

## CoT vs Standard Prompting Performance

### Standard Prompting

- Direct answer generation
- No intermediate steps
- Faster but less accurate
- Harder to debug errors

### CoT Prompting

- Step-by-step reasoning
- Transparent logic
- Higher accuracy
- Auditable process

### Benchmark Results

**MedQA:** 56% → 72% (+16% with CoT)

**PubMedQA:** 68% → 81% (+13% with CoT)

**Clinical Diagnosis:** 62% → 78% (+16% with CoT)

**Treatment Planning:** 70% → 84% (+14% with CoT)

**Average Improvement:** +15% accuracy across medical tasks

**Trade-off:** 2-3x slower but significantly more accurate and trustworthy