

# Regulatory Compliance: Red-Team Test Results

(N=200 adversarial prompts, medical advice detection)

