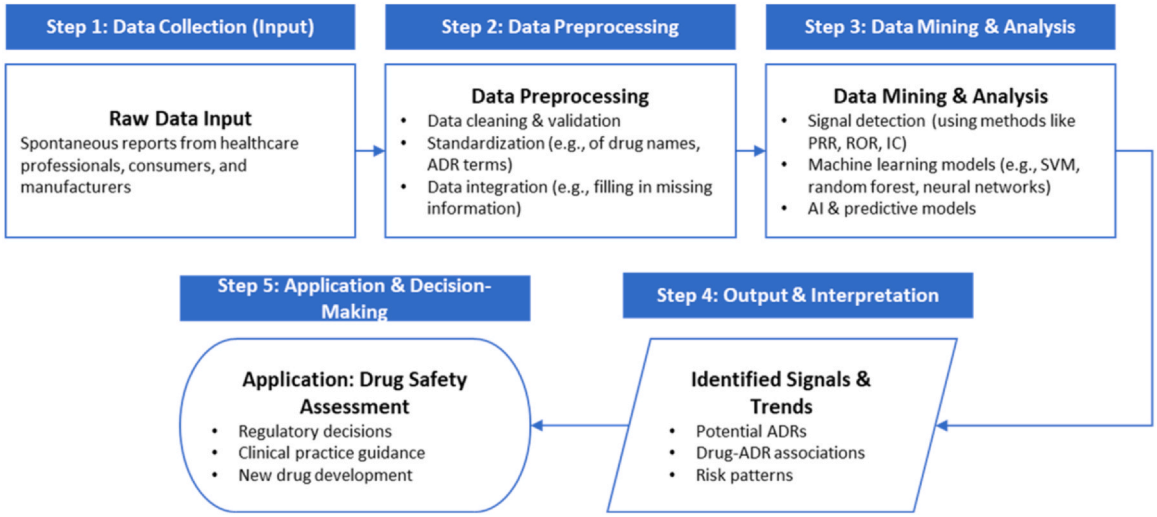

Fig. 1. The Ideal FAERS Data Mining Process. This flowchart outlines the systematic stages from data collection to safety signal detection and application.

Comparative context of global pharmacovigilance database

The landscape of global pharmacovigilance relies on multiple key databases. VigiBase (the WHO Global Individual Case Safety Report Database) and EudraVigilance (the European Medicines Agency's database) are two major counterparts to the FAERS.[5] VigiBase, containing data from over 130 countries, offers the broadest geographical coverage but often faces high variability in reporting standards.[6] EudraVigilance focuses on the European Economic Area, providing a robust, highly regulated system.[7]] FAERS, specifically managed by the FDA, serves as a cornerstone for drug safety in the U.S. market. While these databases share the common goal of ADR monitoring, the specific reporting culture, data fields, and regulatory requirements of each system lead to distinct challenges in data quality, consistency, and computational processing. Consequently, while many data mining methods are transferable, the optimization strategies must often be tailored to the specific characteristics of FAERS, which is the focus of this review.

Application of data mining in adverse drug reactions

Data mining techniques in the FAERS database significantly enhance the efficiency and accuracy of ADR monitoring. Data mining not only aids in signal detection but also helps with pattern recognition and trend analysis, uncovering potential patterns within complex data. With advances in machine learning, FAERS data mining has reached a new stage. Predictive analysis and automated models have significantly improved ADR detection and warning capabilities. Researchers can use machine learning models to predict which drug combinations may cause ADRs and identify risk factors for these reactions.[4] This forward-looking analytical approach not only enhances the sensitivity and specificity of data mining but also helps develop safer medication plans, reducing the likelihood of patients being exposed to potential risks.[8]

Existing FAERS data mining methods and their limitations

The FAERS database uses various data mining techniques to identify potential safety signals. Common methods include signal detection and machine learning models. Signal detection methods like Proportional Reporting Ratio (PRR), Reporting Odds Ratio (ROR), and Information Component (IC) are widely used in large datasets to identify potential ADRs.[9] These methods rely on statistical analysis to determine whether the occurrence frequency of a particular ADR is significantly higher than expected for a given drug.

Machine learning methods play an increasingly important role in complex signal detection and prediction tasks. Algorithms such as Support Vector Machines (SVM), Decision Trees, Random Forests, and Neural Networks are used to create predictive models for detecting new ADRs.[10] These models extract features from large datasets and provide higher classification accuracy. In practical applications, researchers often combine signal detection and machine learning methods to achieve more reliable and precise results. A detailed comparison of the performance characteristics and limitations of these traditional and AI-based methods is summarized in Table 1.

FAERS data mining challenges and inherent limitations

Core challenges: data quality, integrity, and model selection

Data quality and integrity represent the most persistent challenges facing FAERS data mining. The primary cause lies in the spontaneous and voluntary nature of the reports, which are submitted by healthcare professionals, manufacturers, and ordinary consumers. The diversity of report sources and the lack of consistent professional training for all reporters frequently lead to data integrity and accuracy issues.[11,13]

Table 1
Comparative Performance of Traditional vs. AI-Based Methods in Adverse Drug Reaction Signal Detection.

Performance Metric	PRR/ROR (Traditional Methods)	Machine Learning / Deep Learning (AI-Based)
Detection of Rare ADRs	Low (Highly dependent on large sample size)	Medium to High (Leverages feature transfer; less constrained by zero counts)
Sensitivity	Medium	High (Due to capacity to capture complex patterns)
Specificity	Medium	High (Improved by noise reduction and multi-feature analysis)
Handling Data Heterogeneity	Low (Sensitive to inconsistent reporting/coding)	High (Utilizes NLP for free-text processing)
Feature Interaction Analysis	Limited to low-order interactions	Extensive (Detects complex, non-linear relationships in high-dimensional data)
Interpretability	High (Easily understood by regulators/ clinicians)	Medium to Low (Depends on model type, e.g., Deep Learning models are "black-box")
False Positive Rate	Medium to High	Low (With sufficient training and validation)

Reports often lack crucial information, such as precise patient dosage, duration of drug use, and detailed accompanying symptoms, which are vital for establishing a clear drug-ADR association. Furthermore, inconsistent coding or manual input errors in drug names and ADR descriptions (despite the use of standardized terminologies like MedDRA) can introduce significant noise and affect data completeness and accuracy.[13] The absence of standardization and systematic feedback mechanisms further exacerbates instability and inconsistency in the data input.[14] These data deficiencies directly impact the reliability of signal detection. Noisy or incomplete data can lead to biases in statistical models, resulting in false positives (identifying a signal where none exists) or missed signals (failing to detect a genuine, often rare, ADR).[11]

Beyond data quality, the selection and application of analytical tools present a second layer of challenge. Different algorithms have varying applicability, and improper selection may lead to poor detection results or misguidance.[12] For instance, while traditional signal detection methods (PRR, ROR) are efficient, they may struggle to capture subtle patterns in complex, high-dimensional data. Conversely, more complex machine learning models may perform poorly when the underlying data is insufficient or noisy.[1] Thus, researchers must navigate the trade-off between model complexity, interpretability, computational cost, and the specific goals of clinical application, making algorithm selection a persistent hurdle in achieving optimal detection results. The summary of the main content is shown in Fig. 2.

Impact of data consistency on results

Data consistency directly affects the accuracy and reliability of data mining analysis. In ADR signal detection, inconsistent data may lead to false positives or missed signals. If there is a lack of consistency between data sources, even advanced data mining algorithms will struggle to extract accurate signals. Furthermore, the inability to trace changes in report data limits the verifiability and traceability of the data, which has a significant impact on clinical research and regulatory decision-making.[15,16] To address these challenges, stricter data auditing and standard processes need to be established to ensure the accuracy and consistency of information during data entry and transmission. Machine learning and natural language processing (NLP) technologies can automatically detect and correct certain inconsistencies at the data input stage. At a deeper level, enhanced data sharing mechanisms and feedback loops can help reporters improve the accuracy and completeness of their reports, thereby increasing data consistency.[16]

Adopting technical measures and formulating improvement strategies are indispensable for continuously enhancing the comprehensiveness and accuracy of the data. Through these improvements, the FAERS database can better realize its potential, supporting the analysis and early warning of adverse drug reactions and improving the effectiveness of drug safety monitoring.

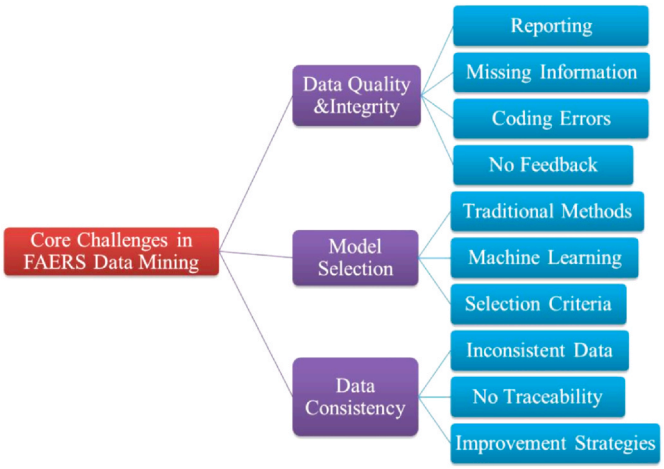

Fig. 2. Core Challenges in FAERS Data Mining.

Suggestions for optimizing the FAERS database

Improvements in data preprocessing techniques

Data preprocessing is the first step in utilizing the FAERS database. Effective data preprocessing can enhance data cleanliness and consistency, laying the foundation for subsequent data mining. Automated data extraction and cleaning processes can improve data accuracy and completeness. NLP technologies can be utilized to identify key entities and relationships in text, reducing data noise and redundancy caused by human errors or inconsistent reports. In addition, applying rule engines and machine learning algorithms can facilitate complex data correction and anomaly detection, ensuring efficient data processing and exclusion of abnormal data.[17] Furthermore, incorporating multi-source data integration during the data preprocessing stage is also crucial. By integrating data from different sources using consistent standards and protocols, the coverage and richness of data in the FAERS database can be improved. Additionally, semantic web technologies can be used to identify and link related data, thus increasing data relevance and usability.[18]

Improving algorithm model accuracy

To improve the effectiveness of ADR signal detection in the FAERS database, it is necessary to enhance the accuracy of the algorithms used. Traditional data mining methods can provide useful results to some extent, but their limitations in handling multi-dimensional data often lead to false positives or missed signals. In recent years, the introduction of machine learning and deep learning technologies has driven improvements in sensitivity and specificity in signal detection.[19] Recurrent Neural Networks (RNN) and Convolutional Neural Networks (CNN) are capable of deeper analysis and mining of complex time-series and spatial data. These methods show great potential in clinical trials and post-market drug safety monitoring.[17] In terms of algorithm model selection, multiple algorithms can be combined to form an ensemble learning model, thereby improving prediction accuracy. Additionally, using transfer learning and adaptive learning pruning strategies can help maintain the model's interpretability and computational efficiency to some extent. This approach not only saves training time but also performs excellently in adapting to new data. For achieving high-precision predictions, it is crucial to design a reasonable experimental plan and conduct benchmark validation using a large volume of historical data.[20]

By continuously updating and optimizing algorithms and data mining strategies based on FAERS, drug safety monitoring capabilities can be further enhanced, effectively reducing the occurrence of ADRs and associated healthcare costs. Each advancement made will provide significant clinical value and scientific importance for the pharmaceutical industry.

Emerging technologies and methods to address FAERS database challenges

Application of AI in data mining

The introduction of AI technology has provided new solutions to improve the efficiency and accuracy of data mining. AI can automatically process large amounts of data and build predictive models by learning from historical data to identify potential ADRs. This method not only increases the sensitivity and specificity of detection but also reduces human errors.[21,22] Studies using deep learning models to analyze data from the FAERS database have shown that these models can accurately predict a variety of ADRs, including some rare reactions.

AI models, particularly those based on deep learning, enable improved prediction accuracy by leveraging key mechanisms such as pattern recognition in high-dimensional data and NLP of free-text reports. High-dimensional pattern recognition allows complex algorithms to detect subtle, non-linear relationships between multiple variables (e.g., patient demographics, concomitant medications, and adverse event terms) that traditional statistical methods often miss. NLP is crucial for converting the unstructured, varied, and often nuanced narratives in the FAERS free-text fields into quantifiable, standardized data. This capability significantly enhances data preprocessing and feature extraction, leading to the identification of signals that would otherwise be obscured by the heterogeneity and variability of the data.[21]

However, the application of AI also comes with challenges, such as the high quality requirements for model training data and the need for algorithm transparency and interpretability. Current research focuses on improving the generalization ability of the model and reducing overfitting. Moreover, how to achieve efficient data processing while reducing computational resource consumption remains an unresolved issue. Addressing these challenges requires researchers to continuously optimize AI algorithms and collaborate with medical experts to better understand clinical needs.

Multi-source data fusion methods and practices

In addition to AI technology, multi-source data fusion is another important method to address the challenges of the FAERS database. Multi-source data fusion aims to integrate data from different sources to provide a more comprehensive drug safety evaluation. These data sources include not only FAERS but also VigiBase, electronic health records (EHRs), literature databases, and social media data. Multi-source data fusion technology can effectively eliminate the biases associated with a single data source, improving the accuracy of ADR detection.[22,23] However, challenges in data sharing and privacy protection, standardization of different data formats, and potential technical barriers in information integration remain. To address these challenges, efforts can be made to strengthen the development of regulatory frameworks, promote standardization of data formats, and use efficient cross-platform data transfer protocols.[24]

Optimizing signal detection in data-limited scenarios

The challenge of limited data, particularly in populations like pediatrics or for rare diseases, is a significant constraint in FAERS data mining. In these contexts, traditional statistical methods often suffer from low power and high false discovery rates due to small sample sizes. Addressing this requires specialized approaches that leverage the capabilities of emerging technologies.

Transfer Learning and Federated Learning are highly promising. Transfer Learning allows models pre-trained on large, general FAERS datasets (e.g., common drugs) to be fine-tuned on smaller, specialized datasets (e.g., a specific rare disease or pediatric drug), transferring learned features and reducing the need for massive amounts of training data in the target domain.[25] Federated Learning enables multiple institutions to collaboratively train a shared model without centrally pooling their sensitive patient data, which is crucial for rare conditions where data is fragmented across different hospitals.[26] Furthermore, advanced multi-source data fusion can help mitigate data sparsity by integrating evidence from EHRs and genomic data, providing a richer context for each individual case report.

Development trends of the FAERS database in ADR mining

Promoting database standardization and sharing

The effectiveness and practical utility of the FAERS database in drug safety monitoring are frequently hampered by insufficient data standardization and sharing. The current lack of standardization hinders direct data comparison across different studies, thereby impeding comprehensive analysis and the formulation of robust scientific conclusions. Improving standardization is crucial, as it can significantly reduce bias in data interpretation, leading to more accurate ADR signal detection, promoting research consistency, and enhancing the transparency of monitoring processes.

Similarly, limitations in data sharing restrict the widespread impact of pharmacovigilance research, especially in global safety monitoring projects. Effective data sharing not only boosts research efficiency but also increases the reliability and generalizability of results by allowing the integration of evidence from multiple sources.

Achieving these goals necessitates international collaboration. Establishing a globally unified data standard system is paramount. This system must balance the need for data privacy and security with considerations for diverse regulatory frameworks and cultural differences. Operating within such a coordinated global framework is essential for all stakeholders to effectively share data resources and advance the field of drug safety regulatory science worldwide.

Strengthening the validation and application of data mining results

A second critical future direction, alongside standardization and sharing, is to improve the validation and practical application of data mining results. Although FAERS data contains valuable ADR signals, it also includes significant noise. Therefore, the results of data mining must be rigorously validated across diverse research contexts to ensure their reliability and applicability.

To enhance the credibility of findings, initial data mining models must be continually evaluated and optimized. This includes the use of techniques such as cross-validation, validation against external datasets, and comparison with clinical trial results. For predicting adverse drug reactions, machine learning-based algorithms offer dynamic solutions. Their continuous iteration and optimization are vital for enhancing adaptability and accuracy across various clinical scenarios.[27]

The application of data mining results must also transition from the research domain directly into clinical practice. The ADR signals identified should inform and guide the adjustment of drug usage strategies, especially for high-risk medications or vulnerable patient groups. The feedback loop involving pharmacists and clinicians is indispensable in this process, as their insights are key to refining data models and ensuring the practical relevance of results. Through this continuous cycle of validation and application, the FAERS database will realize its full potential, providing both theoretical foundation and tangible value in clinical operations.

As the full potential of the FAERS database is realized, it becomes necessary to explore new technologies and methods to address persistent challenges. This includes the integration of AI and big data technologies with the analysis of consolidated multi-source data. By connecting FAERS data with EHR systems, it will be possible to identify potential ADRs more comprehensively, thereby improving the specificity and sensitivity of signal detection.[28] This integrated, data-rich approach not only offers deeper insights into drug safety monitoring but also actively promotes the development of precision medicine, laying a robust foundation for the formulation of individualized treatment plans.

Conclusion

Achieving the full potential of the FAERS database requires innovation and optimization across multiple aspects of pharmacovigilance. First, promoting database standardization and data sharing, establishing unified data standards and protocols, can improve the availability and interoperability of data, promoting international drug safety monitoring collaboration. Secondly, introducing AI and

multi-source data fusion methods at the technical level will significantly enhance the FAERS database's analytical capabilities. AI technology can improve signal detection sensitivity and specificity, helping to identify rare ADRs and their underlying mechanisms. Multi-source data fusion can provide more comprehensive background information for data mining, improving the accuracy and reliability of results. Additionally, strengthening the validation and application of data mining results is equally important. This requires validating the analysis results through appropriate clinical trials and real-world data to ensure their effectiveness and safety in clinical decision-making.

From a policy perspective, it is essential to formulate policies that encourage data standardization, mandate data sharing frameworks where ethically permissible, and fund the development and rigorous validation of advanced analytical tools. Achieving these goals requires continuous research investment and interdisciplinary collaboration. Future research should focus on developing new algorithms, enhancing the analytical capabilities of existing tools, and fostering broad collaboration across various sectors to advance the field of drug safety monitoring, thereby transitioning pharmacovigilance from a reactive to a proactive system capable of delivering personalized drug safety assessments in the era of precision medicine.

Declarations

Not applicable.

CRediT authorship contribution statement

Xuelin Sun: Conceptualization, Methodology, Writing - Original draft preparation. **Yatong Zhang:** Literature retrieval and manuscript revision. **Dongfang Qian:** Visualization, Investigation.

Ethics approval and consent to participate

Not applicable.

Consent for publication

All authors have read and agreed to the published version of the manuscript and give their consent for publication in this journal.

Availability of data and materials

Not applicable.

Funding

This work was supported by the National High Level Hospital Clinical Research Funding (No. BJ-2023-200).

Declaration of Competing interest

The authors declare no competing interests.

Acknowledgements

Not applicable.

Authors' other information

Not applicable.

References

1. Duan R, Zhang X, Du J, et al. Post-marketing drug safety evaluation using data mining based on FAERS. *Data Min Big Data.* 2017:379–389 2017;2017.
2. Arku D, Yousef C, Abraham I. Changing paradigms in detecting rare adverse drug reactions: from disproportionality analysis, old and new, to machine learning. *Expert Opin Drug Saf.* 2022;21(10):1235–1238.
3. Nango D, Sekizuka T, Goto M, et al. Analysis of information on drug adverse reactions using U.S. Food and Drug Administration Adverse Event Reporting System (FAERS). *Yakugaku Zasshi.* 2022;142(4):341–344.
4. Sakaeda T, Tamon A, Kadoyama K, et al. Data mining of the public version of the FDA Adverse Event Reporting System. *Int J Med Sci.* 2013;10(7):796–803.
5. Ventola CL. Big Data and Pharmacovigilance: Data Mining for Adverse Drug Events and Interactions. *P T.* 2018;43(6):340–351.
6. HJELMSTRÖM Peter, BOWRING Geoffrey, YUE Qun-Ying, et al. Methods for signal management using the global safety database VigiBase. *Chin J Pharmacovigil.* 2024;21(7):836–840.
7. Postigo R, Brosch S, Slattery J, et al. EudraVigilance Medicines Safety Database: Publicly Accessible Data for Research and Public Health Protection. *Drug Saf.* 2018;41(7):665–675.
8. Gibbons RD, Amatya AK, Brown CH, et al. Post-approval drug safety surveillance. *Annu Rev Public Health.* 2010;31:419–437.
9. Martínez-Abad F. Identification of factors associated with school effectiveness with data mining techniques: testing a new approach. *Front Psychol.* 2019;10:2583.
10. Jiménez-Carvelo AM, González-Casado A, Bagur-González MG, et al. Alternative data mining/machine learning methods for the analytical evaluation of food quality and authenticity – a review. *Food Res Int.* 2019;122:25–39.
11. Li C, Gao J, Pan Q, et al. Design and development of a big data platform for disease burden based on the Spark engine. *Comput Intell Neurosci.* 2023;2023:8963053.
12. Liao PH, Chu W, Chu WC. Evaluation of the mining techniques in constructing a traditional Chinese-language nursing recording system. *Comput Inf Nurs.* 2014;32(5):223–231.
13. Veronin MA, Schumaker RP, Dixit R. The irony of MedWatch and the FAERS database: an assessment of data input errors and potential consequences. *J Pharm Technol.* 2020;36(4):164–167.
14. Cho S, Ensari I, Weng C, et al. Factors affecting the quality of person-generated wearable device data and associated challenges: rapid systematic review. *JMIR Mhealth Uhealth.* 2021;9(3):e20738.
15. Harkener S, Stausberg J, Hagel C, et al. Towards a core set of indicators for data quality of registries. *Stud Health Technol Inf.* 2019;267:39–45.
16. Kim MK, Rouphael C, McMichael J, et al. Challenges in and opportunities for electronic health record-based data analysis and interpretation. *Gut Liver.* 2024;18(2):201–208.
17. Raju SH, Rao MN. Application of a data mining task called data preprocessing on the input data and efficient external sorting using refinement of existing algorithm. *Int J Pharm Technol.* 2016;8(3):18395–18407.
18. Hung E, Hauben M, Essex H, et al. More extreme duplication in FDA Adverse Event Reporting System detected by literature reference normalization and fuzzy string matching. *Pharmacoepidemiol Drug Saf.* 2023;32(3):387–391.
19. Carrizosa E, Molero-Río C, Romero Morales D. Mathematical optimization in classification and regression trees. *Top (Berl).* 2021;29(1):5–33.
20. Loeffler C, Karlsberg A, Martin LS, et al. Correction to: Improving the usability and comprehensiveness of microbial databases. *BMC Biol.* 2020;18(1):92.
21. Muthuraj, Singla S. Artificial intelligence and machine learning. *MedLeg Update.* 2023;23(5):6–11.
22. Damar M, Yüksel İ, Çetinkol AE, et al. Advancements and integration: a comprehensive review of health informatics and its diverse subdomains with a focus on technological trends. *Health Technol.* 2024;14(4):635–648.
23. Dong P, Mao A, Qiu W, et al. Improvement of cancer prevention and control: reflection on the role of emerging information technologies. *J Med Internet Res.* 2024;26:e50000.
24. Zdravevski E, Pires IM. Advancing methods in big data capture, integration, classification and liberation. *BMC Res Notes.* 2023;16(1):64.
25. Sacco SJ, Chen K, Wang F, et al. Using transfer learning to improve prediction of suicide risk in acute care hospitals. *J Am Med Inf Assoc.* 2026;33(1):159–166.
26. Kazlouski A, Montoya Perez I, Noor F, et al. Towards practical federated learning and evaluation for medical prediction models. *Int J Med Inf.* 2025;204:106046.
27. Nguyen DA, Nguyen CH, Mamitsuka H. A survey on adverse drug reaction studies: data, tasks and machine learning methods. *Brief Bioinform.* 2021;22(1):164–177.
28. Choi YH, Han CY, Kim KS, et al. Future directions of pharmacovigilance studies using electronic medical recording and human genetic databases. *Toxicol Res.* 2019;35(4):319–330.