PRIVACY GUARANTEES REPORT

Bipolar Mixed States Synthetic Dataset v1.0

Privacy Metric	Value	Target	Status
k-Anonymity	k = 12	k ≥ 12	✓ CERTIFIED
Differential Privacy	ε = 0.78	ε < 1.0	✓ CERTIFIED
MIA Resistance	AUC = 0.52	AUC < 0.55	✓ CERTIFIED
l-Diversity	l = 1	l ≥ 2	■ PARTIAL

Overall Privacy Status: *ACCEPTABLE WITH CAVEATS*

1. K-Anonymity Analysis

K-Anonymity ensures that every record is indistinguishable from at least k–1 other records based on quasi-identifiers. This dataset achieves k = 12, meaning every combination of identifying attributes appears in at least 12 records.

1.1 Quasi-Identifiers

Variable	Type	Generalization Method
age_bucket	Categorical	Bucketed to 18–30, 31–45, 46–65
sex	Categorical	Some values suppressed
diagnosis	Categorical	7 ICD-10 codes (naturally k-anonymous)

Table 1.1: Quasi-Identifier Configuration

1.2 Equivalence Class Statistics

Metric	Value
Total Equivalence Classes	42
Minimum Class Size	12
Maximum Class Size	245
Mean Class Size	132.1
Records with Suppressed Sex	280 (5.0%)

Table 1.2: Equivalence Class Distribution

2. Differential Privacy Analysis

Differential privacy provides a mathematical guarantee that any individual's data has minimal impact on the output. We used DP-SGD (Differentially Private Stochastic Gradient Descent) during model training.

2.1 DP-SGD Configuration

Parameter	Value	Purpose
Gradient Clipping	max_norm = 1.0	Bound sensitivity
Noise Multiplier	σ = 0.5	Gaussian noise addition
Delta	δ = 1e−5	Probability of failure
Total Epsilon	ε = 0.78	Privacy budget used

Table 2.1: Differential Privacy Parameters

Interpretation: With ε = 0.78, the probability of any output is at most e^0.78 ≈ 2.18× more likely with vs. without any individual record. This is considered strong privacy (ε < 1 is the gold standard for sensitive data).

3. Membership Inference Attack Resistance

Membership Inference Attacks (MIA) attempt to determine whether a specific record was used in training the generative model. A successful defense makes such attacks no better than random guessing (AUC = 0.50).

Metric	Value	Interpretation
Attack AUC	0.52	Near random guess (0.50)
Attack Accuracy	51.8%	Minimal advantage over 50%
True Positive Rate	0.48	Cannot reliably identify members
False Positive Rate	0.46	High false alarm rate

Table 3.1: Membership Inference Attack Results

4. Risk Assessment

Attack Type	Risk Level	Mitigation
Re-identification	LOW	k-anonymity (k = 12)
Membership Inference	LOW	DP (ε = 0.78), synthetic generation
Attribute Inference	MEDIUM	l-diversity partial failure
Linkage Attack	LOW	No external identifiers
Model Inversion	LOW	Synthetic data, not model release

Table 4.1: Privacy Risk Matrix

5. Regulatory Compliance

5.1 GDPR Perspective

Requirement	Status	Notes
Lawful Basis	N/A	Synthetic, no personal data
Data Minimization	✓	Only necessary variables
Purpose Limitation	✓	Research/education only
Anonymization	✓	Fully synthetic generation

Table 5.1: GDPR Compliance Assessment