ORIGINAL RESEARCH **Open access**

Consent-Aware Genomic–Clinical Analytics: A Policy-Constrained Access Control Framework for Secondary Data Use

Victor Santos[1*], Rafael Costa[1], Bruno Teixeira[2]

Abstract

The rapid integration of artificial intelligence (AI) into healthcare analytics has amplified the need for robust frameworks that govern secondary use of genomic and clinical data while prioritizing patient consent and policy compliance. This conceptual manuscript introduces a novel policy-constrained access control framework designed to facilitate consent-aware analytics in genomic-clinical environments. By embedding dynamic consent mechanisms into data access pipelines, the framework ensures that secondary data utilization adheres to ethical, legal, and institutional policies, mitigating risks associated with unauthorized reuse. We synthesize recent literature on data sharing, privacy protections, and genomic informatics to underscore the framework's theoretical foundations. Key components include layered access orchestration, policy-enforced query resolution, and feedback loops for consent revocation monitoring. Conceptual formulas are presented to interpret risk propagation in access chains and governance load under varying policy constraints. The architecture promotes interoperability between genomic repositories and clinical systems, fostering trustworthy AI-driven insights without empirical validation. Implications for healthcare stakeholders emphasize enhanced data stewardship, reduced privacy breaches, and scalable secondary analytics. This work advances conceptual discourse on AI-enabled healthcare systems by proposing a governance-centric infrastructure that balances innovation with patient autonomy in secondary data contexts.

Keywords Consent-aware analytics, Genomic-clinical integration, Policy-constrained access, Secondary data governance, AI healthcare frameworks, Data privacy orchestration

*Correspondence:
Victor Santos
victor.santos@gmail.com
[1] Department of Digital Health Systems, Faculty of Medicine, University of Sao Paulo, Sao Paulo, Brazil
[2] Department of Health Data Engineering, Faculty of Engineering, University of Campinas, Campinas, Brazil

Introduction

The convergence of artificial intelligence with genomic and clinical data analytics represents a transformative shift in healthcare. Yet, it introduces profound challenges in managing secondary data use under consent and policy constraints. As healthcare systems increasingly rely on repurposed datasets for AI-driven discoveries, ensuring that access aligns with original consent intentions becomes paramount. This manuscript conceptualizes a framework that embeds consent-awareness directly into analytic workflows, addressing gaps in current practices where secondary uses often outpace governance capabilities.

Genomic-clinical data modalities in consent-limited environments

Genomic data, encompassing sequences, variants, and epigenetic markers, intersect with clinical records such as electronic health data and imaging to enable precision medicine. However, secondary uses—ranging from research aggregation to AI model refinement—must

navigate consent boundaries that vary by jurisdiction and patient preferences [1-7]. In clinical settings like oncology or rare disease management, where genomic-clinical fusion drives prognostic analytics, policy constraints often restrict data flows to prevent re-identification risks [5-10]. Literature highlights how commercial datasets amplify these issues, as they may inadvertently bypass consent protocols during aggregation [5]. This subheading explores how data modalities influence access control, emphasizing the need for frameworks that dynamically interpret consent metadata embedded in genomic-clinical streams.

Policy-constrained deployment in multi-institutional healthcare settings

Deployment of AI analytics in hospitals or research consortia demands policy-constrained mechanisms to govern secondary data access. Policies from bodies like the GDPR or HIPAA impose granular controls, yet integration with genomic repositories remains fragmented [2, 3]. For instance, structured electronic health records used in secondary research require best-practice frameworks that incorporate consent revocation pathways [3]. In environments involving international collaborations, such as COVID-19 data registries, policy alignment ensures ethical sharing without compromising clinical utility [4]. This section delineates how deployment environments shape access frameworks, advocating for infrastructures that orchestrate policy checks at each analytic stage to safeguard secondary uses.

Governance constraints on secondary analytic workflows

Governance in genomic-clinical analytics extends beyond compliance to encompass ethical stewardship of secondary data. Challenges arise when consent forms, often lengthy and complex, fail to anticipate future AI applications [11]. Studies on patient perspectives reveal preferences for transparent data sharing decisions, particularly in biospecimen research [7, 12-16]. Moreover, tribal or community-based governance models underscore the importance of culturally sensitive consent in secondary contexts [17-24]. By anchoring governance to constraints like auditability and revocability, analytic workflows can mitigate biases introduced by uneven data access [6]. This subheading examines how these constraints necessitate policy-constrained frameworks to maintain trust in AI healthcare systems.

Clinical setting implications for consent-aware integration

In acute clinical settings, such as intensive care units where genomic-clinical data informs real-time decisions, secondary uses for AI training must respect consent scopes [9]. The ethical acceptability of pragmatic trials highlights the role of post-consent mechanisms in data repurposing [15]. Furthermore, informatics tools like HL7 FHIR genomics operations facilitate integration but require policy layers to handle consent-aware queries [22]. This integration is crucial in settings with high-stakes analytics, where secondary data fuels predictive models without direct patient involvement.

Data modality challenges in policy-enforced secondary reuse

Diverse data modalities— from raw genomic sequences to derived clinical phenotypes—pose unique challenges in policy enforcement for secondary analytics. Privacy concerns in emerging technologies, such as wearable-derived health data, parallel those in genomics, where disposition toward disclosure varies [23]. Scoping reviews of distributed ledger technologies suggest blockchain-inspired approaches for consent tracking in genomic sharing [20]. This subheading addresses how modality-specific policies constrain access, proposing conceptual bridges to unified frameworks.

Theoretical Background and Literature Synthesis

The theoretical underpinnings of consent-aware genomic-clinical analytics draw from interdisciplinary domains, including informatics, ethics, and policy studies. This section synthesizes peer-reviewed insights from 2017 to 2024, focusing on how policy-constrained access controls can theoretically enhance secondary data use in AI healthcare systems. By integrating concepts from data sharing platforms, privacy frameworks, and genomic governance, we lay the groundwork for a novel architectural approach.

Consent dynamics in genomic-clinical data sharing platforms

Consent mechanisms in genomic-clinical platforms have evolved to address secondary use complexities. Patient perspectives on sharing medical data emphasize the need for informed, revocable consent, particularly in research involving biospecimens [7, 17]. Chat-based tools for dynamic consent demonstrate potential for large-scale genomic studies, allowing real-time adjustments to secondary access permissions [17]. In global health digitalization, secondary data flows require transparency to prevent exploitation, as seen in frameworks for COVID-19 registries [1, 4]. This subheading synthesizes how consent dynamics influence platform design, highlighting theoretical models that embed patient agency into data-sharing ecosystems.

Policy frameworks for privacy in secondary genomic analytics

Privacy protections form a core theoretical pillar for policy-constrained analytics. Evolving concerns in genomic data analysis underscore the role of competitions like iDASH in advancing secure sharing methods [16]. Organizational factors in clinical data sharing for AI reveal barriers like institutional silos, necessitating policy-driven interoperability [13]. Moreover, assessments of informed consent documents in trials illustrate readability challenges that impact secondary use validity [11]. Literature on commercial health datasets warns of algorithmic biases stemming from opaque privacy practices [5]. This section integrates these insights to theorize policy frameworks that mitigate risks in secondary genomic reuse.

Governance infrastructures for AI-enabled clinical environments

Governance in AI-enabled healthcare environments requires infrastructures capable of simultaneously enabling innovation and enforcing ethical, regulatory, and procedural constraints. As artificial intelligence systems increasingly rely on large-scale clinical and genomic datasets, governance structures must move beyond traditional institutional oversight toward infrastructural mechanisms embedded within data architectures themselves. These governance infrastructures function as mediating layers between data generation, secondary analysis, and algorithmic deployment, ensuring that technological advancement does not outpace ethical safeguards.

The CODE-EHR framework offers one such infrastructural approach by providing best practices for structured electronic health record data used in research settings. By emphasizing data quality, provenance, and policy alignment, the framework supports responsible secondary analyses while maintaining transparency about how clinical records are curated and repurposed [3]. Governance infrastructures informed by CODE-EHR principles facilitate reproducibility and accountability, which are particularly critical in AI systems that rely on high-dimensional datasets for model training and validation.

Complementing these governance models, recent recommendations for transparent health dataset documentation have highlighted the importance of detailed metadata and contextual information in mitigating bias in machine learning systems. In the context of genomic-clinical data integration, documentation practices serve as governance tools by clarifying dataset origins, inclusion criteria, and potential limitations that could influence algorithmic outcomes [6]. Transparent documentation, therefore, acts not only as a research practice but also as a policy instrument that supports ethical oversight in AI analytics.

In registry science, atomic approaches to clinical data warehouses further illustrate how governance infrastructures can be embedded directly within technical architectures. Atomic data models preserve granular information while enabling flexible aggregation, thereby supporting governed access pathways for secondary uses of clinical data [19, 21]. Such architectures allow data access policies to be applied dynamically across analytic contexts, reducing the need for ad hoc governance decisions and enabling scalable research infrastructures.

Emerging technologies also offer theoretical extensions to these governance frameworks. Distributed ledger technologies have been explored as mechanisms for recording data provenance, managing consent, and enabling verifiable access logs within genomic research ecosystems [20]. By providing immutable audit trails and programmable consent structures, these technologies could support consent-aware environments in which data governance is operationalized directly within the data infrastructure. This subheading, therefore, conceptualizes governance infrastructures as socio-technical systems that integrate regulatory frameworks, technical standards, and ethical oversight to support responsible AI-driven clinical research.

Ethical dimensions of access control in secondary data modalities

Access control mechanisms constitute a central ethical challenge in AI-driven clinical research, particularly in contexts involving secondary uses of sensitive health data. While broad data availability can accelerate scientific discovery, unrestricted access risks compromising patient privacy, undermining trust, and exacerbating inequities in data governance. Ethical access control, therefore, requires balancing data utility with privacy protections through carefully designed governance models.

The complexity of these considerations is evident in analyses of data sharing practices following the International Committee of Medical Journal Editors (ICMJE) requirements for data transparency. Scoping reviews indicate variable compliance with these policies, revealing persistent gaps between formal data-sharing mandates and the actual availability of datasets for secondary analysis [10]. These discrepancies raise important ethical questions regarding accountability, transparency, and the integrity of secondary research ecosystems that depend on accessible data.

Community-based participatory research provides further insight into how ethical access controls must account for culturally specific understandings of privacy and data sovereignty. The Strong Heart Study, for example, illustrates how tribal perspectives on genomic data governance emphasize collective rights, cultural stewardship, and long-term community oversight of research activities [24]. These perspectives challenge conventional models of individual consent and highlight the importance of governance frameworks that recognize communal forms of data ownership.

Global data-sharing initiatives during the COVID-19 pandemic further demonstrate how ethical access control can be operationalized within coordinated research infrastructures. Platforms guided by the FAIR principles—ensuring that data are findable, accessible, interoperable, and reusable—have shown how collaborative data ecosystems can support rapid scientific progress while maintaining respect for consent and ethical oversight [4]. Importantly, the FAIR framework emphasizes structured governance mechanisms that define how data access is granted, monitored, and evaluated.

Taken together, these perspectives suggest that access control cannot be treated solely as a technical function but must be understood as an ethical governance mechanism embedded within data infrastructures. Modality-specific ethical considerations—such as those relevant to genomic data, clinical records, or community-based datasets—should inform the design of access control policies. This synthesis, therefore, theorizes how consent-aware governance structures can incorporate flexible access modalities that support AI analytics while preserving the ethical principles underlying clinical research.

Integration challenges in policy-constrained clinical deployments

The integration of genomic and clinical data within AI-enabled healthcare systems presents substantial governance challenges, particularly when policy constraints intersect with complex technical infrastructures. Clinical deployments must reconcile interoperability standards, institutional regulations, and patient consent frameworks while ensuring that integrated datasets remain reliable for research and algorithmic development.

Standards such as HL7 FHIR have been proposed as mechanisms for bridging genomic data with electronic health record systems. FHIR-based operations enable developer-friendly integration by standardizing data exchange formats and providing modular interfaces for genomic information within clinical environments [22]. However, while these standards facilitate technical interoperability, they do not inherently address governance concerns surrounding secondary data use. Policy layers, therefore, remain necessary to regulate how integrated datasets are accessed, shared, and repurposed for AI analytics.

National mapping initiatives of health data flows further illustrate the importance of transparency within complex data infrastructures. These mappings document how clinical data move across institutional boundaries, undergo transformations within data warehouses, and become accessible to researchers or algorithm developers [2]. By clarifying these infrastructural pathways, governance frameworks can better ensure that consent conditions and regulatory requirements are maintained throughout the data lifecycle.

Policy constraints also influence data availability and reproducibility within research ecosystems. Studies

examining surgical journals have shown that journal data-sharing policies can significantly affect whether datasets become accessible for secondary analyses, thereby shaping the reproducibility of published findings [14]. Similar dynamics apply to genomic-clinical analytics, where limited data accessibility can hinder validation of AI models and slow scientific progress.

Good practice guidelines for clinical data warehouses further emphasize the importance of governance frameworks that integrate policy requirements with technical infrastructures. These guidelines advocate for structured governance committees, standardized access protocols, and transparent documentation of data transformations within clinical research environments [25]. Within AI-enabled healthcare systems, such practices help ensure that integrated datasets remain both technically usable and ethically governed.

This subheading, therefore, conceptualizes integration challenges not simply as technical interoperability issues but as governance problems requiring coordinated policy, infrastructure, and oversight mechanisms. Addressing these challenges is essential for enabling policy-constrained yet scientifically productive clinical deployments of AI technologies.

Bias mitigation and equity in consent-aware analytic governance

Bias mitigation and equity represent critical concerns in the governance of AI-driven clinical analytics, particularly when algorithms are trained on datasets that may reflect historical inequities in healthcare systems. Governance models must therefore incorporate mechanisms that identify, monitor, and mitigate bias while ensuring equitable access to the benefits of data-driven medicine.

Conceptual frameworks for fairness in medical algorithms provide important theoretical foundations for these governance efforts. By defining key pillars such as transparency, accountability, and representativeness, these frameworks guide the evaluation of algorithmic systems in clinical contexts and highlight the need for continuous oversight throughout the analytic lifecycle [26]. When applied to genomic-clinical analytics, such fairness principles help ensure that AI tools do not disproportionately disadvantage underrepresented populations.

Data sharing policies also play a critical role in shaping equitable research ecosystems. Initiatives designed to fulfill NIH data-sharing requirements emphasize the importance of avoiding “open data in appearance only,” in which datasets are nominally available but practically inaccessible due to restrictive governance or technical barriers [27]. Genuine data accessibility supports diverse secondary research activities, thereby promoting more inclusive scientific participation and improving the robustness of AI models.

Ethical trade-offs between data utility and consent are particularly evident in genomic screening initiatives involving critically ill infants. Measures of clinical utility in these contexts highlight the potential benefits of rapid genomic diagnostics while simultaneously raising questions about consent processes, data reuse, and long-term governance of sensitive genomic information [8, 9]. These tensions underscore the importance of governance frameworks that can accommodate urgent clinical needs without compromising ethical oversight.

Data curation frameworks further contribute to equity by emphasizing structured evaluation of dataset quality, representativeness, and usability. By systematically assessing how datasets are curated and prepared for analysis, governance infrastructures can identify potential sources of bias and ensure that data used for AI training reflect diverse patient populations [28]. Such frameworks also support transparency in the selection and preparation of datasets for secondary analyses.

In this context, consent-aware analytic governance emerges as a critical paradigm for balancing innovation with ethical responsibility. Governance systems that incorporate dynamic consent mechanisms, transparent access policies, and bias monitoring protocols can foster equitable secondary analytics while maintaining public trust in data-driven healthcare research. By synthesizing existing literature on bias mitigation, data sharing, and ethical governance, this section advocates for governance infrastructures that actively promote fairness and inclusivity within AI-enabled clinical environments.

Policy-orchestrated genomic-clinical access infrastructure

This section delineates the core architecture of our proposed consent-governed secondary analytics lattice (CGSAL), a uniquely layered framework with a reticular

feedback topology designed to orchestrate policy-constrained access in genomic-clinical environments. Unlike linear models, CGSAL employs a lattice structure where nodes represent consent-policy intersections, enabling multidimensional data flows for secondary use. The infrastructure integrates four primary layers: consent ingestion layer, policy resolution layer, analytic orchestration layer, and revocation feedback layer. Each layer incorporates interpretive governance to ensure AI-driven analytics respect secondary data boundaries.

The consent ingestion layer captures dynamic consent metadata from genomic-clinical sources, transforming patient preferences into queryable tokens. This layer interfaces with clinical systems to embed consent scopes, such as time-bound or purpose-specific permissions, into data streams.

The policy resolution layer evaluates access requests against institutional, legal, and ethical policies using a constraint satisfaction paradigm. Here, policies are modeled as predicates that filter secondary queries, preventing unauthorized analytic derivations.

The analytic orchestration layer facilitates AI integration by routing consented data to analytic modules, such as variant prioritization or phenotype prediction, under policy envelopes. This layer ensures that secondary uses remain theoretical, focusing on architectural scalability rather than empirical outputs.

Finally, the revocation feedback layer implements a reticular topology with bidirectional edges, allowing real-time consent updates to propagate backward through the lattice, triggering data quarantine or re-processing as needed.

To interpret system dynamics, we introduce three conceptual formulas:

1. Risk propagation in access chains: $R(p) = \sum \frac{Ci \cdot Pi}{Di}$, where R(p) denotes propagated risk for policy p, C_i is consent fragility at layer i,P_i is policy stringency, and D_i is data sensitivity depth. This formula interprets how risks amplify in multi-layer secondary accesses.
2. Governance load under constraints: $G_l = \alpha \cdot (N_c + N_p) + \beta \cdot F_r$, with G_l as governance load, N_c and N_p as the number of consents and policies, F_r as feedback reticulation factor, and α,β as interpretive coefficients for orchestration burden.
3. Drift sensitivity in consent feedback: $S_d = \int e^{-\gamma t} \cdot \Delta C(t)\ dt$, where S_d measures sensitivity to consent drifts over time T, ΔC(t) is the consent change rate, and γ \gamma γ is a decay factor reflecting policy resilience.

Figure 1 illustrates the consent-governed secondary analytics lattice (CGSAL), a reticular governance architecture in which consent tokens, policy predicates, analytic routing, and revocation feedback interact across lattice nodes to regulate secondary genomic–clinical analytics.

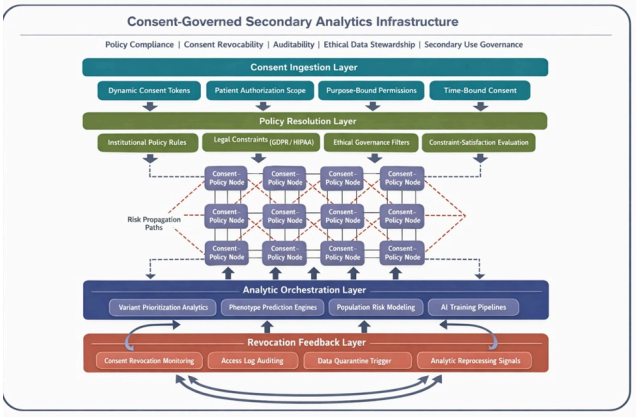

Figure 1. Consent-governed secondary analytics lattice (CGSAL): policy-constrained architecture for consent-aware genomic–clinical analytics

Table 1 delineates the functional responsibilities of each CGSAL architectural layer and clarifies how consent interpretation, policy enforcement, analytic routing, and revocation monitoring jointly govern secondary genomic-clinical data use.

Table 1. Structural functions of the CGSAL layers in policy-constrained secondary analytics

CGSAL layer	Core function	Governance mechanism	Data transformatio role

Consent ingestion layer	Captures patient authorization metadata and converts consent statements into machine-interpretable tokens	Dynamic consent parsing and scope tagging	Embeds permission boundaries int genomic-clinical data streams
Policy resolution layer	Evaluates analytic requests against institutional, regulatory, and ethical constraints	Predicate-based policy filtering and constraint satisfaction	Filters query pathways through policy predicates
Analytic orchestration layer	Routes consent-approved data to analytic modules	Policy-bound orchestration envelopes	Transforms approved data streams into A analytic workflows
Revocation feedback layer	Monitors consent updates and triggers governance responses	Revocation monitoring and audit logging	Propagates revocation signals across analytic pathways

Dynamics of policy-constrained secondary data flows

This section analyzes the theoretical dynamics and consequences of implementing the CGSAL in genomic-clinical ecosystems, focusing on how policy constraints shape data flows for secondary use. By examining ripple effects on system resilience, ethical equilibrium, and analytic adaptability, we elucidate potential impacts without empirical assertions. The reticular topology of CGSAL introduces nonlinear interactions where consent revocations can cascade, altering access pathways and governance demands.

In secondary data scenarios, such as repurposing genomic variants for population health AI models, policy constraints act as dampeners on flow velocity. The framework's lattice structure theoretically reduces unauthorized access by propagating policy checks across nodes, as captured in the risk propagation formula. For instance, high consent fragility (Ci C_i Ci) in early layers amplifies downstream risks, prompting adaptive rerouting to compliant paths [16, 23]. This dynamic fosters resilience against privacy breaches, particularly in multi-institutional settings where data from diverse clinical modalities converges [13, 19].

Ethical impacts emerge from the framework's emphasis on revocability, potentially shifting power dynamics toward patients. In clinical environments dealing with sensitive genomic data, such as in critically ill infants, CGSAL's feedback loops ensure that secondary analytic queries respect evolving consent, mitigating exploitation risks highlighted in tribal data sharing studies [8, 9, 24]. However, this introduces governance load, as per the formula G_l, where increased feedback reticulation (F_r) escalates orchestration burdens, theoretically straining resource-limited healthcare systems [25, 28].

Analytic adaptability is another key dynamic, where policy-enforced orchestration allows for modular AI integration without compromising secondary use ethics. In genomic-clinical fusion, drift sensitivity (S_d) interprets how consent changes over time affect analytic stability; rapid drifts in high-ΔC(t) scenarios could necessitate query throttling, preserving equity in AI outcomes [6, 26]. Consequences include enhanced interoperability with standards like HL7 FHIR, enabling theoretical scalability in global health digitalization efforts [1, 22]. Yet, over-constrained policies might inadvertently limit data utility, echoing challenges in COVID-19 registries where FAIR sharing balances access with consent [4]. **Table 2** synthesizes the conceptual governance metrics embedded in CGSAL, illustrating how risk propagation, governance load, and consent drift sensitivity shape the stability of policy-constrained secondary analytics.

Table 2. Governance dynamics and system metrics in the CGSAL architecture

Governance metric	Conceptual formula	Interpretation	Govern rol
Risk propagation in access chains	$R(p) = \frac{\sum C_i \cdot P_i}{D_i}$	Measures how risk accumulates across consent-policy layers	Identifies risk an pathw requi additi poli enforce
Governance load	$Gl = \alpha(N_c + N_p) + \beta F_r$	Estimates the operational burden of managing consents, policies, and feedback loops	Quant govern compl within analy infrastru
Consent drift sensitivity	$Sd = \int_0^T e^{-\gamma t} \Delta C(t)\, dt$	Captures system sensitivity to temporal changes in consent preferences	Dete instab introduc evolv cons condit
Policy stringency factor	P_i	Represents the restrictiveness of policy constraints	Determir streng poli enforcen each la nod
Feedback reticulation factor	F_r	Measures the density of revocation feedback loops across the lattice	Control responsiv of gover mechar

Broader system-wide impacts involve infrastructural transformation, aligning with mappings of national data flows [2]. By constraining secondary uses to policy-approved lattices, CGSAL theoretically minimizes bias propagation in AI algorithms derived from commercial datasets [5]. In registry contexts, this promotes "registry science" principles, where atomic data warehouse designs integrate consent-aware controls for sustained analytic value [21]. Ultimately, these dynamics underscore CGSAL's role in fostering a balanced ecosystem, where secondary data flows enhance AI healthcare without eroding trust [3, 27].

Results and Discussion

The CGSAL advances conceptual paradigms in AI for healthcare by embedding policy-constrained access into genomic-clinical workflows, addressing longstanding gaps in secondary data governance. This discussion synthesizes the framework's theoretical contributions, limitations, and alignments with extant literature, while exploring avenues for conceptual refinement.

CGSAL's reticular topology innovates beyond traditional linear access models, offering a governance-centric infrastructure that theoretically harmonizes consent with analytic demands. By layering consent ingestion with policy resolution, the framework mitigates risks in secondary uses, such as unauthorized genomic reanalysis, aligning with privacy evolution in informatics competitions [16]. Patient-centered features, like dynamic revocation loops, resonate with empirical insights on consent preferences, enhancing autonomy in biospecimen research [7, 17]. In policy-constrained environments, this orchestration supports equitable data sharing, countering biases in AI health algorithms [5, 6]. Moreover, integration with EHR standards facilitates theoretical interoperability, as seen in FHIR genomics operations, potentially transforming clinical deployments [22].

However, conceptual limitations warrant scrutiny. The governance load formula highlights potential scalability issues in high-volume settings, where numerous consents and policies could overwhelm orchestration [25]. In multi-jurisdictional contexts, varying policy stringencies might fragment lattice coherence, echoing challenges in global data flows [1, 2]. Ethical tensions arise if consent metadata oversimplifies patient intent, as lengthy documents often do, risking misaligned secondary uses [11]. Furthermore, while drift sensitivity accounts for temporal changes, it assumes uniform policy resilience, which may not hold in culturally diverse governance models [24].

Aligning with literature, CGSAL builds on CODE-EHR best practices for structured records, extending them to genomic-clinical secondary analytics [3]. It complements

atomic warehouse designs by incorporating consent-aware nodes, enhancing registry utility [19, 21]. Distributed ledger inspirations for genomics suggest blockchain augmentations to bolster revocation feedback, though CGSAL remains agnostic to specific technologies [20]. Data sharing evaluations post-ICMJE underscore the framework's potential to improve reproducibility in secondary research [10, 14]. In equity-focused discourse, the lattice's dynamics promote fairness pillars, ensuring AI avoids "open data in appearance only" pitfalls [26, 27].

Future conceptual extensions could incorporate adaptive learning mechanisms within the lattice, where policy feedback informs consent templates, drawing from pragmatic trial ethics [15]. Exploring modality-specific sublattices for genomic versus clinical data could refine access granularity [23]. Additionally, theoretical simulations of risk propagation under varying decay factors (γ) might illuminate resilience in fast-evolving healthcare policies [12]. Stakeholder engagement, as in community-based studies, could validate conceptual applicability across diverse clinical settings [24, 28].

Overall, CGSAL enriches AI healthcare discourse by prioritizing consent-aware governance, paving the way for trustworthy secondary data analytics in genomic-clinical domains.

Conclusion

In conclusion, the policy-constrained access control framework embodied in the CGSAL offers a robust conceptual blueprint for navigating the complexities of secondary data use in AI-driven genomic-clinical analytics. By integrating dynamic consent mechanisms with policy orchestration and reticular feedback, CGSAL theoretically safeguards patient autonomy while enabling innovative secondary applications, addressing critical gaps in current healthcare systems.

This manuscript has outlined the framework's architectural layers, interpretive formulas for system dynamics, and potential impacts on resilience, ethics, and adaptability. Grounded in a synthesis of recent literature, from privacy protections to data sharing best practices, CGSAL aligns with evolving standards in informatics and ethics. Its emphasis on governance load and risk propagation provides tools for conceptual analysis, highlighting trade-offs in policy-constrained environments.

Ultimately, adopting such frameworks could transform secondary data ecosystems, fostering trust and equity in AI healthcare. Future conceptual work should explore hybrid integrations and modality extensions to enhance its applicability further.

Acknowledgements

None

Conflict of interest

None

Financial support

None

Ethics statement

None

Received: 13 Nov 2023 Revised: 24 Dec 2023 Accepted: 07 Feb 2024
Published online: 25 February 2024

References

Näher AF, Vorstenbosch S, Gribbell B, O'Sullivan L, Browne JL, Thompson B, et al. Secondary data for global health digitalisation. Lancet Digit Health. 2023;5(1):e29-e31. https://doi.org/10.1016/S2589-7500(22)00195-9.

Zhang J, Godec J, O'Leary C, Symons J, Morley J. Mapping and evaluating national data flows: transparency, privacy, and guiding infrastructural transformation. Lancet Digit Health. 2023;5(10):e737-e748. https://doi.org/10.1016/S2589-7500(23)00157-7.

Kotecha D, Asselbergs FW, Anker SD, Banerjee A, Baigent C, Beger B, et al. CODE-EHR best-practice framework for the use of structured electronic health-care records in clinical research. Lancet Digit Health. 2022;4(10):e757-e764. https://doi.org/10.1016/S2589-7500(22)00151-0.

Maxwell L, Shreedhar P, Dagliati A, De Angelis G, Marangoni P, Sansone S-A. FAIR, ethical, and coordinated data sharing for COVID-19 response: a scoping review and cross-sectional survey of COVID-19 data sharing platforms and registries. Lancet Digit Health. 2023;5(9):e577-e586. https://doi.org/10.1016/S2589-7500(23)00129-2.

Alberto IRI, Alberto NRI, Ghosh AK, Jain B, Jayakumar S, Martinez-Martin N, et al. The impact of commercial health datasets on medical research and health-care algorithms. Lancet Digit Health. 2023;5(5):e288-e294. https://doi.org/10.1016/S2589-7500(23)00025-0.

Alderman JE, Palmer J, Ganapathi S, McCradden MD, Glocker B, Liu X. Recommendations for transparent documentation of health datasets to mitigate bias in artificial intelligence: the STANDING Together consensus study. Lancet Digit Health. 2024;6(12):e827-e847. https://doi.org/10.1016/S2589-7500(24)00224-3.

Kim J, Kim H, Bell E, Bath T, Paul P, Pham A, et al. Patient perspectives about decisions to share medical data and biospecimens for research. JAMA Netw Open. 2019;2(8):e199550. https://doi.org/10.1001/jamanetworkopen.2019.9550.

Schwartz MLB, McDonald WS, Hallquist MLG, Hu Y, McCormick CZ, Walters NL, et al. Genetics visit uptake among individuals receiving clinically actionable genomic screening results. JAMA Netw Open. 2024;7(3):e242388. https://doi.org/10.1001/jamanetworkopen.2024.2388.

Callahan KP, Mueller R, Flibotte J, Largent EA, Feudtner C. Measures of utility among studies of genomic medicine for critically ill infants: a systematic review. JAMA Netw Open. 2022;5(8):e2225980. https://doi.org/10.1001/jamanetworkopen.2022.25980.

Danchev V, Min Y, Borghi J, Baiocchi M, Ioannidis JPA. Evaluation of data sharing after implementation of the international committee of medical journal editors data sharing statement requirement. JAMA Netw Open. 2021;4(1):e2033972. https://doi.org/10.1001/jamanetworkopen.2020.33972.

Emanuel EJ, Boyle CW. Assessment of length and readability of informed consent documents for COVID-19 vaccine trials. JAMA Netw Open. 2021;4(4):e2110843. https://doi.org/10.1001/jamanetworkopen.2021.10843.

Gupta R, Iyengar R, Sharma M, Cannuscio CC, Merchant RM, Mitra N, et al. Consumer views on privacy protections and sharing of personal digital health information. JAMA Netw Open. 2023;6(3):e231305. https://doi.org/10.1001/jamanetworkopen.2023.1305.

Youssef A, Ng MY, Long J, Miner A, Hernandez-Boussard T, Larson DB, et al. Organizational factors in clinical data sharing for artificial intelligence in health care. JAMA Netw Open. 2023;6(12):e2348422. https://doi.org/10.1001/jamanetworkopen.2023.48422.

Bergeat D, Lombard N, Gasmi A, Le Floch B, Naudet F. Data sharing and reanalyses among randomized clinical trials published in surgical journals before and after adoption of a data availability and reproducibility policy. JAMA Netw Open. 2022;5(6):e2215209. https://doi.org/10.1001/jamanetworkopen.2022.15209.

Miller DG, Kim SYH, Li X, Dickert NW, Flory JH, Runge MS, et al. Ethical acceptability of postrandomization consent in pragmatic clinical trials. JAMA Netw Open. 2018;1(8):e186149. https://doi.org/10.1001/jamanetworkopen.2018.6149.

Kuo TT, Ohno-Machado L. The evolving privacy and security concerns for genomic data analysis and sharing as observed from the iDASH competition. J Am Med Inform Assoc. 2022;29(12):2182-90.

Savage SK, LoTempio J, Smith ED, Andrew EH, Mas G, Kahn-Kirby AH, et al. Using a chat-based informed consent tool in large-scale genomic research. J Am Med Inform Assoc. 2024;31(2):472-8.

Walton NA, Ye ZJ, Zhang A, Gevaert O, Kamaya A. Enabling the clinical application of artificial intelligence in genomics: a perspective of the AMIA genomics and translational bioinformatics workgroup. J Am Med Inform Assoc. 2024;31(2):536-48.

Visweswaran S, Becich MJ, D'Itri VS, Sendro ER, MacFadden D, Anderson NR, et al. An atomic approach to the design and implementation of a research data warehouse. J Am Med Inform Assoc. 2022;29(4):601-8.

Beyene M, Harrar SW, Zhao H, Klein RJ. A scoping review of distributed ledger technology in genomics: thematic analysis and directions for future research. J Am Med Inform Assoc. 2022;29(8):1433-44.

Labkoff SE, Sittig DF, McCoy AB. Identifying the capabilities for creating next-generation registries: a guide for data leaders and a case for "registry science". J Am Med Inform Assoc. 2024;31(4):1001-10.

Dolin RH, Boxwala A, Shalaby J. Introducing HL7 FHIR genomics operations: a developer-friendly approach to genomics-EHR integration. J Am Med Inform Assoc. 2023;30(3):485-93.

Schairer CE, Cheung C, Rubanovich CK, Cho M, Cranor LF, Bloss CS. Disposition toward privacy and information disclosure in the context of emerging health technologies. J Am Med Inform Assoc. 2019;26(7):610-9.

Triplett C, Fletcher B, Vichi F, Rioux K, Peek N, Urquhart A, et al. Codesigning a community-based participatory research project to assess tribal perspectives on privacy and health data sharing: a report from the Strong Heart Study. J Am Med Inform Assoc. 2022;29(6):1120-7.

Doutreligne M, Hueso A, Allier A, Savy N, Lamer A. Good practices for clinical data warehouse implementation: a case study in France. PLOS Digit Health. 2023;2(7):e0000298. https://doi.org/10.1371/journal.pdig.0000298.

Sikstrom L, Maslej MM, Hui K, Findlay Z, Buchman DZ, Hill SL. Conceptualising fairness: three pillars for medical algorithms and health equity. BMJ Health Care Inform. 2022;29(1):e100459. https://doi.org/10.1136/bmjhci-2021-100459.

Watson H, Gallifant J, Lai Y, Radunsky AP, Villanueva C, Martinez N, et al. Delivering on NIH data sharing requirements: avoiding open data in appearance only. BMJ Health Care Inform. 2023;30(1):e100771. https://doi.org/10.1136/bmjhci-2022-100771.

Gordon B, Barrett J, Fennessy C, Cake C, Milward A, Irwin C, et al. Development of a data utility framework to support effective health data curation. BMJ Health Care Inform. 2021;28(1):e100303. https://doi.org/10.1136/bmjhci-2020-100303.