Original Article

Synthetic Pharmacy Records without Synthetic Certainty

Yifan Zhao[1]*, Liang Chen[1], Ahmad Fauzi[2], Nor Hayati[3]

[1]Department of Synthetic Data and Pharmacy Records, Faculty of Pharmacy, China Agricultural University, Beijing, China. [2]Department of Synthetic Certainty and Validation, Faculty of Pharmacy, University of Malaya, Kuala Lumpur, Malaysia. [3]Department of Data Quality and AI in Pharmacy, Faculty of Pharmacy, Universiti Putra Malaysia, Serdang, Malaysia.

Abstract

Synthetic pharmacy records may widen access to data for software development, education, interoperability testing, and algorithm research, but plausible records can create false confidence when realism is treated as evidence of privacy, fairness, medication safety, or decision readiness. This article develops an original, non-empirical governance architecture for synthetic pharmacy records. The problem is defined as synthetic certainty: the unwarranted conversion of generated plausibility into claims about observed patients, valid clinical relationships, safe workflow use, or authorized downstream decisions. The proposed architecture begins with an intended-use and consequence specification, then connects the source-data envelope, generator configuration and versioning, provenance, separate privacy–fidelity–bias–disclosure controls, independent use-conditioned evaluation, bounded release states, and continuing monitoring, expiry, withdrawal, and supersession. The architecture separates model output from medication decisions, technical performance from clinical usefulness, association from causation, automation from professional authority, and implementation from validated benefit. Evaluation therefore requires structural, temporal, relational, task-specific, privacy, subgroup, human-factors, medication-safety, and organizational evidence selected for the declared use. The contribution is not a validated standard, regulatory pathway, or deployment manual. Its release states and decision gates require empirical calibration across pharmacy settings, populations, data environments, and jurisdictions. The central proposition is that synthetic records should be governed as versioned, purpose-bounded, auditable artifacts whose uncertainty remains attached throughout generation, release, reuse, and withdrawal.

Keywords: Digital pharmacy, Artificial intelligence, Pharmacy practice, Medication safety, Human–AI collaboration, Clinical decision support

Introduction

Synthetic health data can support innovation and access while retaining unresolved questions about privacy, validity, representation, and governance [1]. In pharmacy, this tension is especially consequential because medication records combine prescriptions, dispensing, administration, laboratory results, diagnoses, adherence signals, alerts, and professional actions. A generated record may look coherent while encoding relationships that were never observed or while omitting rare but safety-critical patterns. This article uses synthetic pharmacy record to mean generated medication-related patient or encounter data that do not directly reproduce an observed individual record.

Evaluation research shows that synthetic-data quality is multidimensional: distributional resemblance, privacy protection, and downstream utility measure different properties and cannot be collapsed safely into visual plausibility [2]. The proposed concept of synthetic certainty therefore describes an epistemic error rather than a generator type. It occurs when a credible-looking record is interpreted as establishing population prevalence, causal treatment response, clinical appropriateness, fairness, anonymity, or workflow fitness beyond the evidence used to generate and evaluate it.

Downstream utility is also task-dependent. Synthetic records may approximate real-data performance for a defined software-assessment task, yet that result remains conditional on the target model, comparator, data-generating process, and evaluation design [3]. Alternative explanations for apparent success include leakage, overfitting, preserved easy patterns, insensitive endpoints, or evaluation on tasks that do not expose medication-related failure modes. Consequently, a synthetic dataset cannot inherit a general label such as “high fidelity” without stating the use for which that claim was tested.

Address for correspondence: Yifan Zhao, Department of Synthetic Data and Pharmacy Records, Faculty of Pharmacy, China Agricultural University, Beijing, China.
yifan.zhao@cau.edu.cn
Received: 16 February 2026; **Accepted:** 28 April 2026

How to cite this article: Zhao Y, Chen L, Fauzi A, Hayati N. Synthetic Pharmacy Records without Synthetic Certainty. Arch Pharm Pract. 2026;17(2):39-46. https://doi.org/10.51847/tzcMgfXH1q

Medication-system evidence further shows why technical performance must be separated from safety and clinical consequence. Electronic medication systems can reduce some errors while leaving heterogeneous effects and system-related harms, so implementation alone does not establish safer care [4]. By analogy, synthetic records may be useful technical artifacts without being adequate evidence for prescribing, dispensing, monitoring, or alert decisions. The original contribution developed here is a proposed governance architecture that binds every synthetic pharmacy-record release to intended use, provenance, control evidence, accountable authority, and explicit decision boundaries.

Synthetic-Record Uses and Risk Categories

Rule-based synthetic records can support software development, education, and interoperability testing without representing actual patients [5]. In pharmacy, corresponding low-consequence uses may include interface testing, terminology mapping, training exercises, and demonstration environments, provided the records are not reinterpreted as epidemiological or clinical evidence. Higher-consequence uses include algorithm training, medication-alert evaluation, workflow simulation, comparative methods research, and any activity that may influence patient-facing decisions.

Longitudinal generation introduces a separate risk class because medication meaning depends on sequence and timing. Event order, initiation and discontinuation, refill intervals, treatment persistence, laboratory follow-up, and censoring must be evaluated explicitly rather than inferred from cross-sectional similarity [6]. A record can have plausible individual fields while presenting an impossible medication trajectory.

Mixed-type records create another risk class. Categorical diagnoses, continuous laboratory values, free-text-derived variables, medication codes, missingness, and repeated measures can appear acceptable separately while their conditional relationships are distorted [7]. Pharmacy-relevant examples include incompatible dose–renal-function patterns, implausible drug–laboratory combinations, or missingness that erases differential monitoring.

Relational generation adds topology and co-occurrence risks. Graph-based synthetic records may preserve some links among encounters, conditions, and interventions, but graph resemblance does not establish clinical validity and may still expose unusual relationship patterns [8]. The proposed taxonomy therefore distinguishes structural, temporal, relational, privacy, bias, provenance, misuse, and workflow-consequence risks.

Risk category is proposed to depend on the interaction of intended use, decision consequence, record structure, access mode, linkage potential, downstream automation, and representational uncertainty [5-7]. These dimensions organize review and release; they do not create universal risk scores or prove that any category is safe.

Table 1 organizes the evidence, constructs, relationships, uncertainties, and boundary conditions required for definitions, components, relationships, and intended uses for the article Synthetic Pharmacy Records without Synthetic Certainty.

Table 1. Definitions, components, relationships, and intended uses for the article Synthetic Pharmacy Records without Synthetic Certainty.

Component or Scientific Dimension	Problem Addressed	Inputs or Determinants	Proposed Mechanism or Relationship	Expected Contribution	Evidence Required	Failure or Uncertainty Risk	Boundary Statement
Synthetic Pharmacy Record	Generated data may be mistaken for observation	Generator, source-data scope, medication variables	Treat as a versioned artifact, not a patient record	Enables bounded non-patient uses	Generation and use documentation [1, 5]	False clinical or epidemiological inference	Plausibility does not establish truth
Intended-Use Class	One dataset may be reused beyond its evidence	User, task, setting, consequence	Proposed use contract precedes evaluation	Aligns tests with decision purpose	Task-specific utility evidence [3]	Scope drift	Approval for one use does not transfer
Temporal Structure	Marginal similarity can conceal impossible sequences	Order, intervals, persistence, censoring	Evaluate medication trajectories separately	Protects longitudinal meaning	Temporal validation [6]	Implausible initiation, discontinuation, or follow-up	Temporal fit is not causal validity
Mixed And Relational Structure	Variables may fit separately but fail jointly	Data types, dependencies, graph links	Test conditional and relational fidelity	Detects hidden incoherence	Mixed-type and graph evaluation [7, 8]	Broken dose–laboratory or diagnosis–therapy relations	Structural fit is not medication safety
Risk Category	"Synthetic" is too broad for governance	Consequence, access, linkage, automation, uncertainty	Proposed multidimensional classification	Supports differentiated controls	Empirical calibration across uses	Unsupported rankings or universal thresholds	Categories organize review only

Professional Boundary	Technical artifacts may influence medication decisions	Interface, user role, workflow, escalation	Keep pharmacist or clinician authority outside generator output	Preserves accountable judgment	Workflow and safety evaluation [4]	Automation bias or unreviewed action	Dataset release is not clinical authorization

Note: Relationships and classifications labelled "proposed" are original synthesis and require empirical validation.

Proposed Synthetic-Data Governance Architecture

Healthcare AI governance requires coordinated oversight across development, validation, implementation, and monitoring rather than a single technical checkpoint [9]. Applied to synthetic pharmacy records, governance should begin before generation. The initiating party must define the intended users, task, setting, decision consequence, access conditions, prohibited uses, and evidence needed to justify release. This intended-use contract becomes the reference point for all later tests.

Responsible health-machine-learning practice also requires explicit problem definition, appropriate data, rigorous evaluation, transparency, and continuing surveillance [10]. The proposed architecture therefore contains seven linked modules: intended use and consequence; source-data envelope and provenance; generator configuration and versioning; a four-domain control bundle; independent use-conditioned evaluation; bounded release; and derivative monitoring, expiry, withdrawal, and supersession. Failed controls return the artifact to scope, source, or generation review rather than being averaged into a favorable composite score.

Organizational accountability requires named roles, traceable decisions, auditing, and enforceable controls [11]. The architecture assigns responsibility to a data steward, generation team, independent evaluator, release authority, and downstream user. These roles may be combined in small organizations, but conflicts and limitations must be disclosed. Every release decision should identify who authorized the use, which evidence was reviewed, what remained uncertain, and what event would trigger suspension or supersession.

Clinical decision-support governance adds separate obligations for validation, transparency, safety monitoring, stakeholder roles, and user training [12]. Accordingly, the architecture distinguishes dataset release from model release and clinical-use authorization. A synthetic dataset may be suitable for a sandbox or methods study while remaining unsuitable for medication alerts, prescribing support, dispensing prioritization, or patient communication.

The proposed architecture integrates lifecycle oversight, responsible-development controls, and auditable organizational accountability [9-11]. Its central visual anchor is an uncertainty ledger that travels with the artifact. This ledger records tested properties, untested uses, residual privacy risk, subgroup limitations, known failure modes, and the current release state.

Figure 1 presents the original conceptual synthesis for synthetic pharmacy-record governance architecture, showing how the article's principal components, evidence relationships, uncertainties, and decision boundaries are connected.

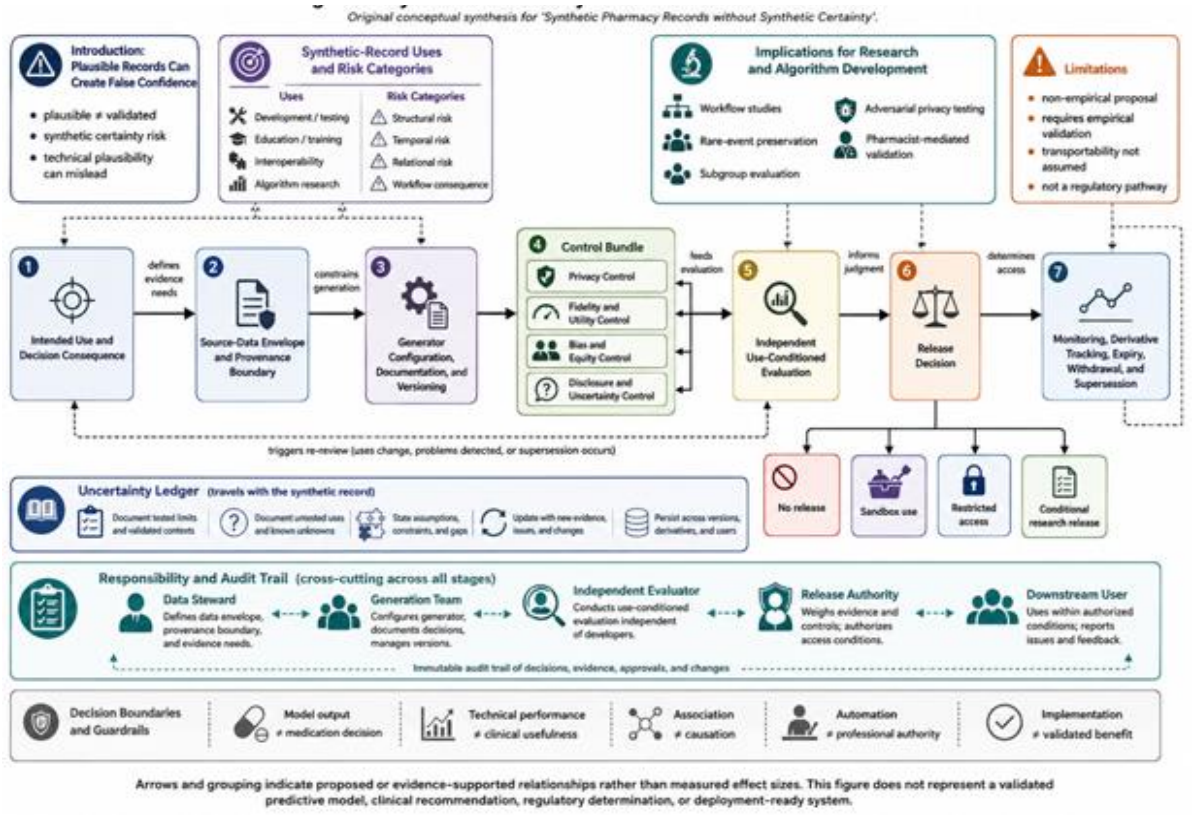

Figure 1. Synthetic Pharmacy-Record Governance Architecture

Table 2 organizes the evidence, constructs, relationships, uncertainties, and boundary conditions required for evaluation criteria, evidence requirements, and failure modes for the article Synthetic Pharmacy Records without Synthetic Certainty.

Table 2. Evaluation criteria, evidence requirements, and failure modes for the article Synthetic Pharmacy Records without Synthetic Certainty.

Architecture Layer or Process Stage	Core Function	Information Flow	Interaction With Other Components	Evaluation Criterion	Potential Failure Mode	Human or Organizational Responsibility	Readiness Boundary
Intended-Use Specification	Define permitted task and consequence	User need → use contract	Controls every later gate	Scope is explicit and testable	Vague or expanding purpose	Sponsor and release authority	No evaluation or release without scope
Source-Data Envelope	Bound represented population and measurements	Source metadata → provenance record	Constrains generator and transportability	Population, time, variables, exclusions documented	Hidden coverage gaps	Data steward	Source availability does not establish representativeness
Generation And Versioning	Produce reproducible artifact	Configuration → versioned output	Feeds all control domains	Model, preprocessing, dependencies, version recorded	Silent model or data change	Generation team	Reproducibility is not validity
Control Bundle	Separate privacy, fidelity, bias, and disclosure	Test results → uncertainty ledger	Parallel evidence; no automatic compensation	Use-matched tests with residual uncertainty	Favorable metric masks another failure	Independent evaluator	No single metric establishes readiness [2, 3]
Independent Evaluation	Challenge claims for declared use	Artifact + contract → evaluation report	Informs bounded release	Methods, comparators, limits, conflicts explicit	Developer-only or insensitive testing	Independent evaluator	Evaluation is conditional on task
Release Decision	Assign no release, sandbox, restricted, or conditional research use	Evidence report → signed decision	Sets access, expiry, and prohibited uses	Conditions and rationale are traceable	Release interpreted as certification	Release authority	Dataset release is not clinical authorization [9, 11, 12]

Monitoring And Supersession	Detect misuse, drift, incidents, and obsolete versions	Use evidence → review or withdrawal	Returns changed uses to earlier gates	Derivatives, incidents, expiry, replacement tracked	Orphaned copies or silent reuse	Steward and downstream user	Continued access requires current conditions

Note: The layer arrangement, release states, feedback rules, and readiness boundaries are proposed original synthesis.

Generation, Documentation, Provenance, and Intended Use

Generation should be treated as a documented scientific event rather than a one-time export. Work on heterogeneous electronic health records shows that sparsity, mixed variables, temporal structure, model design, and the selected privacy and fidelity tests materially shape the generated artifact [13]. A minimum generation dossier should therefore state the source-data envelope, preprocessing, inclusion and exclusion logic, missingness handling, conditioning variables, generator family, training and evaluation partitions, and known unsupported record types.

Generator class and configuration must be versioned because temporal representation, comparison methods, and privacy testing influence record behavior [14]. Each release should have a unique identifier linked to code and dependency versions, configuration, model checkpoint, date, evaluator, and supersession status. Re-running a pipeline after a source-data, preprocessing, or model change should create a new governed object rather than silently replacing the prior release. Versioning improves traceability but does not guarantee reproducibility across computing environments or validity across settings.

Dataset documentation frameworks support recording motivation, composition, creation, recommended uses, limitations, and maintenance [15]. For synthetic pharmacy records, the proposed synthetic-record dossier extends those fields to include medication vocabularies, dose and unit handling, temporal granularity, rare-event treatment, laboratory linkages, adherence assumptions, excluded populations, intended users, prohibited uses, access tier, expiry, and contact points for incident reporting. Documentation must distinguish what was measured from what was assumed and what remains untested.

Responsible use of health data in AI further depends on metadata, transparency labels, stakeholder accountability, bias controls, and continuing updates [16]. The proposed provenance chain links the source cohort, transformations, generator version, evaluator, release authority, access decision, downstream derivative, and any withdrawal or replacement. A corresponding disclosure label should travel with every copy and derivative. It should state that the records are synthetic, identify the permitted use, summarize tested and untested properties, disclose residual privacy and representation risks, and prohibit unsupported linkage or clinical interpretation.

Intended use is therefore not a narrative added after generation. It is a governing constraint that determines which variables must be preserved, which failures matter, which evaluators are needed, and which release conditions are defensible. A change from software testing to algorithm training, or from research to medication decision support, creates a new use claim and requires re-entry into evaluation. The dossier and provenance chain make such changes visible; they do not establish that the new use is valid, safe, fair, lawful, or operationally ready.

Privacy, Fidelity, Bias, and Disclosure Controls

Privacy protection cannot be inferred merely because records are labelled synthetic. Fully synthetic health data may retain identity-disclosure risk when a generator memorizes unusual source patterns or when generated attributes can be linked to external information about real individuals [17]. Evaluation should therefore begin with an explicit threat model describing potential adversaries, available background knowledge, linkage opportunities, access conditions, and the forms of identity or attribute disclosure being tested. Passing one attack test does not establish zero privacy risk.

Fidelity and utility are also distinct. Existing privacy and utility metrics vary substantially in definition, implementation, and interpretation, and no single metric is suitable for every medical synthetic-data use [18]. The proposed control bundle consequently separates marginal fidelity, conditional fidelity, temporal fidelity, relational fidelity, and downstream task utility. Each test should be justified by the intended use rather than selected because it produces a favourable result. Aggregating incompatible measurements into one quality score could obscure clinically important failure.

Claims of anonymity require particular restraint. Regulatory analysis of tabular synthetic health data indicates that privacy claims should be tied to empirical tests, threat assumptions, residual risk, and the context in which records will be accessed and reused [19]. The proposed disclosure label should therefore identify the protection mechanism, attacks performed, assumptions about external information, unresolved exposure, access restrictions, prohibited linkage, and consequences of derivative release. Such disclosure supports governance but does not itself determine legal status or regulatory acceptability.

Bias control must extend beyond matching overall distributions. Healthcare algorithms can reproduce structural inequity when proxy targets encode unequal access, expenditure, documentation, or service use rather than the clinical need of interest [20]. A synthetic generator may preserve, amplify, suppress, or reorganize those patterns. Evaluation should examine subgroup representation,

intersectional sparsity, missingness, proxy labels, rare medication pathways, and differential downstream errors. Similar aggregate performance does not establish equitable utility.

A privacy claim is incomplete unless it reports disclosure-risk testing, utility metrics, threat assumptions, residual risk, and access conditions [17-19]. Privacy, fidelity, bias, and disclosure should therefore remain separate but interacting controls: success in one domain cannot automatically compensate for failure in another. The proposed bundle organizes evidence and uncertainty; it does not prove anonymity, fairness, medication safety, or clinical usefulness.

Evaluation and Release Decisions

Evaluation should be multidimensional and conditioned on the intended use. Medical-AI data-quality research supports assessing relevance, representativeness, completeness, correctness, consistency, and other context-dependent dimensions rather than treating quality as an intrinsic property of a dataset [21]. Where feasible, evaluation should be conducted by reviewers who are sufficiently independent from generator development to challenge assumptions, identify conflicts, and test failure conditions.

High-consequence uses require more than dataset-level metrics. Early clinical evaluation of AI decision support should examine human interaction, workflow integration, errors, safety, usability, and local context alongside technical performance [22]. Accordingly, synthetic pharmacy records intended for medication-related decision-support development require pharmacist-informed scenario testing, error-recovery analysis, workload evaluation, and explicit separation between experimental model performance and professional authority.

Synthetic-EHR methods differ in fidelity, privacy exposure, conditional generation, and downstream task performance, making transparent benchmarking necessary [23]. Appropriate comparisons may include real-only, synthetic-only, and mixed-data development; multiple generator classes; temporally or institutionally separated evaluations; calibration; subgroup performance; and sensitivity to rare medication events. No single benchmark can establish general fitness.

The proposed release decision integrates context-specific data quality [21], sociotechnical evaluation [22], and transparent synthetic-data benchmarking [23]. Four non-numeric states are proposed: no release, sandbox use, restricted access, and conditional research release. Each decision should identify the permitted task, users, access controls, expiry date, monitoring obligations, prohibited uses, and evidence that would trigger withdrawal or re-evaluation. These states require empirical and institutional calibration. They are not certification levels or regulatory determinations.

Implications for Research and Algorithm Development

Research should examine how synthetic records affect professional interpretation, trust, role boundaries, local adaptation, and error recovery. Qualitative evidence from AI decision-support implementation shows that consequences emerge through workflow and organizational conditions that technical testing alone cannot resolve [24]. Pharmacy-specific studies should therefore investigate how pharmacists identify synthetic anomalies, communicate uncertainty, challenge model outputs, and respond when generated records conflict with medication knowledge.

Algorithm developers should distinguish development utility from clinical impact. External validation, usability, workflow integration, pathway change, regulation, and monitoring influence whether technically successful AI produces clinical benefit [25]. Synthetic records may be useful for prototyping or stress testing while remaining insufficient for estimating patient outcomes, causal effects, safety, or performance in a live pharmacy environment.

Medication-alert research illustrates this boundary. AI approaches to alert optimization show heterogeneous performance, limited formal validation, and sparse real-world implementation evidence [26]. Synthetic pharmacy records could support controlled evaluation of alert logic, override scenarios, rare combinations, and interface behaviour, but they should not be used alone to claim reduced alert burden or improved medication safety.

Future validation should combine workflow inquiry [24], translational evaluation [25], and medication-specific decision-support testing [26]. Priorities include rare-event preservation, counterfactual misuse testing, subgroup utility, adversarial privacy evaluation, pharmacist-mediated simulation, cross-institutional transportability, derivative tracking, and prospective assessment of the proposed release gates.

Limitations

The architecture may not anticipate every clinical, cognitive, organizational, or epistemic consequence of machine-learning use [27]. Its categories and responsibility assignments are therefore non-exhaustive.

Apparent performance may also be distorted by biased samples, weak validation, inadequate reporting, or inappropriate analytic choices [28]. The proposed gates cannot correct poor underlying evidence automatically.

Finally, transportability across populations, institutions, pharmacy systems, jurisdictions, and time cannot be assumed [29]. The architecture and its release states require local adaptation and empirical validation. This article does not validate a generator, synthetic dataset, clinical workflow,

governance committee, privacy method, or regulatory pathway.

Conclusion

Synthetic pharmacy records can support valuable forms of development, education, testing, and research, but their plausibility can generate unwarranted certainty. This article proposes a governance architecture that treats each synthetic record set as a purpose-bounded, versioned, auditable, and reversible artifact rather than as a substitute for observed clinical evidence.

The architecture connects intended use and consequence to the source-data envelope, generator configuration, provenance, separate privacy–fidelity–bias–disclosure controls, independent evaluation, bounded release, and continuing monitoring and supersession. It preserves boundaries between generated output and medication decisions, technical performance and clinical usefulness, association and causation, automation and professional authority, and implementation and validated benefit.

Its scientific contribution is an integrated set of proposed relationships, responsibilities, evidence gates, release states, and uncertainty disclosures that can be tested in future pharmacy research. Its professional and organizational contribution is to make responsibility and permitted use traceable. Its safety and equity contribution is to prevent aggregate realism or performance from concealing rare medication failures, subgroup limitations, privacy exposure, or workflow consequences.

The architecture is not validated, clinically prescriptive, regulator-approved, universally applicable, or deployment-ready. Its value depends on future empirical evaluation, transparent adaptation, and continued willingness to withhold, restrict, withdraw, or supersede synthetic records when their evidence no longer supports their intended use.

ACKNOWLEDGMENTS: None
CONFLICT OF INTEREST: None
FINANCIAL SUPPORT: None
ETHICS STATEMENT: None

References

1. Giuffrè M, Shung DL. Harnessing the power of synthetic data in healthcare: Innovation, application, and privacy. NPJ Digit Med. 2023;6(1):186. doi:10.1038/s41746-023-00927-3
2. Goncalves A, Ray P, Soper B, Stevens J, Coyle L, Sales AP. Generation and evaluation of synthetic patient data. BMC Med Res Methodol. 2020;20(1):108. doi:10.1186/s12874-020-00977-1
3. Tucker A, Wang Z, Rotalinti Y, Myles P. Generating high-fidelity synthetic patient data for assessing machine learning healthcare software. NPJ Digit Med. 2020;3(1):147. doi:10.1038/s41746-020-00353-9
4. Gates PJ, Hardie RA, Raban MZ, Li L, Westbrook JI. How effective are electronic medication systems in reducing medication error rates and associated harm among hospital inpatients? A systematic review and meta-analysis. J Am Med Inform Assoc. 2021;28(1):167-76. doi:10.1093/jamia/ocaa230
5. Walonoski J, Kramer M, Nichols J, Quina A, Moesel C, Hall D, et al. Synthea: An approach, method, and software mechanism for generating synthetic patients and the synthetic electronic health care record. J Am Med Inform Assoc. 2018;25(3):230-8. doi:10.1093/jamia/ocx079
6. Zhang Z, Yan C, Lasko TA, Sun J, Malin BA. SynTEG: A framework for temporal structured electronic health data simulation. J Am Med Inform Assoc. 2021;28(3):596-604. doi:10.1093/jamia/ocaa262
7. Li J, Cairns BJ, Li J, Zhu T. Generating synthetic mixed-type longitudinal electronic health records for artificial intelligent applications. NPJ Digit Med. 2023;6(1):98. doi:10.1038/s41746-023-00834-7
8. Nikolentzos G, Vazirgiannis M, Xypolopoulos C, Lingman M, Brandt EG. Synthetic electronic health records generated with variational graph autoencoders. NPJ Digit Med. 2023;6(1):83. doi:10.1038/s41746-023-00822-x
9. Reddy S, Allan S, Coghlan S, Cooper P. A governance model for the application of AI in health care. J Am Med Inform Assoc. 2020;27(3):491-7. doi:10.1093/jamia/ocz192
10. Wiens J, Saria S, Sendak M, Ghassemi M, Liu VX, Doshi-Velez F, et al. Do no harm: A roadmap for responsible machine learning for health care. Nat Med. 2019;25(9):1337-40. doi:10.1038/s41591-019-0548-6
11. Roski J, Maier EJ, Vigilante K, Kane EA, Matheny ME. Enhancing trust in AI through industry self-governance. J Am Med Inform Assoc. 2021;28(7):1582-90. doi:10.1093/jamia/ocab065
12. Labkoff S, Oladimeji B, Kannry J, Solomonides A, Leftwich R, Koski E, et al. Toward a responsible future: Recommendations for AI-enabled clinical decision support. J Am Med Inform Assoc. 2024;31(11):2730-9. doi:10.1093/jamia/ocae209
13. Yoon J, Mizrahi M, Ghalaty NF, Jarvinen T, Ravi AS, Brune P, et al. EHR-Safe: Generating high-fidelity and privacy-preserving synthetic electronic health records. NPJ Digit Med. 2023;6(1):141. doi:10.1038/s41746-023-00888-7
14. Tian M, Chen B, Guo A, Jiang S, Zhang AR. Reliable generation of privacy-preserving synthetic electronic health record time series via diffusion models. J Am Med Inform Assoc. 2024;31(11):2529-39. doi:10.1093/jamia/ocae229
15. Gebru T, Morgenstern J, Vecchione B, Vaughan JW, Wallach H, Daumé H 3rd, et al. Datasheets for datasets. Commun ACM. 2021;64(12):86-92. doi:10.1145/3458723
16. Koski E, Das A, Hsueh PYS, Solomonides A, Joseph AL, Srivastava G, et al. Towards responsible artificial intelligence in healthcare—getting real about real-world data and evidence. J Am Med Inform Assoc. 2025;32(11):1746-55. doi:10.1093/jamia/ocaf133
17. El Emam K, Mosquera L, Bass J. Evaluating identity disclosure risk in fully synthetic health data: Model development and validation. J Med Internet Res. 2020;22(11). doi:10.2196/23139
18. Kaabachi B, Despraz J, Meurers T, Otte K, Halilovic M, Kulynych B, et al. A scoping review of privacy and utility metrics in medical synthetic data. NPJ Digit Med. 2025;8(1):60. doi:10.1038/s41746-024-01359-3
19. Pilgram L, Ko H, Tung A, El Emam K. Protecting patient privacy in tabular synthetic health data: A regulatory perspective. NPJ Digit Med. 2025;8(1):732. doi:10.1038/s41746-025-02112-0
20. Obermeyer Z, Powers B, Vogeli C, Mullainathan S. Dissecting racial bias in an algorithm used to manage the health of populations. Science. 2019;366(6464):447-53. doi:10.1126/science.aax2342
21. Schwabe D, Becker K, Seyferth M, Klaß A, Schaeffter T. The METRIC-framework for assessing data quality for trustworthy AI in medicine: A systematic review. NPJ Digit Med. 2024;7(1):203. doi:10.1038/s41746-024-01196-4
22. Vasey B, Nagendran M, Campbell B, Clifton DA, Collins GS, Denaxas S, et al. Reporting guideline for the early-stage clinical evaluation of decision support systems driven by artificial intelligence: DECIDE-AI. Nat Med. 2022;28(5):924-33. doi:10.1038/s41591-022-01772-9
23. Chen X, Wu Z, Shi X, Cho H, Mukherjee B. Generating synthetic electronic health record data: A methodological scoping review with benchmarking on phenotype data and open-source software. J Am Med Inform Assoc. 2025;32(7):1227-40. doi:10.1093/jamia/ocaf082
24. Farič N, Hinder S, Williams R, Ramaesh R, Bernabeu MO, van Beek E, et al. Early experiences of integrating an artificial intelligence-based diagnostic decision support system into radiology settings: A

qualitative study. J Am Med Inform Assoc. 2024;31(1):24-34. doi:10.1093/jamia/ocad191
25. Kelly CJ, Karthikesalingam A, Suleyman M, Corrado G, King D. Key challenges for delivering clinical impact with artificial intelligence. BMC Med. 2019;17(1):195. doi:10.1186/s12916-019-1426-2
26. Graafsma J, Murphy RM, van de Garde EMW, Karapinar-Çarkit F, Derijks HJ, Hoge RHL, et al. The use of artificial intelligence to optimize medication alerts generated by clinical decision support systems: A scoping review. J Am Med Inform Assoc. 2024;31(6):1411-22. doi:10.1093/jamia/ocae076
27. Cabitza F, Rasoini R, Gensini GF. Unintended consequences of machine learning in medicine. JAMA. 2017;318(6):517-8. doi:10.1001/jama.2017.7797
28. Wynants L, Van Calster B, Collins GS, Riley RD, Heinze G, Schuit E, et al. Prediction models for diagnosis and prognosis of COVID-19 infection: Systematic review and critical appraisal. BMJ. 2020;369. doi:10.1136/bmj.m1328
29. Futoma J, Simons M, Panch T, Doshi-Velez F, Celi LA. The myth of generalisability in clinical research and machine learning in health care. Lancet Digit Health. 2020;2(9). doi:10.1016/S2589-7500(20)30186-2