

Policy Proposal: Ethical AI Use in Healthcare – Guidelines

1. Patient Consent Protocols

- Informed Consent: Patients must be clearly informed when AI is used in diagnosis, treatment planning, or decision support. Consent forms should specify the AI's role, data usage, and potential limitations.
- Opt-In/Opt-Out Rights: Patients must have the right to opt out of AI-assisted care without compromising access to standard medical services.
- Data Usage Clarity: Consent must distinguish between data used for direct care, model training, and research. Secondary uses (e.g., training) require separate, explicit consent.
- Dynamic Consent: Allow patients to update or withdraw consent at any time via accessible digital or in-person channels.

2. Bias Mitigation Strategies

- Diverse & Representative Data: Training datasets must include balanced representation across age, gender, ethnicity, socioeconomic status, and geographic region to avoid health disparities.
- Pre-Deployment Audits: Conduct rigorous bias testing using fairness metrics (e.g., equal error rates across demographic groups) before clinical deployment.
- Continuous Monitoring: Implement real-time performance tracking to detect and correct emerging biases during use (e.g., lower accuracy for rural or minority patients).
- Multidisciplinary Review: Include clinicians, ethicists, and community representatives in AI design and validation to identify context-specific fairness concerns.

3. Transparency Requirements

- Explainability: AI systems used in clinical settings must provide interpretable outputs (e.g., “This diagnosis is based on X-ray patterns similar to 95% of confirmed pneumonia cases”). Black-box models are prohibited for high-stakes decisions unless validated with external explainability tools.
- Disclosure to Clinicians and Patients: Healthcare providers must be trained to understand AI limitations and communicate them to patients. AI involvement in care must be documented in medical records.
- Public Documentation: Deployed AI tools must have publicly available documentation detailing purpose, data sources, validation results, known limitations, and contact for reporting issues.
- Third-Party Audits: Require annual independent audits of AI performance, bias, and clinical impact, with summaries made available to regulators and the public.

These guidelines ensure AI in healthcare upholds the core ethical principles of autonomy (via consent), justice (via bias mitigation), and transparency, aligning with international standards such as the EU's Ethics Guidelines for Trustworthy AI and the principles outlined in *Introduction to AI Ethics*.