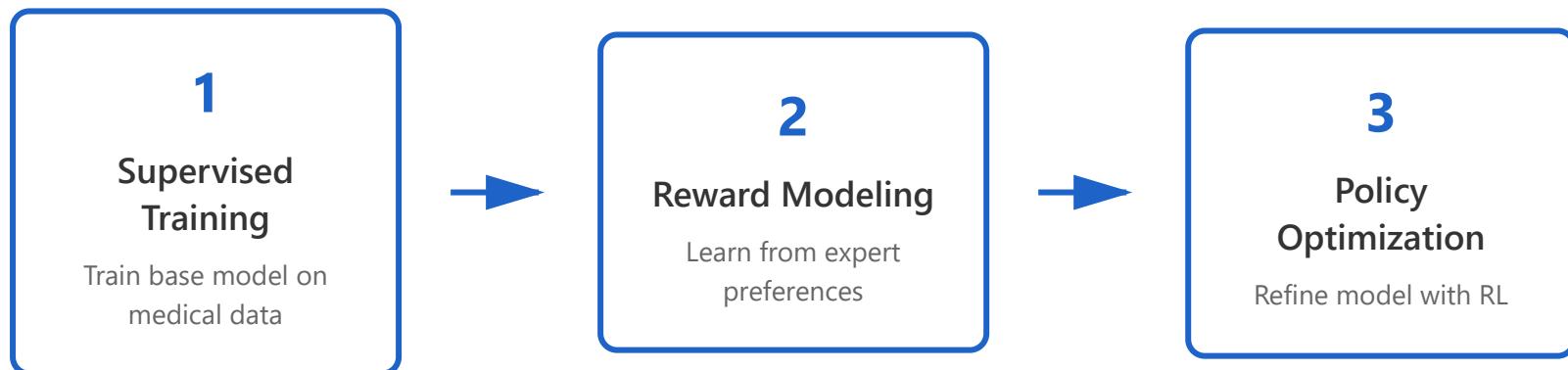


# RLHF Overview in Healthcare

## What is RLHF?

Reinforcement Learning from Human Feedback aligns AI models with medical expertise through iterative feedback loops, ensuring outputs match clinical standards and safety requirements.



### Clinical Alignment

Ensures AI decisions match medical standards

### Safety Enhancement

Reduces harmful or inappropriate outputs

### Continuous Improvement

Adapts to evolving medical knowledge

### Expert Integration

Incorporates physician judgment directly