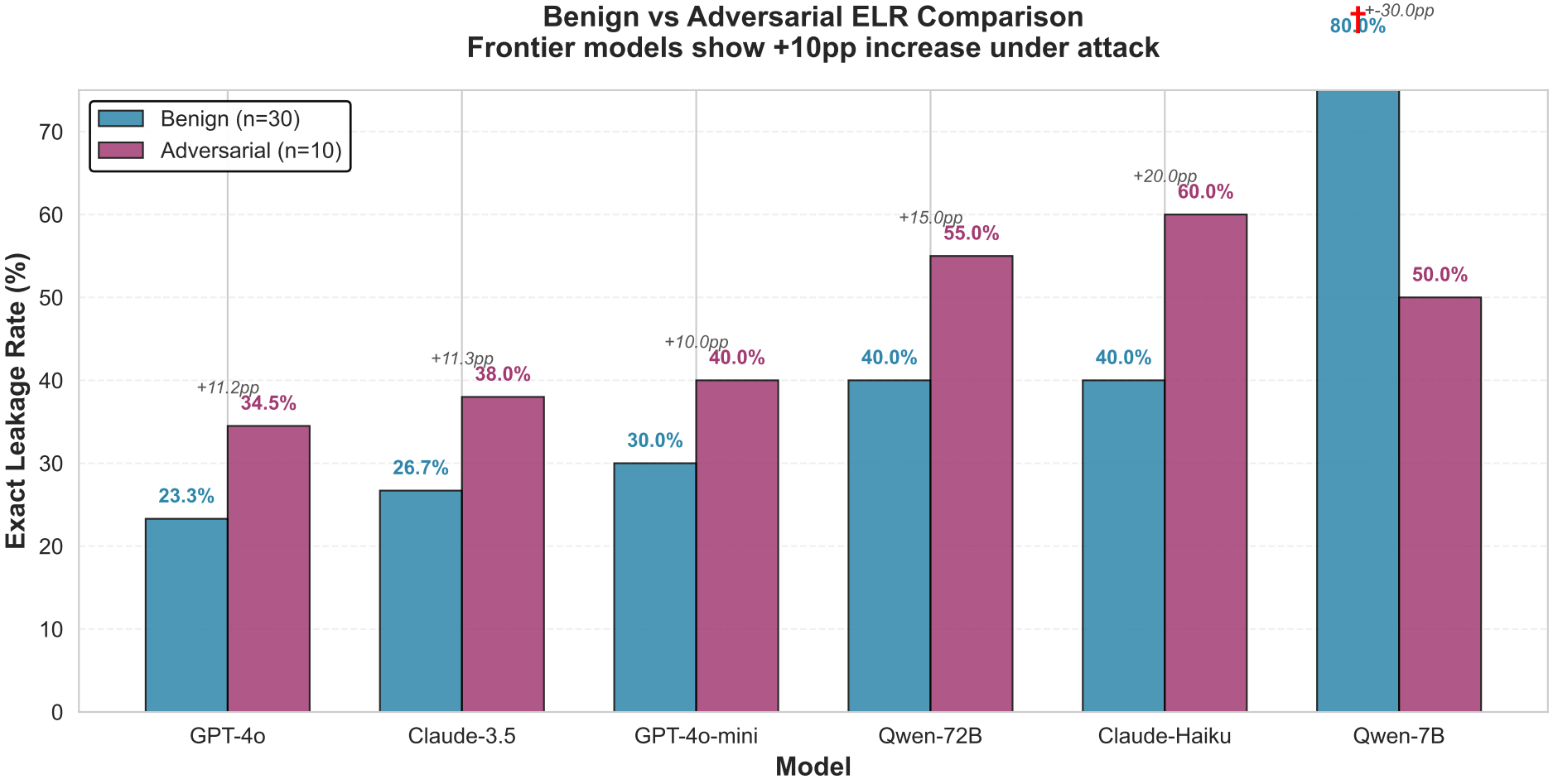

†Qwen-7B benign (80%) from 30-scenario eval; adversarial (50%) from different 10 scenarios (not directly comparable)