Two-Tier Mathematical Risk Quantification (for AI-driven PatC* ICS**)

Our healthcare-specific risk definition provides theoretical foundation for a comprehensive mathematical framework. Consider medication safety assessment — clinical trials establish a drug's inherent risk profile, but prescribing decisions must account for patient-specific factors like age, comorbidities, and care coordination complexity that significantly modify those base risks. We developed a two-tier mathematical approach that mirrors this clinical reasoning process.

Healthcare AI risk assessment faces a fundamental challenge: while core risk relationships apply universally across clinical contexts, actual risk manifestation depends heavily on implementation circumstances. Our mathematical approach addresses this challenge through two complementary tiers. The core model quantifies fundamental risk relationships that apply across all healthcare AI implementations. The extended model then adjusts these base calculations for the contextual factors that determine how theoretical risks actually manifest in specific care settings.

Core Risk Model Foundation

Each risk component follows the probabilistic relationship:

$E[R_x(t)] = P_x(t) \times S_x(t)$ (Equation 1), where $P_x(t)$ represents occurrence probability and $S_x(t)$ represent severity magnitude.

The core model combines three essential risk components derived from our healthcare definition:

$R_{total}(t) = \alpha \times E[R_{implementation}(t)] + \beta \times E[R_{opportunity}(t)] + \gamma \times E[R_{interaction}(t)]$ (Equation 2), where α, β, γ are empirically-determined weights satisfying $\alpha + \beta + \gamma = 1$.

The explicit inclusion of both commission and omission risks in our healthcare definition provides theoretical justification for the α and β weighting coefficients.

Risk Components:

- **Implementation Risk:** Expected harm from AI system deployment, including false positives, negatives, and cascade effects.
- **Opportunity Risk:** Expected harm from delayed or absent AI implementation when beneficial interventions are available.
- **Interaction Risk:** Risks from human-AI workflow integration, including automation bias, alert fatigue, and override behaviors.

Extended Contextual Risk Model

While the core model captures universal risk relationships applicable across healthcare AI implementations, real-world deployment occurs within complex organizational, temporal, and social contexts that significantly modify risk outcomes. A diagnostic AI tool may have consistent performance characteristics (core risk), but its actual impact on patient care depends on factors like implementation delays in rural settings, differential access across populations, and uncertainty in local deployment conditions.

The extended model incorporates these implementation realities that determine how theoretical risks actually manifest in patient care:

$R\_total\_weighted(t) = R\_total(t) \times W\_temporal(\Delta t) \times W\_equity(g) \times W\_uncertainty(\sigma)$ (Equation 3)

Contextual Weighting Functions:

- **Temporal:** $W\_temporal(\Delta t) = e^{(-\lambda \Delta t)}$ captures time preference for healthcare interventions, operationalizing our definition's temporal dimension
- **Equity:** $W\_equity(g) = 1 + \delta \times H(g)$ applies penalties for health disparities across population groups
- **Uncertainty:** $W\_uncertainty(\sigma) = 1 + \varepsilon \times \sigma^2$ reflects risk-averse decision-making under model uncertainty

Mathematical Constraints: All probabilities and severities normalized to [0,1]; component weights sum to 1; contextual weights ≥ 1 (never reduce baseline risk). Complete specifications available in GitHub repository.

Our risk equations need real-world data to work. The following section shows how each mathematical component connects to existing healthcare data standards, enabling hospitals to implement our framework using data they already collect.

Abbreviations

*PatC - patient centered

**ICS – integrated healthcare systems