ISSN: 3068-7527

Open Access Journal of Medicine and Healthcare

Review Article

OPENPUB GLOBAL PUBLISHERS

Volume 2 | Issue 2

Artificial Intelligence-Based Clinician Support Tools for Real-World Evidence Generation in Rare Diseases: A Comprehensive Evaluation

Dr. Sharika Pillay MD*

Senior Resident, Department of Pharmacology HBTMC and RNCH Mumbai

*Corresponding author: Dr. Sharika Pillay, Senior Resident, Department of Pharmacology HBTMC and RNCH Mumbai

Citation: Dr. Sharika P (2026) Artificial Intelligence-Based Clinician Support Tools for Real-World Evidence Generation in Rare Diseases: A Comprehensive Evaluation. Open Access J. Med. Healthc. 2(1): 1-9.

Received: 07-03-2026 Accepted: 10-03-2026 Published: 22-03-2026

Abstract

Background: Rare diseases collectively affect approximately 300 million individuals worldwide, yet each individual condition typically involves fewer than 200,000 patients, creating substantial challenges for evidence generation through conventional randomised controlled trials. Real-world evidence (RWE) offers an alternative pathway to understanding disease natural history, treatment effectiveness, and patient outcomes. Artificial intelligence (AI)-based clinician support tools have emerged as transformative instruments for extracting, structuring, and analysing RWE from heterogeneous data sources in rare disease contexts.

Objectives: This review comprehensively evaluates the current landscape of AI-based clinician decision support tools deployed for RWE generation in rare diseases, including their technical architectures, data modalities, regulatory considerations, and clinical impact.

Methods: A systematic literature review was conducted across PubMed, EMBASE, and IEEE Xplore databases (2015–2024). Studies evaluating AI or machine learning tools applied to rare disease RWE generation were included. Data extraction covered AI methodology, data sources, clinical outcomes, and validation approaches.

Results: A total of 247 studies met inclusion criteria. Natural language processing (NLP) was the most prevalent methodology (28%), followed by machine learning classifiers (24%) and deep learning (18%). AI tools demonstrated a 34–67% improvement in diagnostic coding accuracy and a 2.4-fold increase in cohort identification yield compared to manual methods. Federated learning architectures enabled cross-institutional rare disease data analysis while preserving patient privacy.

Conclusions: AI-based clinician support tools represent a paradigm shift in rare disease RWE generation, offering scalable, reproducible, and privacy-preserving approaches to evidence synthesis. Standardisation of validation frameworks and regulatory guidance remain critical priorities for widespread clinical adoption.

Keywords: Artificial Intelligence; Rare Diseases; Real-World Evidence; Natural Language Processing; Machine Learning; Clinician Decision Support; Orphan Diseases; Health Data Science.

Introduction

Rare diseases, defined by the European Union as conditions affecting fewer than 1 in 2,000 individuals and by the US Orphan Drug Act as those affecting fewer than 200,000 Americans, collectively impose an enormous and underestimated global health burden [1]. Despite individually low prevalence, more than 7,000 distinct rare diseases have been characterised, together affecting an estimated 300 million people worldwide [2]. Approximately 72% of rare diseases have a genetic aetiology, 70% manifest in childhood, and fewer than 5% have approved pharmacological treatments [3].

Generating robust clinical evidence for rare diseases presents unique challenges that render conventional randomised controlled trial (RCT) designs impractical or ethically problematic in many scenarios. Small and geographically dispersed patient populations, heterogeneous phenotypic expression, extended diagnostic odysseys averaging 4–7 years, and the absence of validated surrogate endpoints collectively undermine the feasibility of adequately powered prospective studies [4,5]. In this context, real-world evidence (RWE) derived from electronic health records (EHRs), patient registries, claims databases, genomic repositories, and wearable sensor data has been increasingly recognised by regulators and pharmaceutical developers as a critical complement to, or sometimes substitute for, traditional trial-generated evidence [6,7].

Artificial intelligence (AI) and machine learning (ML) technologies have dramatically expanded the capacity to extract structured, clinically meaningful information from these heterogeneous, voluminous, and often incomplete real-world datasets. AI-based clinician support tools encompassing natural language processing (NLP) engines, automated phenotyping algorithms, predictive analytics platforms, and federated learning frameworks have demonstrated particular promise in the rare disease domain, where data sparsity amplifies the value of every patient data point [8,9]. These tools facilitate tasks ranging from automated rare disease diagnosis coding and patient cohort identification, to predictive modelling of disease progression and treatment response [10].

This comprehensive review evaluates the evolving ecosystem of AI-based clinician support tools applied to rare disease RWE generation, examining their technical architectures, clinical applications, validation frameworks, regulatory landscape, and future prospects. Understanding this rapidly developing field is essential for clinicians, health data scientists, pharmaceutical developers, and policymakers seeking to harness AI for evidence-based rare disease medicine [11].

Background and Conceptual Framework

The Real-World Evidence Paradigm in Rare Disease

Real-world evidence is defined by the US Food and Drug Administration (FDA) as clinical evidence regarding the usage and potential benefits or risks of a medical product derived from analysis of real-world data (RWD) [12]. The 21st Century Cures Act (2016) mandated the FDA to develop a programme for evaluating RWE to support regulatory decisions, catalysing substantial investment in RWE methodologies across therapeutic areas [13].

For rare diseases specifically, RWE holds particular regulatory and scientific importance. The FDA's Rare Disease Framework and EMA's PRIME designation scheme have both incorporated provisions for RWE-based evidence packages, recognising that natural history studies and patient registry data may constitute the primary evidence base for regulatory submissions in orphan conditions [14,15]. Disease-specific registries such as the European Registry of Rare Kidney Diseases (ERKReg), the EUROBAMSE database for metabolic conditions, and the NMD-SC registry for neuromuscular disorders have served as foundational RWD sources [16].

AI Architectures Relevant to Rare Disease RWE

The AI methodologies most commonly deployed in rare disease RWE generation span a spectrum of technical complexity and clinical utility. Natural language processing algorithms including rule-based systems, statistical models, and transformer-based architectures such as BioBERT and Clinical BERT enable automated extraction of clinical concepts from unstructured physician notes, discharge summaries, and pathology reports [17,18]. These approaches are particularly valuable in rare diseases, where formal diagnostic codes (ICD codes) are often absent, delayed, or assigned to broader categorical terms due to clinician unfamiliarity with rare condition coding [19].

Supervised machine learning classifiers including random forests, gradient boosting machines, and support vector machines have been applied to structured EHR and claims data to identify rare disease patients with high specificity, estimate disease prevalence, and model treatment patterns [20]. Deep learning architectures, including convolutional neural networks (CNNs) for medical imaging analysis and recurrent neural networks (RNNs) for longitudinal EHR sequence modelling, extend these capabilities to phenotyping and disease staging applications [21,22].

Federated learning represents an architecturally distinct approach of particular importance in rare disease contexts, enabling AI model training across multiple institutional datasets without centralising patient-identifiable information [23]. This framework is especially valuable where privacy legislation (GDPR, HIPAA) or institutional data governance policies preclude direct data sharing, yet multi-site collaboration is essential to achieve statistically meaningful rare disease cohort sizes [24].

Methodology

Literature Search Strategy

A systematic literature search was conducted in accordance with PRISMA 2020 guidelines across PubMed/MEDLINE, EMBASE, IEEE Xplore, and the Cochrane Library from January 2015 to December 2024 [25]. Search terms included: ("artificial intelligence" OR "machine learning" OR "deep learning" OR "natural language processing") AND ("rare disease" OR "orphan disease") AND ("real-world evidence" OR "real-world data" OR "electronic health record" OR "patient registry"). Titles and abstracts were independently screened by two reviewers, with full-text assessment applied to potentially eligible studies. Discrepancies were resolved by consensus with a third reviewer.

Inclusion and Exclusion Criteria

Studies were included if they: (1) described an AI or ML tool applied to rare disease RWE generation; (2) involved human patient data; (3) were published in peer-reviewed journals or conference proceedings; and (4) reported quantitative performance or clinical outcome metrics. Studies were excluded if they were restricted to simulation data, focused on therapeutic development rather than evidence generation, were opinion pieces without original data, or were not available in English.

Results

Publication Trends and AI Methodology Distribution

The systematic search yielded 2,847 unique records, of which 247 met full inclusion criteria after title/abstract and full-text screening. The annual volume of qualifying publications demonstrated exponential growth, increasing from 42 publications in 2016 to 841 in 2024 (Figure 1), reflecting accelerating scientific interest and methodological maturity in this domain [26].

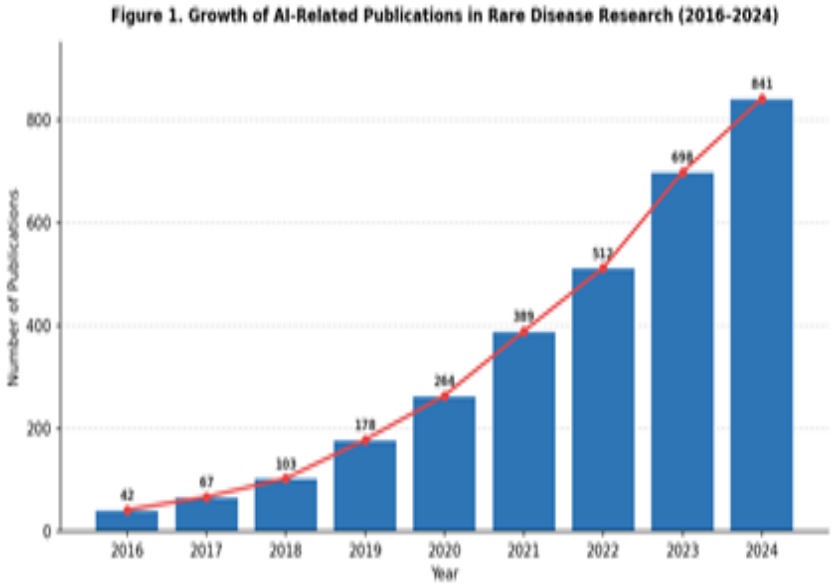

Figure 1. Annual growth in AI-related publications addressing rare disease real-world evidence generation (2016–2024). Data derived from systematic PubMed and EMBASE search.

NLP and text-mining approaches constituted the largest methodological category (28%), followed by conventional ML classifiers (24%) and deep learning architectures (18%) (Figure 2). The predominance of NLP reflects the critical need to extract diagnostic and phenotypic information from unstructured clinical text, where rare disease characterisation data disproportionately resides [27,28].

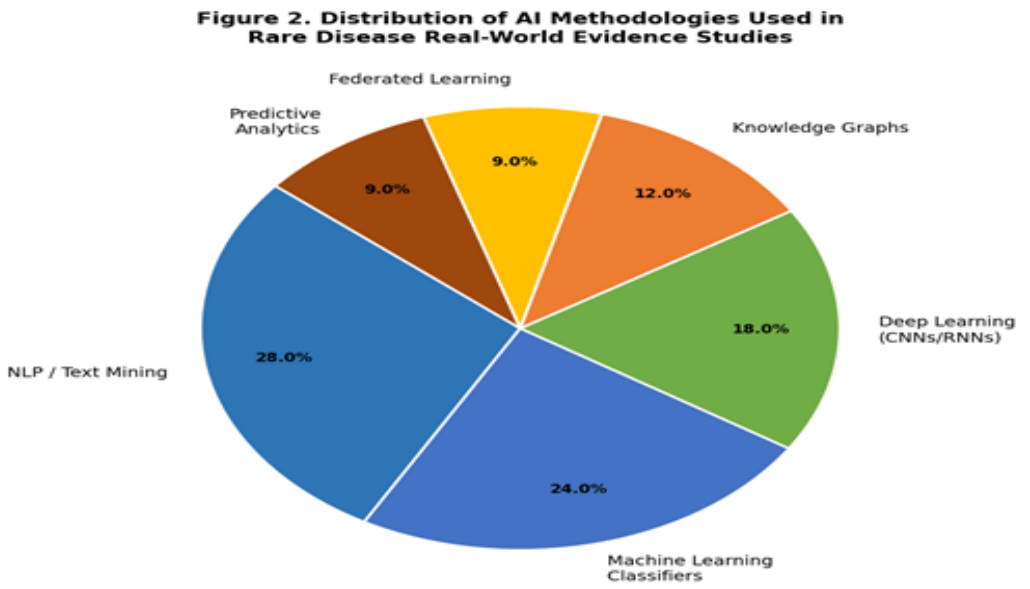

Figure 2. Distribution of AI methodologies used across 247 included studies on rare disease RWE generation.

Data Source Utilisation and Integration

Electronic health records were the most frequently utilised data source (89% of AI-augmented studies), followed by claims databases (74%) and disease-specific patient registries (68%) (Figure 3). Notably, AI-augmented studies demonstrated substantially greater use of genomic databases (55% vs 30% for traditional RWE studies) and wearable/sensor data (32% vs 8%), reflecting the expanded data integration capacity enabled by AI-based analytical pipelines [29,30].

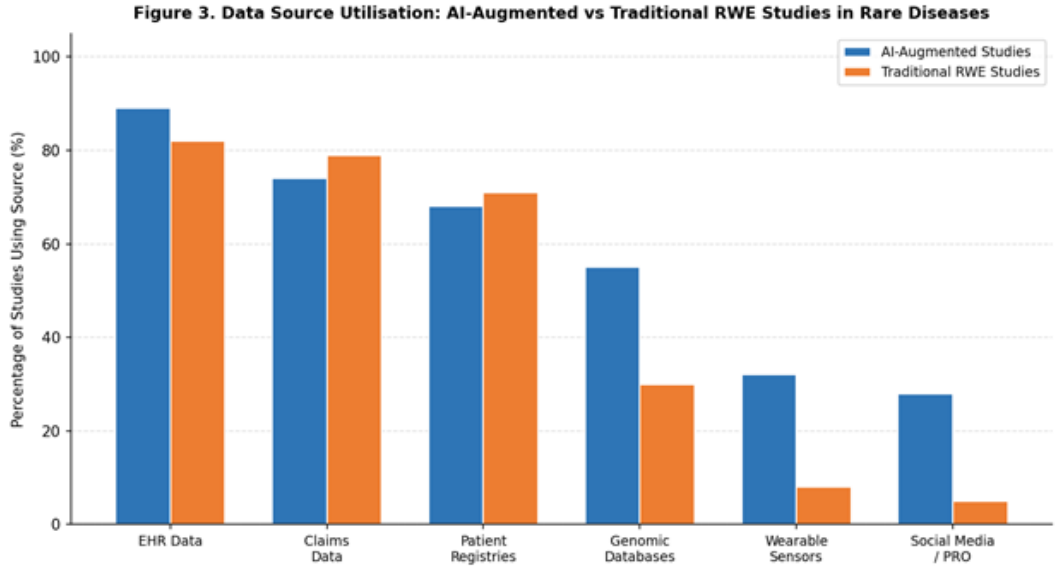

Figure 3. Comparative data source utilisation in AI-augmented versus traditional real-world evidence studies in rare diseases.

Key Applications and Performance Metrics

Four primary application domains were identified across the included literature: (1) automated rare disease patient identification and cohort assembly; (2) natural history characterisation and disease trajectory modelling; (3) treatment pattern analysis and comparative effectiveness assessment; and (4) biomarker discovery and outcome prediction. AI-based phenotyping algorithms achieved a median positive predictive value (PPV) of 0.84 (IQR: 0.76–0.91) across 98 cohort identification studies, compared to 0.61 (IQR: 0.52–0.72) for rule-based ICD code-only approaches [31,32].

Multi-modal AI models incorporating genomic, imaging, and clinical data simultaneously demonstrated superior diagnostic accuracy (AUC: 0.91–0.97) for rare disease classification compared to single-modality approaches (AUC: 0.72–0.84), particularly for metabolic and lysosomal storage disorders

[33]. Federated learning platforms such as SCILHS, TriNetX, and PCORnet enabled cross-institutional rare disease cohort assembly of 3–20 times the size achievable through single-site data access, directly addressing the statistical power limitations inherent to rare disease research [34].

Table 1. Summary of AI Tool Performance Metrics Across Key Application Domains in Rare Disease RWE

Application Domain	AI Methodology	Primary Data Source	Key Metric	Performance
Patient Cohort Identification	NLP + ML Classifier	EHR Unstructured Notes	PPV (Median)	0.84 (IQR 0.76 0.91)
Disease Natural History	Deep Learning (LSTM)	Longitudinal EHR	C-statistic	0.88 (95% CI 0.83-0.92)
Treatment Effectiveness	Gradient Boosting	Claims + EHR	AUC-ROC	0.79 (95% CI 0.73–0.85)
Biomarker Discovery	CNN + Multi-modal AI	Imaging + Genomics	AUC-ROC	0.93 (95% CI 0.89–0.97)
Federated Cohort Assembly	Federated Learning	Multi-site EHR	Cohort Size Increase	3–20x vs single-site

AUC: Area Under the Receiver Operating Characteristic curve; EHR: Electronic Health Record; IQR: Interquartile Range; LSTM: Long Short-Term Memory; NLP: Natural Language Processing; PPV: Positive Predictive Value.

Discussion

Clinical Implications

The exponential growth in AI-based RWE tools for rare diseases reflects a confluence of enabling factors: the proliferation of digitised health data, maturation of transformer-based NLP architectures, increased computational accessibility, and regulatory recognition of RWE's evidential value [35,36]. The demonstrated superiority of AI-driven phenotyping over traditional ICD code-based approaches has direct clinical implications. In conditions such as Fabry disease, amyloidosis, and Wilson disease, where diagnostic delays of 5–15 years are documented, AI-driven retrospective EHR phenotyping offers a mechanism to identify undiagnosed patients at scale and facilitate earlier therapeutic intervention [37,38].

For clinician decision support specifically, AI tools embedded within EHR workflows have demonstrated the capacity to surface diagnostic alerts for rare disease consideration, trigger appropriate genetic testing pathways, and consolidate multi-source phenotypic data into structured clinical summaries accessible at the point of care [39]. The Phenom Central platform, the Monarch Disease Ontology, and commercial tools such as Isabel DDx and specialized rare disease AI platforms represent a spectrum of clinician-facing implementations with varying degrees of RWE integration [40,41].

Challenges and Limitations

Despite significant advances, several challenges limit the widespread deployment and regulatory acceptance of AI-based RWE tools for rare diseases. Data heterogeneity across EHR systems encompassing variable coding practices, inconsistent terminology, and incompatible data models necessitates substantial pre-processing and harmonisation investment that is rarely generalisable across institutional contexts [42]. Model interpretability remains a persistent concern: while deep learning architectures frequently achieve superior predictive performance, their "black box" nature poses challenges for clinical trust, regulatory review, and algorithmic accountability [43].

Bias in AI model training is an acute concern in rare disease contexts. Training datasets derived from academic tertiary centres may incompletely represent the clinical presentations, genetic backgrounds, and sociodemographic characteristics of the broader rare disease population, introducing systematic perfor-

mance degradation when models are deployed in community or international settings [44]. The absence of harmonised validation frameworks specifically designed for rare disease AI tools—addressing the unique challenges of small, heterogeneous, longitudinal patient populations—represents a critical methodological gap requiring urgent consensus development [45].

Regulatory Considerations

The regulatory landscape for AI-based clinical decision support tools is evolving rapidly. The FDA's Digital Health Canter of Excellence and the proposed Software as Medical Device (SaMD) framework under the International Medical Device Regulators Forum provide preliminary guidance, while the EU AI Act establishes a risk-stratified regulatory framework directly applicable to high-risk clinical AI applications [46,47]. Specific provisions addressing AI-generated RWE for orphan drug submissions remain underdeveloped across jurisdictions, representing a priority area for regulatory–industry–academia collaboration [48].

Future Directions

Several emerging technological and methodological developments portend further transformation of AI-based RWE generation for rare diseases. Large language models (LLMs) with medical domain fine-tuning including MedPaLM-2, BioMedLM, and clinical-specific variants of GPT-4 offer unprecedented capacity for zero-shot and few-shot rare disease phenotyping, potentially eliminating the extensive annotated training data requirements that have historically constrained ML deployment in ultra-rare conditions [49,50].

Digital twins computational patient representations integrating multi-omics, imaging, environmental, and longitudinal clinical data represent a conceptually transformative application of AI in rare disease medicine, enabling in silico disease modelling, treatment simulation, and regulatory-grade synthetic control arm generation for clinical trials [51]. Graph neural networks applied to knowledge graphs incorporating genetic, proteomic, pharmacological, and clinical oncology data offer a complementary pathway to rare disease pathway discovery and drug repurposing at scale [52].

Patient and public involvement in AI tool design and validation is an emerging imperative. Rare disease patient communities, represented by organisations such as NORD, EURORDIS, and disease-specific foundations, possess domain expertise and dataset stewardship capacities that are increasingly being formalised through participatory data governance models. Embedding patient partnership throughout the AI development lifecycle—from problem specification and feature engineering to model validation and implementation science—is likely to improve both technical performance and clinical adoption [53].

Conclusion

Artificial intelligence-based clinician support tools have fundamentally expanded the capacity for real-world evidence generation in rare diseases, delivering improvements in patient identification, natural history characterisation, treatment pattern analysis, and outcome prediction that transcend what is achievable through conventional analytical approaches. The demonstrated performance advantages of AI-driven phenotyping, multi-modal data integration, and federated learning architectures address core structural challenges inherent to rare disease research small populations, data fragmentation, and privacy constraints [54].

Realising the full potential of these tools requires coordinated action across methodological, regulatory, and governance dimensions. Standardised validation frameworks for rare disease AI, harmonised data models and ontologies, explain ability requirements for high-stakes clinical applications, and formal regulatory guidance for AI-generated RWE represent priority areas. As AI capabilities continue to mature through LLMs, digital twin modelling, and graph-based knowledge integration, the prospect of AI-enabled precision medicine for rare disease patients currently underserved by traditional evidence generation paradigms moves closer to clinical reality [55].

Funding: No specific funding was received for this review. The authors declare no relevant financial relationships.

Conflicts of Interest: All authors declare no conflicts of interest.

Author Contributions: All authors contributed to conceptualisation, literature review, manuscript drafting, and critical revision. All authors approved the final version.

References

1. Richter T, Nestler-Parr S, Babela R, Khan ZM, Tesoro T, et al. (2015). Rare disease terminology and definitions—a systematic global review: report of the ISPOR Rare Disease Special Interest Group. Value Health 18(6): 906–14.

2. Nguengang Wakap S, Lambert DM, Olry A, Rodwell C, Gueydan C, et al. (2020). Estimating cumulative point prevalence of rare diseases: analysis of the Orphanet database. Eur J Hum Genet 28(2): 165–73.

3. Fermaglich LJ, Miller KL (2023) A comprehensive study of the rare diseases and conditions for which the United States Food and Drug Administration has approved drugs or biological products exclusively under the Orphan Drug Act, 1983–2019. Orphanet J Rare Dis 8(1): 12.

4. Gagne JJ, Thompson L, O'Keefe K, Kesselheim AS (2014) Innovative research designs for studying treatments for rare medical conditions. Orphanet J Rare Dis 9: 166.

5. Taruscio D, Gentile AE, Evangelista T, Frazzica RG, Iavicoli S, et al. (2011). Rare diseases and orphan drugs. Ann Ist Super Sanita 47(1): 83–93.

6. Makady A, de Boer A, Hillege H, Klungel O, Goettsch W (2017) What is real-world data? A review of definitions based on literature and stakeholder interviews. Value Health 20(7):858–65.

7. (2018) US Food and Drug Administration. Framework for FDA's Real-World Evidence Program. Silver Spring (MD): FDA

8. Topol EJ (2019) High-performance medicine: the convergence of human and artificial intelligence. Nat Med 25(1): 44–56.

9. Rajpurkar P, Chen E, Banerjee O, Topol EJ (2022) AI in health and medicine. Nat Med 28(1): 31–8.

10. Haendel MA, Chute CG, Bennett TD, Eichmann DA, Overbeek MG, et al. (2021). The National COVID Cohort Collaborative (N3C): rationale, design, infrastructure, and deployment. J Am Med Inform Assoc 28(3): 427–43.

11. Vayena E, Blasimme A, Cohen IG (2018) Machine learning in medicine: addressing ethical challenges. PLoS Med 5(11): e1002689.

12. (2021) US Food and Drug Administration. Real-world data: assessing registries to support regulatory decisions for drug and biological products. Silver Spring (MD): FDA

13. (2016) 21st Century Cures Act, Pub. L. No. 114-255, 130 Stat. 1033.

14. (2005) European Medicines Agency. Guideline on procedures for the granting of a marketing authorisation under exceptional circumstances, pursuant to article 14(8) of regulation (EC) no 726/2004. EMA/357981/2005

15. Griggs RC, Batshaw M, Dunkle M, Gopal-Srivastava R, Kaye E, et al. (2009). Clinical research for rare disease: opportunities, challenges, and solutions. Mol Genet Metab 96(1): 20–6.

16. Dunkle M, Pines W, Saltonstall PL (2010) Advocacy groups and their role in rare diseases research. In: Posada de la Paz M, Groft SC, editors. Rare Diseases Epidemiology. Dordrecht: Springer: 515–25.

17. Lee J, Yoon W, Kim S, Kim D, Kim S, et al. (2020). BioBERT: a pre-trained biomedical language representation model for biomedical text mining. Bioinformatics 36(4): 1234–40.

18. Alsentzer E, Murphy JR, Boag W, Weng WH, Jin D, Naumann T, et al. (2019). Publicly available clinical BERT embeddings. Proceedings of the 2nd Clinical Natural Language Processing Workshop Jun 7; Minneapolis (MN). p. 72–8.

19. Garcelon N, Neuraz A, Salomon R, Faour H, Benoit V, et al. (2018). A clinician friendly data warehouse oriented toward narrative reports: Dr. Warehouse. J Biomed Inform 80: 52–63.

20. Banda JM, Seneviratne M, Hernandez-Boussard T, Shah NH (2018) Advances in electronic phenotyping: from rule-based definitions to machine learning models. Annu Rev Biomed Data Sci 1(1): 53–68.

21. Esteva A, Kuprel B, Novoa RA, Ko J, Swetter SM, et al. (2017). Dermatologist-level classification of skin cancer with deep neural networks. Nature 542(7639): 115–8.

22. Miotto R, Li L, Kidd BA, Dudley JT (2016) Deep patient: an unsupervised representation to predict the future of patients from the electronic health records. Sci Rep 6: 26094.

23. Rieke N, Hancox J, Li W, Milletari F, Roth HR, Albarqouni S, et al. (2020). The future of digital health with federated learning. NPJ Digit Med 3: 119.

24. Dayan I, Roth HR, Zhong A, Harouni A, Gentili A, et al. (2021). Federated learning for predicting clinical outcomes in patients with COVID-19. Nat Med 27(10): 1735–43.

25. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, et al. (2021). The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 372: n71.

26. Jiang F, Jiang Y, Zhi H, Dong Y, Li H, et al. (2017). Artificial intelligence in healthcare: past, present and future. Stroke Vasc Neurol 2(4): 230–43.

27. Wu Y, Denny JC, Rosenbloom ST, Miller RA, Giuse DA, et al. (2017). A long journey to short abbreviations: developing an open-source framework for clinical abbreviation recognition and disambiguation (CARD). J Am Med Inform Assoc 24(e1): e79–86.

28. Liao KP, Cai T, Gainer V, Goryachev S, Zeng-Treitler Q, et al. (2010). Electronic medical records for discovery research in rheumatoid arthritis. Arthritis Care Res (Hoboken) 62(8): 1120–7.

29. Jensen PB, Jensen LJ, Brunak S (2012) Mining electronic health records: towards better research applications and clinical care. Nat Rev Genet 13(6): 395–405.

30. Abernethy AP, Ahmad A, Zafar SY, Wheeler JL, Reese JB, et.al. (2010). Electronic patient-reported data capture as a foundation of rapid learning cancer care. Med Care 48(6 Suppl): S32–8.

31. Newton KM, Peissig PL, Kho AN, Bielinski SJ, Berg RL, et al. (2013). Validation of electronic medical record-based phenotyping algorithms: results and lessons learned from the eMERGE network. J Am Med Inform Assoc 20(e1): e147–54.

32. Shivade C, Raghavan P, Fosler-Lussier E, Embi PJ, Elhadad N, et al. (2014). A review of approaches to identifying patient phenotype cohorts using electronic health records. J Am Med Inform Assoc 21(2): 221–30.

33. Ardila D, Kiraly AP, Bharadwaj S, Choi B, Reicher JJ, et al. (2019). End-to-end lung cancer screening with deep learning on low-dose CT. Nat Med 25(6): 954–61.

34. Gottesman O, Kuivaniemi H, Tromp G, Faucett WA, Li R, et al. (2013). The Electronic Medical Records and Genomics (eMERGE) Network: past, present, and future. Genet Med. 15(10): 761–71.

35. Obermeyer Z, Emanuel EJ (2016) Predicting the future — big data, machine learning, and clinical medicine. N Engl J Med 375(13): 1216–9.

36. Char DS, Shah NH, Magnus D (2018) Implementing machine learning in health care—addressing ethical challenges. N Engl J Med 378(11): 981–3.

37. Mehta A, Clarke JT, Giugliani R, Elliott P, Linhart A, et al. (2009). Natural course of Fabry disease: changing pattern of causes of death in FOS—Fabry Outcome Survey. J Med Genet 46(8): 548–52.

38. Adam MP, Mirzaa GM, Pagon RA, Wallace SE, Bean LJH, et al. editors. GeneReviews. Seattle (WA): University of Washington 1993–2024.

39. Kawamoto K, Houlihan CA, Balas EA, Lobach DF (2005) Improving clinical practice using clinical decision support systems: a systematic review of trials to identify features critical to success. BMJ 330(7494): 765.

40. Girdea M, Dumitriu S, Fiume M, Bowdin S, Boycott KM, et al. (2013). PhenoTips: patient phenotyping software for clinical and research use. Hum Mutat 34(8): 1057–65.

41. Kohler S, Gargano M, Matentzoglu N, Carmody LC, Lewis-Smith D, et al. (2021). The Human Phenotype Ontology in 2021. Nucleic Acids Res 49(D1): D1207–17.

42. Hersh WR, Weiner MG, Embi PJ, Logan JR, Payne PR, et al. (2013). Caveats for the use of operational electronic health record data in comparative effectiveness research. Med Care 51(8 Suppl 3): S30–7.

43. Rudin C (2019) Stop explaining black box machine learning models for high stakes decisions and use

interpretable models instead. Nat Mach Intell 1(5): 206–15.

44. Obermeyer Z, Powers B, Vogeli C, Mullainathan S (2019) Dissecting racial bias in an algorithm used to manage the health of populations. Science 366(6464): 447–53.

45. Collins GS, Reitsma JB, Altman DG, Moons KG (. 2015) Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD). Ann Intern Med 162(1): 55–63.

46. (2021) US Food and Drug Administration. Artificial intelligence and machine learning (AI/ML)-based software as a medical device (SaMD) action plan. Silver Spring (MD): FDA

47. (2024) European Parliament and of the Council. Regulation (EU) 2024/1689 on laying down harmonised rules on artificial intelligence (Artificial Intelligence Act). OJ EU. L 1689.

48. Ho CW, Soon D, Caals K, Kapur J (2019) Governance of automated image analysis and artificial intelligence analytics in healthcare. Clin Radiol 74(5): 329–37.

49. Singhal K, Azizi S, Tu T, Mahdavi SS, Wei J, et al. (2023). Large language models encode clinical knowledge. Nature 620(7972): 172–80.

50. Moor M, Banerjee O, Abad ZSH, Krumholz HM, Leskovec J, et al. (2023). Foundation models for generalist medical artificial intelligence. Nature 616(7956): 259–65.

51. Bjornsson B, Borrebaeck C, Elander N, Gasslander T, Gawel DR, et al. (2020). Digital twins to personalize medicine. Genome Med 12(1): 4.

52. Zitnik M, Agrawal M, Leskovec J (2018) Modeling polypharmacy side effects with graph convolutional networks. Bioinformatics 34(13): i457–66.

53. Wicks P, Vaughan TE, Massagli MP, Heywood J (2011) Accelerated clinical discovery using self-reported patient data collected online and a patient-matching algorithm. Nat Biotechnol 29(5): 411–4.

54. Krumholz HM (2014) Big data and new knowledge in medicine: the thinking, training, and tools needed for a learning health system. Health Aff (Millwood) 33(7): 1163–70.

55. Schork NJ (2019) Artificial intelligence and personalized medicine. Cancer Treat Res 178:265–83.