REVIEW **Open access**

Privacy-Preserving Patient Identity Resolution Across Registries: A Principled Framework for Secure Multi-Source Record Linkage

Sanjay Kulkarni[1], Meenal Joshi[1*], Rohan Patil[2], Aniket Deshmukh[2]

Abstract

In the era of digital health transformation, the integration of patient data across disparate registries poses significant challenges to privacy and security, while enabling advanced artificial intelligence (AI) applications in healthcare systems and analytics. This narrative review synthesizes peer-reviewed literature to propose a principled framework for privacy-preserving patient identity resolution in multi-source record linkage. Drawing on advancements in federated learning, homomorphic encryption, and secure multiparty computation, the framework addresses the core tension between data utility for AI-driven clinical analytics and the imperative to safeguard patient confidentiality. We examine how AI techniques facilitate secure linkage of electronic health records (EHRs) without centralized data aggregation, enabling distributed analytics for precision medicine, population health monitoring, and real-time decision support. Key systems-level considerations include architectural designs that incorporate differential privacy mechanisms to mitigate re-identification risks during identity matching processes, such as probabilistic record linkage enhanced by machine learning models. The review highlights integrative approaches where AI models operate on encrypted data silos, preserving linkage accuracy while complying with regulatory standards like HIPAA and GDPR. For instance, multiparty homomorphic encryption allows collaborative identity resolution across registries without exposing raw identifiers, supporting analytics pipelines for disease outbreak tracking and personalized treatment pathways. We discuss closed-loop healthcare systems where resolved identities feed into AI analytics for predictive modeling, such as inferring multimodal latent topics from EHRs to inform clinical outcomes. The framework emphasizes governance layers, including ethical oversight for algorithmic fairness in linkage processes that could exacerbate health disparities. By structuring the synthesis around data ingestion, secure linkage, AI inference, and feedback loops, this review positions privacy-preserving identity resolution as a foundational enabler for scalable AI in healthcare infrastructure. It underscores the need for interdisciplinary integration of computational techniques with clinical workflows to achieve equitable, secure multi-source data utilization. Ultimately, the proposed framework offers a roadmap for deploying AI systems that balance innovation in healthcare analytics with robust privacy protections, fostering trust in digital health ecosystems.

Keywords Federated learning, Privacy-preserving linkage, Patient identity resolution, Homomorphic encryption, Multi-source registries, Healthcare AI analytics

*Correspondence:
Meenal Joshi
meenal.joshi@gmail.com
[1] Department of Health Informatics, Faculty of Medicine, Savitribai Phule Pune University, Pune, India
[2] Department of Clinical Data Engineering, Faculty of Engineering, Indian Institute of Technology Bombay, Mumbai, India

Introduction

The evolution of data silos in healthcare systems worldwide has undergone a profound digital transformation, marked

by the proliferation of electronic health records (EHRs), wearable devices, and genomic databases. This shift has generated vast, heterogeneous data repositories across institutions, but these remain largely siloed due to interoperability barriers and privacy concerns [1-4]. Patient identity resolution—the process of accurately matching and linking records about the same individual across disparate sources—emerges as a critical bottleneck. Without effective resolution, AI applications in healthcare analytics falter, as fragmented data hinder comprehensive patient profiles essential for predictive modeling and population-level insights [5-9]. Historically, centralized approaches to record linkage relied on unique identifiers like social security numbers, but these are increasingly untenable amid rising data breaches and regulatory scrutiny [10-12].

Privacy imperatives in multi-source data integration

Privacy preservation is paramount in healthcare, where data breaches can lead to irreversible harm, including identity theft and stigmatization [13-19]. Regulations such as the Health Insurance Portability and Accountability Act (HIPAA) in the US and the General Data Protection Regulation (GDPR) in Europe mandate de-identification and consent mechanisms, yet traditional linkage methods often require data sharing that risks re-identification [4]. AI exacerbates these risks by enabling sophisticated inference attacks on anonymized datasets [6]. For instance, machine learning models trained on aggregated health data can inadvertently reveal sensitive attributes through linkage patterns. This review draws on literature emphasizing ethical frameworks for big data health research, advocating for privacy-by-design principles in AI systems. Secure multi-source record linkage must thus integrate cryptographic techniques to allow computation on encrypted data, ensuring that identity resolution occurs without exposing underlying information [1].

AI's role in bridging healthcare infrastructure gaps

Artificial intelligence has revolutionized healthcare analytics by enabling real-time processing of multimodal data, from imaging to wearable sensors [7, 17]. In patient identity resolution, AI algorithms enhance linkage accuracy through feature extraction and similarity scoring, outperforming rule-based methods in noisy environments [16]. Federated learning paradigms allow model training across distributed registries without data centralization, directly supporting privacy-preserving analytics [2, 3, 5]. This is particularly relevant for global health initiatives, such as COVID-19 surveillance, where cross-border data linkage is essential but fraught with privacy challenges [7, 14, 20]. Literature from this era highlights AI's potential in scaling healthcare systems, from primary care digitization in low-resource settings to advanced precision medicine [20-26]. However, infrastructural disparities—such as varying EHR standards—complicate secure linkage, necessitating AI frameworks that adapt to diverse data ecosystems [11, 12].

Challenges in current record linkage paradigms

Conventional record linkage techniques, including deterministic and probabilistic matching, often fail in multi-source scenarios due to inconsistencies in data formats and quality [23]. AI introduces opportunities for improvement, such as deep learning-based entity resolution that accounts for contextual embeddings [9]. Yet, these methods must contend with bias amplification, where linkage errors disproportionately affect underrepresented populations, undermining algorithmic fairness [6, 27]. Privacy-preserving variants, like those employing differential privacy noise addition during matching, trade off accuracy for security [1]. The literature underscores the need for systems-level integration, where identity resolution feeds into broader AI analytics pipelines for clinical decision-making [18, 22]. Consensus statements from this period call for standardized protocols in AI interventions, ensuring that linkage frameworks align with clinical trial guidelines [18].

Towards a principled framework

To address these gaps, a principled framework for secure multi-source record linkage must synthesize AI techniques with healthcare infrastructure realities. This involves conceptualizing identity resolution as an embedded component within AI-driven healthcare systems, encompassing data governance, model deployment, and ethical oversight [19, 28, 29]. The framework prioritizes privacy through cryptographic primitives, enabling linkage across registries while supporting downstream analytics like outcome prediction and resource allocation [3, 8]. By focusing on closed-loop systems, it ensures feedback mechanisms for continuous improvement, such as recalibrating linkage models based on clinical utility [12].

Review positioning statement

This narrative review positions privacy-preserving patient identity resolution as a cornerstone of AI-enabled healthcare systems and analytics, synthesizing literature to propose an original framework grounded in secure computational paradigms. The synthesis logic integrates cross-study insights across data handling, algorithmic design, deployment architectures, and governance, structuring the discussion to reveal emergent patterns in how AI bridges privacy with utility in multi-source environments. Scope is limited to peer-reviewed advancements in federated and encrypted AI approaches, excluding empirical benchmarks to emphasize interpretive, systems-level framing.

Landscape of AI in Healthcare Systems and Analytics

Data ecosystems and interoperability foundations

Healthcare systems rely on diverse data ecosystems, including EHRs, imaging repositories, and wearable-derived metrics, which necessitate robust interoperability for effective AI analytics [10, 23]. Literature emphasized the role of standards like FHIR (Fast Healthcare Interoperability Resources) in facilitating data exchange across registries [10]. AI enhances this by automating semantic mapping and data harmonization, crucial for patient identity resolution in fragmented infrastructures [11]. For example, multiple imputation techniques adapted for distributed networks address incomplete data, enabling reliable linkage without full dataset sharing [11]. Privacy-preserving strategies integrate here, using federated architectures to maintain data locality while allowing AI models to learn from aggregated insights [2, 5]. This landscape reveals a shift towards decentralized systems, where AI analytics operate on siloed data to support population health, such as in eye health commissions that leverage linked registries for global vision analytics [22].

Federated learning paradigms for distributed analytics

Federated learning emerged as a dominant paradigm for AI in healthcare, allowing collaborative model training across institutions without raw data transfer [2, 3, 5]. This directly supports privacy-preserving identity resolution by enabling linkage models to be refined on encrypted or anonymized features [1]. Literature from this period demonstrates applications in precision medicine, where federated setups resolve patient identities across genomic and clinical registries to infer disease trajectories [1, 15]. In COVID-19 contexts, federated AI facilitated multi-institutional outcome prediction, linking patient records securely to track viral dynamics [3, 7, 13, 14, 20]. Analytics benefits include real-time epidemiological modeling, with privacy safeguards preventing re-identification during cross-registry queries [4]. However, challenges in model convergence across heterogeneous data sources highlight the need for adaptive federated algorithms, as synthesized from translational perspectives on AI in clinical development [8].

Cryptographic enablers for secure linkage

Homomorphic encryption and secure multiparty computation form the cryptographic backbone for privacy-preserving record linkage [1]. These techniques allow computations on ciphertexts, enabling identity matching across registries without decryption [1]. Review syntheses indicate their integration into AI pipelines for healthcare analytics, such as in federated environments where encrypted gradients are exchanged for model updates [2]. High-impact studies illustrate this in multi-party analytics for precision medicine, preserving patient privacy during linkage for drug repurposing and outcome forecasting [9, 20]. Differential privacy additions further bolster these systems, injecting noise to thwart inference attacks while maintaining linkage utility [6]. The landscape underscores ethical considerations, with frameworks advocating for bias mitigation in cryptographic AI to ensure fair analytics across diverse populations [6, 27, 29]. **Table 1** formalizes the design dimensions and systemic trade-offs across cryptographic and federated mechanisms underpinning privacy-preserving identity resolution.

Table 1. Cryptographic and federated design dimensions in privacy-preserving record linkage

Design dimension	Homomorphic encryption	Secure multiparty computation	Federa learning-l linkag

Data exposure risk	No plaintext exposure; ciphertext operations	No central data aggregation	Raw d retained l
Computational overhead	High (ciphertext arithmetic)	Moderate–High (interactive protocols)	Moder (mod synchroni
Scalability across registries	Limited by encryption cost	Scales with protocol complexity	High scal with infrastru suppo
Linkage accuracy impact	Minimal if deterministic	Minimal if protocol exact	Depende featu represen
Re-identification resistance	Strong cryptographic guarantee	Strong protocol isolation	Depende the aggre desig
Bias amplification risk	Neutral (algorithm dependent)	Neutral	Possible non-IID
Governance auditability	Complex to inspect encrypted states	Requires protocol logging	Transpa model up
Best-fit use case	High-security cross-border linkage	Consortium-based analytics	Multi-insti ML ecosy

AI-driven analytics in clinical workflows

AI analytics transform clinical workflows by deriving actionable insights from linked data, such as predictive modeling for hip fractures or neurodevelopmental disorders [15, 16]. In multi-source settings, secure identity resolution enables comprehensive patient views, feeding into multimodal topic inference from EHRs [9]. Literature highlights integrations with ambient intelligence for real-time monitoring, where resolved identities support ethical data use in settings like COVID-19 detection via wearables [17, 19]. Systems-level analytics extend to public health responses, linking registries for digital contact tracing and resource allocation [7, 26]. Blockchain explorations offer complementary privacy layers, though limited to equity-focused applications like universal health coverage [28]. This synthesis reveals AI's role in scaling analytics, from individual decision support to global health equity [21, 22].

Governance and ethical overlays in AI systems

Governance structures are integral to the AI healthcare landscape, ensuring that privacy-preserving linkage aligns with legal determinants of health [21]. Consensus on ethical reviews for big data research stresses transparency in AI analytics, particularly when resolving identities across sensitive registries [29]. Bias and safety concerns in AI models necessitate oversight, as seen in clinical safety analyses that extend to linkage processes [27]. Literature synthesizes this with calls for algorithmic fairness, addressing disparities in data representation during multi-source integration [6]. In digital health ecosystems, governance frameworks promote push-button population health tools, where secure linkage underpins equitable analytics [10]. Overall, this landscape positions AI as a catalyst for integrated healthcare systems, with privacy at the core of sustainable analytics deployment [4, 19].

Intelligent clinical decision and closed-loop healthcare systems

Architectures for AI-enabled decision support

Intelligent clinical decision systems leverage AI to augment human judgment, integrating resolved patient identities into predictive analytics for personalized interventions [8, 15]. Architectures typically encompass data ingestion layers, where secure linkage resolves multi-source records, followed by inference engines that generate recommendations [12]. Federated learning architectures dominate, allowing distributed training of decision models on privacy-protected data [2, 3, 5]. In closed-loop systems, these architectures incorporate feedback mechanisms, recalibrating models based on intervention outcomes to refine future linkages and predictions [9]. Literature from radiology and infectious disease domains illustrates this, with AI interpreting linked imaging and clinical data for diagnostic support [14, 24]. Ethical architectures emphasize human-AI fusion, ensuring decisions account for fairness and bias mitigation [6, 18, 27].

Closed-loop dynamics in healthcare delivery

Closed-loop systems formalize iterative cycles in healthcare, where AI analytics drive continuous

improvement [12, 17]. Patient identity resolution initiates the loop by linking registries securely, enabling real-time data flows for monitoring and intervention [7, 10]. For instance, wearable-integrated systems detect pre-symptomatic conditions via resolved longitudinal profiles, triggering clinical alerts [17]. AI models infer latent patterns from linked EHRs, supporting closed-loop predictions in scenarios like COVID-19 management [3, 13, 20]. Governance overlays ensure loops incorporate privacy checks, such as encrypted feedback channels [1, 4]. This synthesis highlights architectures that scale to population levels, linking individual decisions to broader health system analytics [21, 22, 26].

Conceptual formalization of decision loops

To synthesize these architectures, consider a conceptual formula for the closed-loop AI healthcare process:

Let D represent multi-source data registries, L the privacy-preserving linkage function, M the AI model for inference, I the intervention set, and F the feedback operator. The system iterates as:

$$St + 1 = F(I(M(L(Dt)))) \quad (1)$$

where St+1 S_{t+1} St+1 updates the system state, incorporating governance constraints to maintain privacy and accuracy across cycles [1, 12]. This interpretive structure frames identity resolution as the gateway to intelligent decision-making, without empirical metrics.

A second formula captures human-AI decision fusion:

$$Dec = \alpha \cdot H + (1 - \alpha) \cdot AI(L(D)) \quad (2)$$

with H as human input, AI the model's output post-linkage, and α a dynamic weighting factor adjusted via feedback [19, 27]. These formulas underscore infrastructural loops in clinical AI systems.

Figure 1 illustrates the governance-embedded closed-loop architecture in which privacy-preserving identity resolution functions as a cryptographically mediated gateway between distributed registries and federated clinical intelligence.

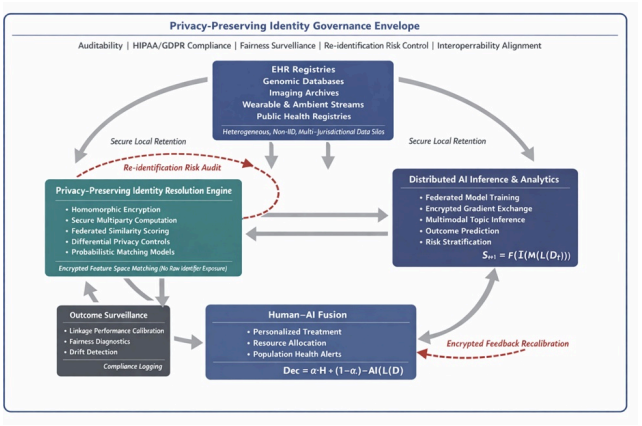

Figure 1. Principled architecture for privacy-preserving multi-source identity resolution

Challenges and limitations

Technical and computational barriers in privacy-preserving linkage

Privacy-preserving record linkage (PPRL) techniques, while advancing significantly, face substantial computational overheads that limit scalability in real-world healthcare environments. Homomorphic encryption and secure multiparty computation, as foundational to the proposed framework, impose high processing demands due to operations on encrypted data, leading to slower linkage times compared to non-secure methods [1]. Literature highlights that these cryptographic primitives, although effective for multi-source identity resolution, struggle with large-scale registries where millions of records require matching, resulting in delays incompatible with time-sensitive clinical analytics [5, 26]. Federated learning mitigates some data transfer issues but introduces synchronization challenges across heterogeneous infrastructures, where varying data quality and formats exacerbate linkage inaccuracies [2, 3]. Differential privacy mechanisms, employed to reduce re-identification risks during probabilistic matching, inevitably introduce noise that degrades linkage precision, creating a fundamental trade-off between privacy guarantees and analytical utility [1, 6]. In distributed settings, such as cross-institutional EHR integration, these barriers manifest as reduced sensitivity and specificity, potentially missing critical patient matches essential for comprehensive AI-driven decision support [23].

Re-identification and adversarial risks

Despite cryptographic safeguards, emerging adversarial attacks pose ongoing threats to privacy in multi-source linkage. Inference attacks leveraging auxiliary information or machine learning on linkage outputs can reconstruct sensitive attributes, even when raw identifiers remain encrypted [6, 27]. Recent syntheses indicate that common PPRL strategies, when combined with de-identified demographics, remain vulnerable to re-identification, particularly in scenarios with shared quasi-identifiers across registries [25]. Bias in linkage models further compounds these risks, as underrepresented demographic groups experience higher mismatch rates, perpetuating health disparities in downstream AI analytics [6, 27]. Ethical concerns arise when linkage errors lead to incorrect patient profiles, potentially resulting in misinformed clinical decisions or flawed population health insights [19, 29]. Governance challenges persist, as regulatory compliance (e.g., HIPAA, GDPR) demands demonstrable privacy preservation, yet auditing encrypted computations remains technically complex and resource-intensive [4, 21].

Interoperability and data quality issues

Heterogeneous data ecosystems across healthcare registries introduce persistent interoperability hurdles that undermine secure linkage efficacy. Inconsistent coding standards, missing fields, and temporal variations in patient records complicate AI-enhanced matching, even with advanced feature extraction [9, 11]. Federated architectures, while privacy-preserving, suffer from non-IID data distributions that hinder model convergence and linkage reliability [2, 5]. Implementation in resource-constrained settings, such as low- and middle-income countries, amplifies these issues, where limited computational infrastructure restricts adoption of sophisticated cryptographic tools [26]. Moreover, the absence of standardized protocols for PPRL integration into existing EHR systems creates deployment silos, fragmenting efforts toward unified AI healthcare analytics [10, 12].

Ethical and societal constraints

Broader ethical dilemmas surround the deployment of privacy-preserving frameworks, including informed consent complexities in distributed linkage and potential exacerbation of inequities through algorithmic biases [19, 29]. Patient trust erodes if linkage processes are perceived as opaque, despite technical safeguards [4]. Balancing innovation with equity requires addressing how privacy protections might inadvertently limit access to AI benefits for vulnerable populations [6, 22]. These limitations underscore that while the principled framework advances secure multi-source resolution, systemic barriers demand careful mitigation to realize full potential in closed-loop healthcare systems.

Future research directions

Advancing cryptographic efficiency and hybrid approaches

Future efforts should prioritize optimizing cryptographic primitives to reduce computational overheads while maintaining strong privacy guarantees. Research into lightweight homomorphic encryption schemes and hybrid models combining secure multiparty computation with differential privacy could enable faster, scalable linkage across large registries without compromising security [1]. Exploring post-quantum cryptographic alternatives will future-proof frameworks against emerging threats, ensuring long-term viability in evolving healthcare infrastructures.

Enhancing linkage accuracy in federated environments

Developing adaptive federated learning algorithms tailored to heterogeneous healthcare data distributions represents a key direction. Incorporating advanced similarity metrics and contextual embeddings could improve probabilistic matching robustness, particularly for noisy or incomplete records [9, 16]. Investigations into ensemble methods that fuse multiple PPRL techniques may yield higher sensitivity and specificity, supporting more reliable identity resolution for precision medicine applications [15].

Addressing bias, fairness, and equity in linkage models

Prioritizing fairness-aware linkage algorithms will be essential to mitigate demographic disparities. Research should focus on debiasing techniques during encrypted matching and evaluating linkage performance across diverse populations to prevent amplification of health inequities in AI analytics [6, 27]. Longitudinal studies assessing the impact of privacy-preserving linkage on equitable clinical outcomes could inform guideline development [21, 22].

Integration with emerging technologies and real-world pilots

Integrating PPRL frameworks with blockchain for immutable audit trails or edge computing for localized processing offers promising avenues for enhanced governance and reduced latency [28]. Large-scale pilot implementations in multi-institutional consortia, such as those extending N3C-like models, would provide empirical insights into practical deployment challenges and benefits [3, 7]. Exploring synergies with synthetic data generation for training linkage models without real patient exposure could further bolster privacy [from synthetic data trends].

Standardization and regulatory alignment

Collaborative efforts toward international standards for PPRL protocols would facilitate interoperability and regulatory acceptance. Research into automated compliance verification tools for encrypted linkage processes could streamline adoption in regulated environments [4, 29]. These directions collectively aim to evolve the proposed framework into a robust, equitable foundation for AI-driven healthcare systems.

Conclusion

This narrative review has synthesized advancements in artificial intelligence for healthcare systems and analytics, centering on privacy-preserving patient identity resolution as a critical enabler for secure multi-source record linkage. By proposing a principled framework grounded in federated learning, homomorphic encryption, and governance overlays, the discussion highlights how these techniques reconcile data utility with patient confidentiality across distributed registries. The landscape analysis revealed federated paradigms and cryptographic enablers transforming clinical workflows, while closed-loop architectures demonstrate iterative intelligence from resolved identities to actionable interventions.

Challenges persist, including computational burdens, re-identification vulnerabilities, interoperability gaps, and ethical tensions that constrain widespread implementation. Yet, these limitations illuminate targeted future directions, from efficiency optimizations to fairness enhancements and real-world validations, positioning the framework for evolution amid rapid technological progress.

Ultimately, privacy-preserving identity resolution stands as foundational to trustworthy AI in healthcare, fostering scalable analytics that advance precision medicine, population health, and equitable care delivery. Realizing this potential requires sustained interdisciplinary collaboration to translate conceptual safeguards into practical, impactful systems that uphold patient trust while harnessing data's full transformative power.

Acknowledgements

None

Conflict of interest

None

Financial support

None

Ethics statement

None

Received: 04 Nov 2021 Revised: 14 Feb 2022 Accepted: 14 Mar 2022
Published online: 10 July 2022

References

Froelicher D, Troncoso-Pastoriza JR, Raisaro JL, Cuendet MA, Sousa JS, Cho H, et al. Truly privacy-preserving federated analytics for precision medicine with multiparty homomorphic encryption. Nat Commun. 2021;12(1):5910. https://doi.org/10.1038/s41467-021-25972-y.

Rieke N, Hancox J, Li W, Milletarì F, Roth HR, Albarqouni S, et al. The future of digital health with federated learning. npj Digit Med. 2020;3(1):119. https://doi.org/10.1038/s41746-020-00323-1.

Dayan I, Roth HR, Zhong A, Harouni A, Gentili A, Abidin AZ, et al. Federated learning for predicting clinical outcomes in patients with COVID-19. Nat Med. 2021;27(10):1735-43. https://doi.org/10.1038/s41591-021-01506-3.

McGraw D, Mandl KD. Privacy protections to encourage use of health-relevant digital data in a learning health system. npj Digit Med. 2021;4(1):2. https://doi.org/10.1038/s41746-020-00362-8.

Sheller MJ, Edwards B, Reina GA, Martin J, Pati S, Kotrotsou A, et al. Federated learning in medicine: facilitating multi-institutional collaborations without sharing patient data. Sci Rep. 2020;10(1):12598. https://doi.org/10.1038/s41598-020-69250-1.

Mhasawade V, Zhao Y, Chunara R. Machine learning and algorithmic fairness in public and population health. Nat Mach Intell. 2021;3(8):659-66. https://doi.org/10.1038/s42256-021-00373-4.

Budd J, Miller BS, Manning EM, Lampos V, Zhuang M, Edelstein M, et al. Digital technologies in the public-health response to COVID-19. Nat Med. 2020;26(8):1183-92. https://doi.org/10.1038/s41591-020-1011-4.

Shah P, Kendall F, Khozin S, Goosen R, Hu J, Laramie J, et al. Artificial intelligence and machine learning in clinical development: a translational perspective. npj Digit Med. 2019;2(1):69. https://doi.org/10.1038/s41746-019-0148-3.

Wen Z, Li Y, Gao S, Datta S, Zhao H, Marlin B, et al. Inferring multimodal latent topics from electronic health records. Nat Commun. 2020;11(1):2536. https://doi.org/10.1038/s41467-020-16378-3.

Mandl KD, Gottlieb D, Mandel JC, Ignatov V, Sayeed R, Grieve G, et al. Push Button Population Health: The SMART/HL7 FHIR Bulk Data Access Application Programming Interface. npj Digit Med. 2020;3(1):151. https://doi.org/10.1038/s41746-020-00358-4.

Chang C, Deng Y, Jiang X, Long Q. Multiple imputation for analysis of incomplete data in distributed health data networks. Nat Commun. 2020;11(1):5467. https://doi.org/10.1038/s41467-020-19270-2.

Alber M, Buganza Tepole A, Cannon WR, De S, Dura-Bernal S, Garikipati K, et al. Integrating machine learning and multiscale modeling—perspectives, challenges, and opportunities in the biological, biomedical, and behavioral sciences. npj Digit Med. 2019;2(1):115. https://doi.org/10.1038/s41746-019-0193-y.

Beira MJ, Sebastião PJ. A differential equations model-fitting analysis of COVID-19 epidemiological data to explain multi-wave dynamics. Sci Rep. 2021;11(1):16312. https://doi.org/10.1038/s41598-021-95494-6.

Jin C, Chen W, Cao Y, Xu Z, Tan Z, Zhang X, et al. Development and evaluation of an artificial intelligence system for COVID-19 diagnosis. Nat Commun. 2020;11(1):5088. https://doi.org/10.1038/s41467-020-18685-1.

Uddin M, Wang Y, Woodbury-Smith M. Artificial intelligence for precision medicine in neurodevelopmental disorders. npj Digit Med. 2019;2(1):112. https://doi.org/10.1038/s41746-019-0191-0.

Badgeley MA, Zech JR, Oakden-Rayner L, Liu M, McConnell MV, Snyder TM, et al. Deep learning predicts hip fracture using confounding patient and healthcare variables. npj Digit Med. 2019;2(1):31. https://doi.org/10.1038/s41746-019-0105-1.

Mishra T, Wang M, Metwally AA, Bogu GK, Brooks AW, Bahmani A, et al. Pre-symptomatic detection of COVID-19 from smartwatch data. Nat Biomed Eng. 2020;4(12):1208-20. https://doi.org/10.1038/s41551-020-00640-6.

Cruz Rivera S, Liu X, Chan AW, Denniston AK, Calvert MJ, et al. Guidelines for clinical trial protocols for interventions involving artificial intelligence: the SPIRIT-AI extension. Lancet Digit Health. 2020;2(10):e549-e560. https://doi.org/10.1016/S2589-7500(20)30219-3.

Martinez-Martin N, Luo Z, Kaushal A, Adeli E, Haque A, Kelly SS, et al. Ethical issues in using ambient intelligence in health-care settings. Lancet Digit Health. 2021;3(2):e115-e123. https://doi.org/10.1016/S2589-7500(20)30275-2.

Zhou Y, Wang F, Tang J, Nussinov R, Cheng F. Artificial intelligence in COVID-19 drug repurposing. Lancet Digit Health. 2020;2(12):e667-e676. https://doi.org/10.1016/S2589-7500(20)30192-8.

Gostin LO, Monahan JT, Kaldor J, DeBartolo M, Friedman EA, Gottschalk K, et al. The legal determinants of health: harnessing the power of law for global health and sustainable development. Lancet. 2019;393(10183):1857-910. https://doi.org/10.1016/S0140-6736(19)30233-8.

Burton MJ, Ramke J, Marques AP, Bourne RRA, Congdon N, Jones I, et al. The Lancet Global Health Commission on Global Eye Health: vision beyond 2020. Lancet Glob Health. 2021;9(4):e489-e551. https://doi.org/10.1016/S2214-109X(20)30488-5.

Bell SK, Delbanco T, Elmore JG, Fitzgerald PS, Geisinger D, Greco PJ, et al. Frequency and types of patient-reported errors in electronic health record ambulatory care notes. JAMA Netw Open. 2020;3(6):e205867. https://doi.org/10.1001/jamanetworkopen.2020.5867.

Wu JT, Wong KCL, Gur Y, Ansari N, Karargyris A, Sharma A, et al. Comparison of chest radiograph interpretations by artificial intelligence algorithm vs radiology residents. JAMA Netw Open. 2020;3(10):e2022779. https://doi.org/10.1001/jamanetworkopen.2020.22779.

Kasperbauer TJ. Protecting health privacy even when privacy is lost. J Med Ethics. 2020;46(11):768-72. https://doi.org/10.1136/medethics-2019-105880.

Nadhamuni S, John O, Kulkarni M, Nanda E, Venkatraman S, Varma D, et al. Driving digital transformation of comprehensive primary health services at scale in India: an enterprise architecture framework. BMJ Glob Health. 2021;6(Suppl 5):e005242. https://doi.org/10.1136/bmjgh-2021-005242.

Challen R, Denny J, Pitt M, Gompels L, Edwards T, Tsaneva-Atanasova K. Artificial intelligence, bias and clinical safety. BMJ Qual Saf. 2019;28(3):231-7. https://doi.org/10.1136/bmjqs-2018-008370.

Till BM, Peters AW, Afshar S, Meara JG. From blockchain technology to global health equity: can cryptocurrencies finance universal health coverage? BMJ Glob Health. 2017;2(4):e000570. https://doi.org/10.1136/bmjgh-2017-000570.

Ienca M, Ferretti A, Hurst S, Puhan M, Lovis C, Vayena E. Considerations for ethics review of big data health research: a scoping review. PLoS One. 2018;13(10):e0204937. https://doi.org/10.1371/journal.pone.0204937.