Privacy without Loss of Utility: Evaluation of De-identification Techniques in Deep Learning for Intensive Care Units

Vitor Matheus Valandro da Rosa[1] **, Giovana Nunes Inocêncio**[1] **, Jean Everson Martina**[1]

[1]Universidade Federal de Santa Catarina (UFSC)
Florianópolis – Brazil

vitorvalandro.dev@gmail.com, gioinocencio017@gmail.com,

jean.martina@ufsc.br

Abstract. *The growing adoption of Deep Learning in Intensive Care Units depends on access to high-granularity clinical data, which conflicts with privacy regulations such as Brazil's LGPD. This work evaluates the trade-off between de-identification techniques and clinical AI utility. Using the MIMIC-III database, we applied k-anonymity, l-diversity, and Differential Privacy to demographic attributes and trained the ConCare model for in-hospital mortality prediction. Re-identification risk was audited under Prosecutor, Journalist, and Marketer attack models. Results demonstrate that robust pseudonymization ($k = 10, l = 2$) reduces maximum risk from 100% to 1.09% without degrading predictive performance (AUROC 0.861 vs. 0.868 baseline). We conclude that preserving sample size is critical; Differential Privacy ($\epsilon = 2.0$) discarded 16.3% of data, yielding inferior clinical utility (0.853) compared to syntactic approaches.*

1. Introduction

Modern precision medicine is grounded in the massive analysis of data, where AI applied to healthcare acts as a catalyst for clinical decision support, ranging from risk triage to outcome prediction in Intensive Care Units (ICUs) [Junior et al. 2024, Pessoa et al. 2024]. The success of *Deep Learning* architectures in this domain, however, depends on the availability of high-granularity data — including physiological time series and detailed demographic contexts — which inevitably involves the processing of sensitive personal data.

This scenario conflicts with the current regulatory framework, notably Brazil's General Data Protection Law (Lei Geral de Proteção de Dados Pessoais, LGPD), generating direct challenges for the security of Health Information Systems and patient-centric confidence [Inocêncio et al. 2026]. The LGPD explicitly enumerates health-related data ("dado referente à saúde") among the categories of sensitive personal data (Art. 5, II), subjecting their processing to stricter legal bases — including explicit and specific consent — to mitigate risks of stigmatization and discrimination, a central concern in recent regulatory approaches [Gonçalo et al. 2025]. For scientific research, the LGPD itself (Art. 11, II, c) already provides a legal basis that may dispense with individual consent, provided that de-identification safeguards are guaranteed. Resolution CNS No. 738/2024

operationalizes this provision for the research ethics system, detailing the procedural conditions under which retrospective studies with de-identified data may proceed through the CEP/CONEP framework [Machado 2025].

To resolve this technical-legal dilemma — whose regulation is a priority on the 2025–2026 Agenda of Brazil's National Data Protection Authority (ANPD) [Autoridade Nacional de Proteção de Dados (ANPD) 2024] — it is fundamental to distinguish two concepts. Anonymization, as defined by the LGPD, uses reasonable technical means to permanently remove the possibility of associating data with an individual [Brasil 2018]. In contrast, pseudonymization masks identity by replacing direct identifiers with artificial codes; re-identification only becomes possible if the attacker crosses this data with an additional information base (such as a key mapping table), maintained and protected separately under strict control.

For this study, we adopt the term 'anonymization' in compliance with the LGPD as a risk mitigation process, in which techniques such as generalization, suppression, and k-anonymity fall under robust pseudonymization: they do not guarantee strict mathematical irreversibility, but significantly reduce the probability of re-identification in empirical clinical research contexts [Sousa et al. 2020].

Although there are extensive theoretical studies on de-identification techniques in healthcare, there is a lack of quantitative evidence in real ICU scenarios that simultaneously evaluate residual re-identification risk and performance degradation in predictive *Deep Learning* models. The bibliographic selection underpinning this study was conducted on PubMed, IEEE Xplore, and ACM Digital Library (2018–2026), combining the terms *anonymization*, *de-identification*, *deep learning*, and *ICU*. We prioritized systematic literature reviews and proceedings from reference conferences at the intersection of health and computing, such as recent editions of SBCAS, supplemented by backward snowballing on seminal works. To fill this gap, this work quantifies the exact balance between patient privacy protection and the maintenance of utility in clinical algorithms.

In practice, we conducted a task-based evaluation of the the in-hospital mortality prediction scenario. Using real data from the MIMIC-III database [Johnson et al. 2016] and the cross-attention ConCare model [Ma et al. 2020], we compared the direct impact of three data protection approaches on demographic data: k-anonymity, l-diversity (syntactic techniques), and Differential Privacy (probabilistic technique). The study measures and contrasts two fundamental axes: (1) the mitigation of re-identification risk, audited under Prosecutor, Journalist, and Marketer attack perspectives; and (2) the preservation of clinical utility, quantified by the model's discrimination (AUROC) and precision (AUPRC) metrics.

In this context, the present study is guided by the following research question: **Can anonymization enable the legal use of sensitive data under the LGPD, while preserving the clinical utility of *Deep Learning* models in ICUs?**

The structure of this article is organized as follows: Section 2 reviews the theoretical framework and related work; Section 3 details the materials and methods, encompassing sample filtering and the configuration of experimental privacy scenarios; Section 4 presents the quantitative results of risk audits and clinical performance, as well as discussing practical implications for the LGPD and the study's limitations; finally, Section 5

presents the conclusion.

2. Theoretical Framework and Related Work

The area of Privacy-Preserving Data Publishing (PPDP), extensively reviewed by Fung [Fung et al. 2010], investigates methods for transforming raw data into safe formats for sharing, aiming to mitigate re-identification risks while maintaining statistical utility for data mining tasks.

Systematic Literature Mappings and recent review studies have exhaustively investigated the impact of clinical data de-identification on the *Machine Learning* ecosystem [Kayaalp et al. 2021, Hansson et al. 2025, Inocêncio and Martina 2026]. These surveys point to a critical consensus: although anonymization of health data is feasible and legally required, the literature lacks metric standardization. A large portion of the solutions focuses on mathematical privacy guarantees without quantifying the severity of utility loss in clinical predictions (*task-based evaluation*).

In the brazilian national scenario, the community has led essential discussions. Recent work evidences the methodological effort to apply anonymization in medical contexts, such as the removal of identifiers in unstructured texts [Gonçalves et al. 2025]. Simultaneously, attention is drawn to the risk that the suppression of demographic attributes inherent to anonymization may introduce or conceal predictive biases, compromising AI equity in healthcare [Rodrigues et al. 2025]. Our work fills this gap between strict privacy metrics and practical diagnostic utility.

2.1. Deep Learning in Clinical Time Series

Processing ICU data poses challenges, including temporal irregularity and high dimensionality. To handle the sequential nature of this data, Recurrent Neural Networks (RNNs) such as *Long Short-Term Memory* (LSTM) and *Gated Recurrent Units* (GRU) have become the dominant architectures, as confirmed by large-scale clinical studies [Rajkomar et al. 2018] and recent systematic reviews [Morid et al. 2023].

Harutyunyan et al. [Harutyunyan et al. 2019] established the benchmark for MIMIC-III, demonstrating that LSTM/GRU architectures outperform classical models in mortality tasks. Their proposed preprocessing pipeline extracts 17 physiological variables (e.g., heart rate, pH, glucose) sampled hourly over the first 48 hours, generating tensors $X \in \mathbb{R}^{48\times17}$.

2.2. ConCare Architecture and the Static Context

Although effective in temporal modeling, traditional LSTMs and GRUs frequently underutilize patients' static demographic data. The **ConCare** model [Ma et al. 2020] innovates by proposing explicit fusion of dynamic data with the clinical profile.

The architecture uses a recurrent network to encode the time series of vital signs. Simultaneously, static features are projected into a latent space. The central innovation is the application of a cross-attention mechanism: the neural network uses the patient's demographic profile as a filter to weight the importance of specific moments in the clinical evolution. This formulation makes the integrity of the context fundamental; if anonymization degrades the precision of these attributes, the model loses the ability to weigh the severity of vital signs correctly.

2.3. Privacy Mechanisms and Guarantees

The literature on privacy-preserving data publishing distinguishes three fundamental paradigms, each offering different security guarantees against specific attack vectors [Fung et al. 2010].

2.3.1. k-Anonymity (Syntactic)

Proposed by Sweeney [Sweeney 2002], it offers protection against linkage attacks. The central guarantee is that an individual can "hide in a crowd", ensuring that any combination of visible demographic attributes (Age, Gender, and Ethnicity) is repeated in at least k records. This is achieved through generalization (e.g., replacing "34 years" with "30–35 years") and suppression (removal of outliers) — patterns recently formalized as reusable design solutions for privacy-preserving data publishing [Monteiro et al. 2024].

2.3.2. l-Diversity (Semantic)

k-anonymity is inherently vulnerable to homogeneity attacks. If all k patients in an anonymized group share exactly the same diagnosis, privacy is violated regardless of the masking. To address this, Machanavajjhala et al. [Machanavajjhala et al. 2007] proposed l-diversity, imposing a semantic constraint: each group must contain at least l distinct diagnoses. This ensures that medical inference never reaches absolute certainty, forcing the creation of clinically heterogeneous groupings.

2.3.3. Differential Privacy (Probabilistic)

Unlike structural approaches, Differential Privacy (DP), formalized by Dwork [Dwork et al. 2006], offers a mathematical guarantee immune to external knowledge accumulated by the attacker, based on the premise of "plausible deniability". The control of this guarantee is given by the privacy budget (ϵ). In this work, we use DP as a rigorous filter: the algorithm evaluates and suppresses records that, due to their clinical or demographic rarity, would require an excess of noise that would corrupt the pre-established budget.

3. Materials and Methods

The experiment followed a three-stage workflow: (1) Cohort definition, (2) Application of privacy algorithms to the total database, and (3) Training and validation of the predictive model.

Figure 1 presents the detailed flow diagram of the method, summarizing the three main stages of the experimental design.

3.1. Sample Definition (MIMIC-III)

We used the MIMIC-III v1.4 database [Johnson et al. 2016], a benchmark repository containing approximately 60,000 ICU admissions from the *Beth Israel Deaconess Medical Center*. The filtering process strictly followed the criteria from Harutyunyan et al.'s benchmark [Harutyunyan et al. 2019] and the ConCare baseline model [Ma et al. 2020].

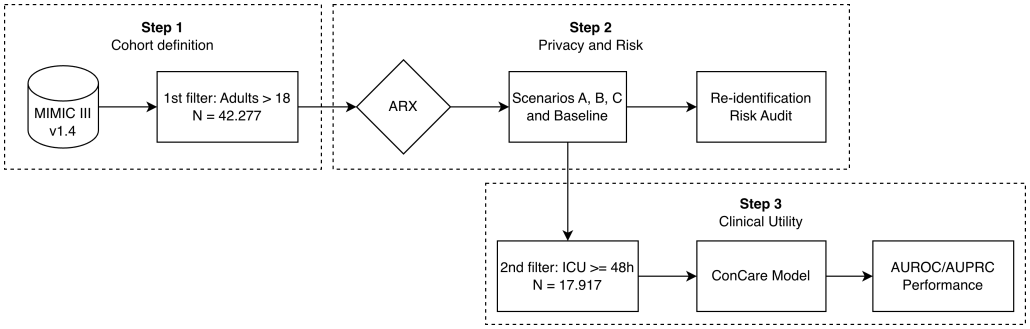

Figure 1. Methodological flow of the study detailing the stages of sample definition, application of privacy techniques, and clinical utility evaluation.

The filtering funnel occurred in two stages. Initially, the sample was restricted to adult patients ($\geq$ 18 years), establishing a primary universe of $N_{total} = 42{,}277$ admissions. This full database was used for privacy protection calculations in the ARX tool [Prasser et al. 2014], following the established principle that de-identification should be applied to the broadest available population to maximize equivalence class sizes and minimize information loss [El Emam and Arbuckle 2013, Pilgram et al. 2024]. The protection applied k-anonymity, l-diversity, and DP ($k = 5, k = 10, l = 2, \epsilon = 2.0$) to explore different levels of the information loss trade-off.

In the second stage, the benchmark's final rule was applied [Harutyunyan et al. 2019]: admissions with ICU stays shorter than 48 hours were removed, a requirement necessary to guarantee the completeness of the physiological time series required by ConCare. Thus, the final analytical sample used for training and testing totaled $N_{task} = 17{,}917$ admissions. Table 1 presents the characterization of this cohort.

Table 1. Characterization of the Training and Test Sample ($N = 17{,}917$)

Characteristic	**Statistic / Value**
Total Admissions	17,917
Mean Age	63.9 ± 17.5 years
Gender	Female: 44.1% / Male: 55.9%
Mortality	10.6% (1,899 deaths)
Top 5 Comorbidities (CCS)	**Prevalence in Sample**
Essential Hypertension	41.6%
Cardiac Arrhythmias	37.5%
Heart Failure	32.9%
Coronary Atherosclerosis	32.8%
Fluid Disorders	31.3%

The data from the final cohort were structured into two distinct vectors to feed the model architecture. The *Dynamic Vector* is a time series matrix containing 34 physiological variables sampled hourly over the first 48 hours (e.g., Heart Rate, Blood Pressure, Oxygen Saturation, Blood pH, Glucose, and Respiratory Rate). The *Static Vector* (Context) is a 42-dimensional vector containing original or anonymized demographic identi-

fiers (Age, Gender, Ethnicity), Admission Unit Type, and 25 prior comorbidities extracted via automatic grouping of ICD-9 codes using the CCS scheme.

3.2. Risk Modeling and Threat Models

For the pseudonymization strategy, *Age*, *Gender*, and *Ethnicity* were defined as Quasi-Identifiers (QIs), *Primary Diagnosis* as the Sensitive Attribute, and *Mortality* as the outcome. The evaluation followed El Emam's taxonomy [El Emam et al. 2009], considering three attack models. The *Prosecutor Risk* represents the worst-case scenario, where the attacker already knows that a specific individual is in the database and attempts to prove which record belongs to the target; the risk is calculated as $1/N$, where N is the size of the record's equivalence class. The *Journalist Risk* models the situation where the attacker knows the target's identity but is not certain whether the target is in the database, mitigated by population-level probability estimates. The *Marketer Risk* quantifies the average success rate of an attacker who attempts to re-identify as many random records as possible for profiling purposes, without a specific target.

The empirical quantification of these metrics was processed in ARX. The Prosecutor Risk is mathematically estimated as $1/k_{min}$, where k_{min} is the smallest grouping generated. The Marketer Risk represents the overall expected success ($E[1/k]$). Finally, the Journalist Risk applies population estimation statistical adjustments to mitigate the empirical Prosecutor risk.

3.3. Experimental Privacy Scenarios

Three protection scenarios were configured (Table 2), applied to the total database ($N = 42{,}277$) prior to the mortality filtering. The choice of privacy hyperparameters was anchored in consolidated guidelines from the health de-identification literature.

For syntactic models, the adoption of $k = 5$ (Scenario A) reflects the empirical safety threshold frequently recommended by clinical data sharing guidelines [El Emam et al. 2009]. The increment to $k = 10$ combined with $l = 2$ (Scenario B) establishes a more rigorous protection level, guaranteeing a minimum uncertainty (preventing deterministic inferences about the diagnosis) [Machanavajjhala et al. 2007]. In the probabilistic scope (Scenario C), the budget $\epsilon = 2.0$ was selected as it represents a common equilibrium point reported in real-world Differential Privacy applications in healthcare, ensuring strong "plausible deniability" without rendering the training of complex *Deep Learning* models unfeasible [Dwork et al. 2006].

The methodological distinction reveals different approaches. Scenario A (Light Syntactic) prioritizes Age granularity (short bands), accepting the total suppression of Ethnicity to achieve the desired anonymity ($k = 5$). Scenario B (Robust Syntactic) raises the requirement to $k = 10$ and applies the semantic constraint on the *Diagnosis* attribute; by forcing the aggregation of groups containing patients with at least two distinct diagnoses ($l = 2$), inference with 100% certainty is prevented, though this degree of heterogeneity requires more aggressive generalization of Age (10-year bands). Scenario C (Probabilistic/DP) prioritizes Ethnicity integrity to mitigate potential secondary algorithmic biases; to satisfy the rigorous privacy budget of $\epsilon = 2.0$ without suppressing this sensitive attribute, the algorithm opts for removing outlier records and reducing Age granularity.

Table 2. Configuration of Experimental and Privacy Scenarios

Scenario	Technique / Configuration	Age Treatment	Ethnicity Treatment
Baseline	Original Data (No anonymization)	**Exact Value** (e.g., 64 years)	**Original** (Detailed)
A	k-Anonymity ($k = 5$) Outlier Suppression	**5-year** bands (e.g., 60–65)	**Suppressed** (*)
B	k-Anonymity ($k = 10$) l-Diversity ($l = 2$)	**10-year** bands (e.g., 60–70)	**Suppressed** (*)
C	Differential Privacy ($\epsilon = 2.0$, $\delta = 10^{-5}$)	**10-year** bands (e.g., 60–70)	**Original** (Preserved)

4. Results and Discussion

The evaluation integrated risk auditing on the full demographic database ($N = 42{,}277$) and clinical utility validation strictly on the mortality subset ($N = 17{,}919$).

4.1. Re-identification Risk Audit ($N = 42k$)

The Baseline analysis confirmed the vulnerability of raw data: 99.95% of records were unique, resulting in a maximum Prosecutor Risk of 100%. Table 3 presents the mitigation results.

Table 3. Re-identification Risk Metrics ($N = 42{,}277$)

Scenario	Prosecutor *(Max)*	Journalist *(Max)*	Marketer *(Success)*	N Total	Loss (%)
Baseline	100%	100%	99.97%	42,277	0%
Scenario A	3.33%	3.33%	0.08%	42,272	≈0%
Scenario B	1.09%	1.09%	0.04%	42,272	≈0%
Scenario C	1.05%	1.05%	0.12%	31,010	16.3%

Scenario A ($k = 5$) reduced the maximum risk to 3.33% with no data loss. **Scenario B** ($k = 10, l = 2$) reduced the risk to 1.09%, rendering random re-identification (Marketer 0.04%) statistically improbable, while maintaining the sample volume intact. In contrast, **Scenario C** (DP) imposed a significant cost: to guarantee $\epsilon = 2.0$ while preserving the *Ethnicity* variable, the algorithm discarded 16.3% of records to create robust safety zones.

4.2. Clinical Utility ($N = 17k$)

Utility was measured by analyzing predicted probabilities ($\hat{y} \in [0, 1]$) vs. observed outcomes ($y \in \{0, 1\}$). We used AUROC (evaluating the overall capacity to classify deaths vs. discharges) and AUPRC, the primary metric due to the severe class imbalance in ICUs (mortality of only 10.6%), since it rigorously penalizes models that generate an excess of false positives. The *Gender* attribute was preserved across all scenarios.

To ensure statistical robustness of the comparisons, 95% Confidence Intervals (95% CI) were calculated using the non-parametric bootstrapping method, with 1,000 resamplings over the test set predictions. Table 4 presents the results.

Table 4. Predictive Performance with Confidence Intervals (95% CI) on the Test Set

Training Scenario	AUROC (95% CI)	AUPRC (95% CI)	Δ AUROC
Baseline (Original)	**0.868 (0.849–0.885)**	**0.425 (0.376–0.479)**	-
Scenario B (k=10, l=2)	0.861 (0.843–0.878)	0.419 (0.373–0.477)	-0.007
Scenario C (ϵ-DP)	0.853 (0.836–0.871)	0.412 (0.365–0.467)	-0.015
Scenario A (k=5)	0.841 (0.822–0.860)	0.402 (0.359–0.457)	-0.027
Control (Blind)	0.806 (0.785–0.828)	0.364 (0.319–0.418)	-0.061

The analysis of the intervals corroborates the stability of the predictions: AUROC showed a narrow variation (average margin of ± 0.018), while AUPRC exhibited wider variance (approximately ± 0.050), an expected and intrinsic statistical behavior in severely imbalanced ICU contexts. Despite this variance, the performance hierarchy remained unchanged, indicating that preserving sample volume (N) is the preponderant factor for clinical performance.

Scenario B achieved the best performance among the protected scenarios (AUPRC 0.419), remaining close to the Baseline. This result reveals a counterintuitive phenomenon: the superiority of B over Scenario A (AUPRC 0.402), even while imposing more severe demographic restrictions ($k = 10$ and $l = 2$). One hypothesis for this behavior is that the semantic constraint imposed by l-diversity acted as an indirect form of regularization during training. By forcing the aggregation of distinct diagnoses in the same group, the model may have been induced to rely less on demographic "shortcuts" and to extract more robust patterns from the physiological time series. In contrast, the finer granularity of Scenario A ($k = 5$) may have favored slight overfitting to the static profile of the training data, marginally impairing its generalization.

Scenario C (ϵ-DP) achieved intermediate performance (AUPRC 0.412). The suppression of 16.3% of records to satisfy the rigorous privacy budget negatively impacted predictive performance. This finding is consistent with Fakeeroodeen and Beeharry [Fakeeroodeen and Beeharry 2021], who showed that stand-alone Differential Privacy can destroy nearly all data utility on clinical datasets, whereas hybrid syntactic combinations (such as k+l) preserve both privacy and information content. The result demonstrates that the reduction in sample volume in imbalanced datasets impairs the modeling of the minority class (deaths), since DP tends to discard rare clinical examples (considered privacy outliers).

Finally, the Blind scenario (0.806) establishes the evaluation floor, with a drastic drop in AUPRC (0.364). This degradation confirms that ignoring the static patient context severely impairs the ability of the ConCare architecture to correctly identify mortality risk.

Practical Implications for LGPD Compliance. The practice of prophylactically removing the entire demographic profile (the Blind approach), frequently adopted out of excessive legal caution, substantially compromises AI utility. Although these findings derive from a single-center database and a specific architecture, the results suggest that applying robust pseudonymization techniques, such as k-anonymity with l-diversity, may offer a viable path forward. Sacrificing the precision of specific demographic attributes rather than discarding entire records can preserve the sample volume critical for digital

health.

4.3. Limitations of the Experimental Design

The study presents limitations inherent to the monocentric database (MIMIC-III) and the specific architecture used (ConCare). The behavior of other attention-based models requires further investigation to assess generalization. Furthermore, the utility metric focused exclusively on global performance. The suppression of the *Ethnicity* variable (Scenarios A and B) maintained acceptable global accuracy, but raises critical questions about algorithmic fairness — a concern amplified by evidence that clinical algorithms can embed systematic racial disparities even when overall performance appears adequate [Obermeyer et al. 2019]. As Chen et al. [Chen et al. 2023] emphasize, evaluating AI in healthcare demands subgroup-level analysis beyond aggregate metrics such as AUROC. It is imperative to investigate whether the suppression of demographic attributes entails the introduction of error disparities across patient subgroups.

5. Conclusion

This study demonstrated the technical viability of using pseudonymized data for training complex predictive models. The application of syntactic and probabilistic techniques reduced the empirical re-identification risk from 100% to levels of robust statistical security ($< 3.33\%$), overcoming the dilemma of "share everything or share nothing".

The central finding reveals a utility hierarchy in *Deep Learning* models: strategies that preserve the total sample volume at the cost of granularity (Scenario B, AUROC 0.861) outperform methods that preserve sensitive attributes at the cost of an abrupt sample reduction (Scenario C, AUROC 0.853). It is concluded that "absolute anonymization" is a theoretical ideal that can be circumvented by risk management through controlled pseudonymization.

As directions for future work, beyond investigating the impact of these modifications through the essential lens of algorithmic fairness, we propose the replication of this methodological evaluation in national clinical databases (such as from Brazil's Unified Health System — SUS). This validation on Brazilian data will allow exploring the practical integration of these anonymization approaches with the guidelines of Article 13 of the LGPD, evaluating how pseudonymization and the separation of identifiers can be orchestrated within the controlled environments required for public health research.

References

Autoridade Nacional de Proteção de Dados (ANPD) (2024). Agenda regulatória para o biênio 2025-2026: Diretrizes para inteligência artificial e dados de saúde. `https://www.gov.br/anpd/`. Acesso em: mar. 2026.

Brasil (2018). Lei nº 13.709, de 14 de agosto de 2018.

Chen, R. J., Wang, J. J., Williamson, D. F. K., Chen, T. Y., Lipkova, J., Lu, M. Y., Saber, S., and Mahmood, F. (2023). Algorithmic fairness in artificial intelligence for medicine and healthcare. *Nature Biomedical Engineering*, 7(6):719–742.

Dwork, C., McSherry, F., Nissim, K., and Smith, A. (2006). Calibrating noise to sensitivity in private data analysis. In *Theory of Cryptography Conference*, pages 265–284. Springer.

El Emam, K. and Arbuckle, L. (2013). *Anonymizing Health Data: Case Studies and Methods to Get You Started*. O'Reilly Media.

El Emam, K., Dankar, F. K., Issa, R., Jonker, E., Amyot, D., et al. (2009). A globally optimal k-anonymity method for the de-identification of health data. *Journal of the American Medical Informatics Association*, 16(5):670–682.

Fakeeroodeen, Y. N. and Beeharry, Y. (2021). Hybrid data privacy and anonymization algorithms for smart health applications. *International Journal of Advanced Computer Science and Applications (IJACSA)*, 12(6).

Fung, B. C., Wang, K., Fu, A. W.-C., and Pei, J. (2010). Privacy-preserving data publishing: A survey of recent developments. *ACM Computing Surveys (CSUR)*, 42(4):1–53.

Gonçalo, W. et al. (2025). Abordagens regulatórias na proteção de dados sensíveis na saúde digital: uma revisão integrativa. *Physis: Revista de Saúde Coletiva*, 35(1):e350113.

Gonçalves, A. C. M. et al. (2025). Anonimização de textos clínicos utilizando llm. In *Anais do Simpósio Brasileiro de Computação Aplicada à Saúde (SBCAS)*. SBC.

Hansson, M. et al. (2025). A systematic review of privacy-preserving techniques for synthetic tabular health data. *Artificial Intelligence in Medicine*.

Harutyunyan, H., Khachatrian, H., Kale, D. C., Ver Steeg, G., and Galstyan, A. (2019). Multitask learning and benchmarking with clinical time series data. *Scientific Data*, 6(1):1–18.

Inocêncio, G. N. and Martina, J. E. (2026). Assuring trustworthy data: A dual-criteria analysis of anonymization and system reliability in digital health (a systematic review). In *Anais do Simpósio Brasileiro de Sistemas de Informação (SBSI)*, Brasil. Sociedade Brasileira de Computação.

Inocêncio, G. N., Severo, L. P. F., and Martina, J. E. (2026). Assessing trustworthiness in digital health: Insights from the brazilian case of "meu sus digital". In *Proceedings of the 19th International Joint Conference on Biomedical Engineering Systems and Technologies - Volume 4: HEALTHINF*, pages 435–442. SCITEPRESS – Science and Technology Publications.

Johnson, A. E. W. et al. (2016). Mimic-iii, a freely accessible critical care database. *Scientific Data*, 3:160035.

Junior, J., Nakaya, H., and Rizzo, L. (2024). A inteligência artificial na medicina. *Revista de Medicina*, 103(1):1–5.

Kayaalp, M. et al. (2021). Data anonymization for pervasive health care: Systematic literature mapping study. *JMIR Medical Informatics*, 9(10):e29871.

Ma, L., Zhang, C., Wang, Y., Ruan, W., Wang, J., Tang, W., Ma, X., Gao, X., and Gao, J. (2020). Concare: Personalized clinical feature embedding via capturing the healthcare context. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 34, pages 844–851.

Machado, B. B. (2025). Proteção de dados e ética em pesquisa clínica: um estudo sobre o impacto da lgpd e da resolução cns/conep 738/2024 na condução de estudos clínicos

em território nacional. Master's thesis, Universidade Federal de São Paulo (UNIFESP), São Paulo, Brasil.

Machanavajjhala, A., Kifer, D., Gehrke, J., and Venkitasubramaniam, M. (2007). l-diversity: Privacy beyond k-anonymity. In *ACM Transactions on Knowledge Discovery from Data (TKDD)*, volume 1, page 3. ACM.

Monteiro, M., Correia, F., Queiroz, P., Ramos, R., Trigo, D., and Gonçalves, G. (2024). Patterns of data anonymization. In *Proceedings of the European Conference on Pattern Languages of Programs (EuroPLoP)*.

Morid, M. A., Sheng, O. R. L., and Dunbar, J. (2023). Time series prediction using deep learning methods in healthcare. *ACM Transactions on Management Information Systems*, 14(1):1–29.

Obermeyer, Z., Powers, B., Vogeli, C., and Mullainathan, S. (2019). Dissecting racial bias in an algorithm used to manage the health of populations. *Science*, 366(6464):447–453.

Pessoa, S. M. B. et al. (2024). Previsão de infecção relacionada à assistência à saúde em pacientes adultos de uti utilizando ferramentas de inteligência artificial. In *Anais do Simpósio Brasileiro de Computação Aplicada à Saúde (SBCAS)*. SBC.

Pilgram, L., Meurers, T., Malin, B., Schaeffner, E., Schwab, P., and Jensen, B. E. O. (2024). The costs of anonymization: case study using clinical data. *Journal of Medical Internet Research*, 26:e49445.

Prasser, F., Kohlmayer, F., and Kuhn, K. A. (2014). Arx—a comprehensive tool for anonymizing biomedical data. *Amia Annual Symposium Proceedings*, 2014:984.

Rajkomar, A., Oren, E., Chen, K., Dai, A. M., Hajaj, N., Hardt, M., Liu, P. J., Liu, X., Marcus, J., Sun, M., et al. (2018). Scalable and accurate deep learning with electronic health records. *NPJ Digital Medicine*, 1(1):18.

Rodrigues, D. D. et al. (2025). Bias propagation in health ai: Measuring pre-training bias and its effect on machine learning model outcomes. In *Anais do Simpósio Brasileiro de Computação Aplicada à Saúde (SBCAS)*. SBC.

Sousa, R. et al. (2020). Técnicas de anonimização de dados em saúde. *Journal of Health Informatics*, 12(3).

Sweeney, L. (2002). k-anonymity: A model for protecting privacy. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 10(05):557–570.