

MedQA Dataset Analysis

11,450

Total Questions

4-Choice

Answer Format

2 Lang

EN / ZH

USMLE

Style

Key Features



Clinical vignette-based questions



Multi-step reasoning required



Bilingual support (English/Chinese)



Difficulty levels from easy to expert

Human Expert Performance: ~87% | GPT-4: ~78% | Random Baseline: 25%