

Clinical Case Benchmarks

Real-World Clinical Scenarios



Patient history and presentation



Differential diagnosis generation



Diagnostic test ordering



Treatment plan development

Complexity Levels



Simple: Single diagnosis, clear presentation



Moderate: Multiple symptoms, common conditions



Complex: Comorbidities, atypical presentation



Expert: Rare diseases, multi-step reasoning

Evaluation: Diagnostic accuracy | Reasoning process | Treatment appropriateness