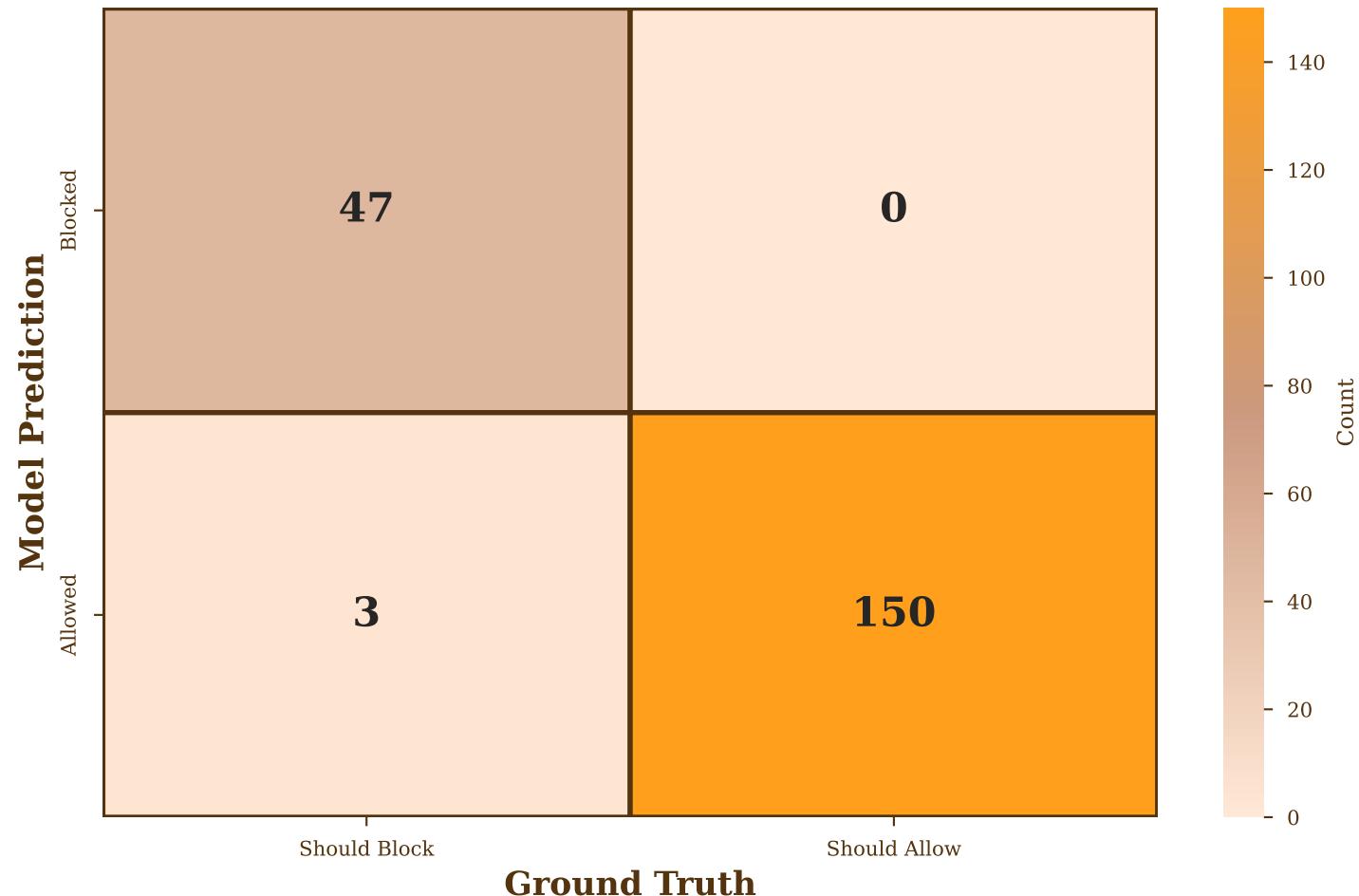


Precision: 94.00% (47/50)
Recall: 100.00% (47/47)
F1: 0.97

Regulatory Compliance: Medical Advice Detection Automated Red-Team Testing (N=200 prompts)



*Note: 3 false positives (1.5%) reflect conservative guardrails
(e.g., blocking "This sounds really hard" - fixed in v0.8.2)*