

# Ethical AI Guideline for Healthcare

## 1. Purpose

This guideline establishes principles for responsible AI use in healthcare, ensuring patient safety, fairness, and trust. It provides protocols for consent, bias mitigation, and transparency to align AI systems with ethical and legal standards.

## 2. Patient Consent Protocols

- Informed consent: Patients must be clearly informed when AI contributes to diagnosis, treatment, or triage. Explanations should cover benefits, risks, and alternatives in accessible language.
- Explicit authorization: Consent must be obtained before using sensitive data (e.g., genetic, biometric).
- Human oversight: Critical decisions must include clinician review; AI should support, not replace, medical judgment.
- Opt-out rights: Patients must be able to refuse AI involvement without losing access to care.

## 3. Bias Mitigation Strategies

- Representative datasets: Train AI on diverse populations to avoid underrepresentation of minorities, genders, or age groups.
- Fairness audits: Regularly measure subgroup performance (sensitivity, specificity, false positive/negative rates).
- Algorithmic safeguards: Apply fairness-aware methods (e.g., reweighing, adversarial debiasing) and recalibrate thresholds when disparities are detected.
- Continuous monitoring: Track model drift and retrain with updated data to prevent bias accumulation.

## 4. Transparency Requirements

- Model cards: Publish documentation detailing intended use, training data sources, performance metrics (overall + subgroup), and limitations.

- Explainability: Provide case-level rationales understandable to clinicians and patients.
- Audit trails: Log AI decisions for accountability and review.
- Incident reporting: Establish channels for reporting errors, adverse events, or bias, with clear remediation timelines.

## 5. Governance & Accountability

- Ethics review: Conduct Data Protection Impact Assessments (DPIAs) and ethics committee reviews for high-risk applications.
- Security & privacy: Enforce strict access controls, encryption, and minimal data retention.
- Lifecycle stewardship: Define roles for deployment, monitoring, and decommissioning; conduct periodic external audits.
- Sustainability: Optimize compute resources to reduce environmental footprint, aligning with healthcare sustainability goals.