Theorem 1 (Risk-Theoretic Dominance)**.** *Let π_τ be a probability threshold policy. Define the loss functional as*

$$L(a, x) = \begin{cases} C_{miss}, & \textit{for a missed detection,} \\ C_{fp}, & \textit{for a false positive.} \end{cases}$$

Then the expected loss of the DAIS-10 policy is lower than or equal to the threshold policy:

$$\mathbb{E}[L|\pi_{DAIS}] \leq \mathbb{E}[L|\pi_\tau].$$

Proof. **Step 1: Define Decision Sets.**
DAIS-10 triggers "safe" under multiple conditions (high risk mass, timeout, persistent conflict, high CVaR), whereas π_τ only triggers under high probability. Therefore:

$$\{s : \pi_{DAIS}(s) = \text{safe}\} \supseteq \{s : \pi_\tau(s) = \text{safe}\}.$$

Step 2: Probability of Error Comparison.
Because DAIS-10 is more conservative:

$$P(M|\pi_{DAIS}) \leq P(M|\pi_\tau), \quad P(FP|\pi_{DAIS}) \geq P(FP|\pi_\tau).$$

Step 3: Expected Loss Difference.
The change in expected loss is:

$$\Delta\mathbb{E}[L] = \Delta P(M) \cdot C_{miss} + \Delta P(FP) \cdot C_{fp}.$$

Step 4: Dominance Conclusion.
By Axiom 1 (Safety Priority), $C_{miss} \gg C_{fp}$ and $\Delta P(M) \leq 0$, so the reduction in misses dominates any increase in false positives:

$$\Delta\mathbb{E}[L] \leq 0.$$

Hence:

$$\mathbb{E}[L|\pi_{DAIS}] \leq \mathbb{E}[L|\pi_\tau].$$

□