



Weike Zhao<sup>1,2\*</sup>; Chaoyi Wu<sup>3\*</sup>; Yanjie Fan<sup>4,5\*</sup>; Pengcheng Qiu<sup>1,2</sup>; Xiaoman Zhang<sup>6</sup>; Yuze Sun<sup>1</sup>; Xiao Zhou<sup>2</sup>; Shuju Zhang<sup>7</sup>; Yu Peng<sup>7</sup>; Yanfeng Wang<sup>3,2</sup>; Xin Sun<sup>4</sup>; Ya Zhang<sup>3,8,2†</sup>; Yongguo Yu<sup>4†</sup>; Kun Sun<sup>4,9†</sup>; Weidi Xie<sup>3,2†</sup>

Email: [zwk0629@sjtu.edu.cn](mailto:zwk0629@sjtu.edu.cn) Web: [deprare.cn](http://deprare.cn)

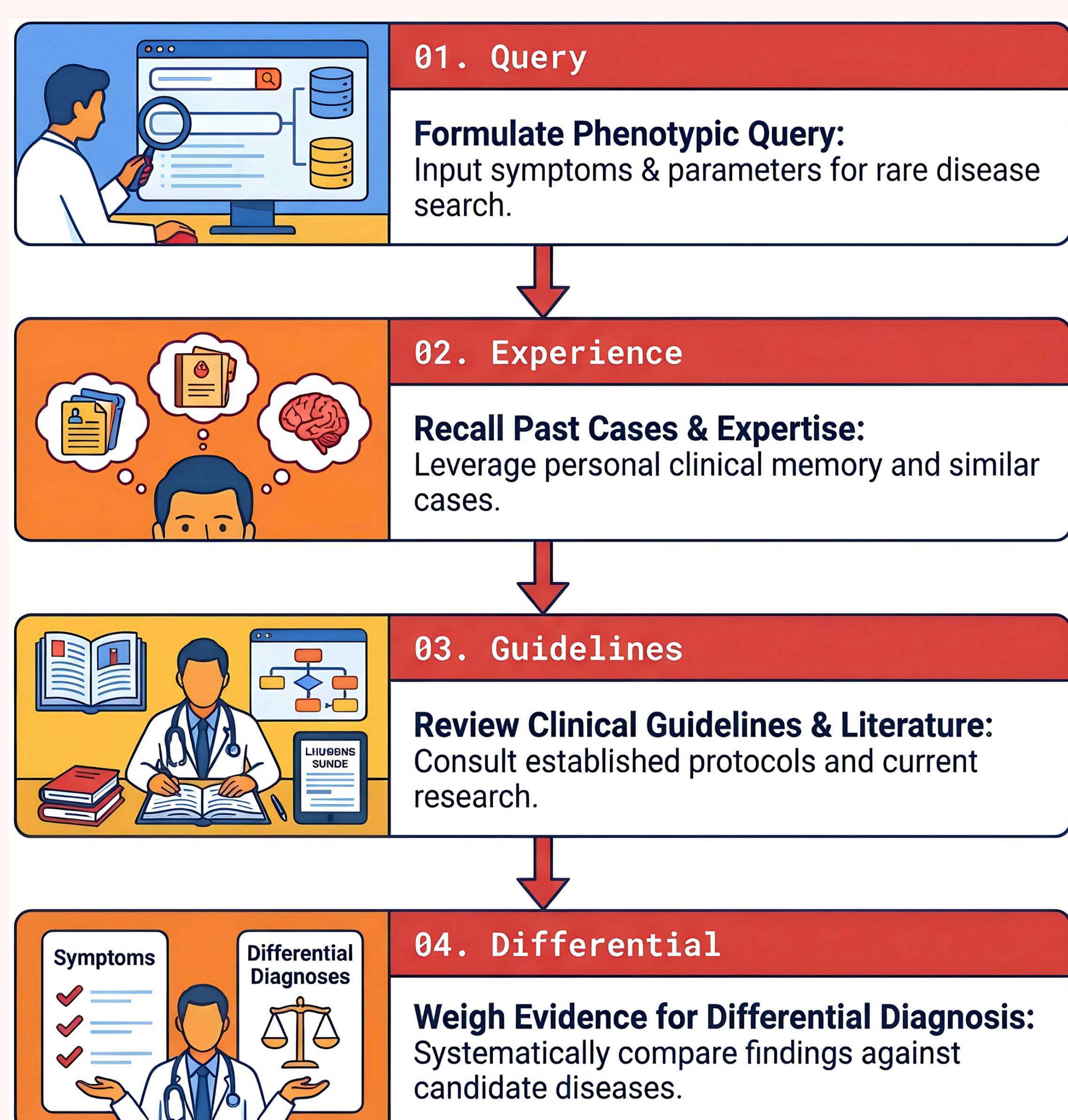
## MOTIVATION

Rare diseases affect over 300 million people worldwide. Patients with rare diseases often endure a “diagnostic odyssey” exceeding 5 years.

### Thinking Like a Physician

We propose a system that mimics a specialist’s cognitive process:

- ❖ **Knowledge Base:** LLMs act as the medical foundation.
- ❖ **Clinical Experience:** Case Searcher recalls similar past patients.
- ❖ **Literature Review:** Knowledge Searcher queries real-time guidelines.
- ❖ **Critical Thinking:** Self-Reflective loop prevents hallucinations.

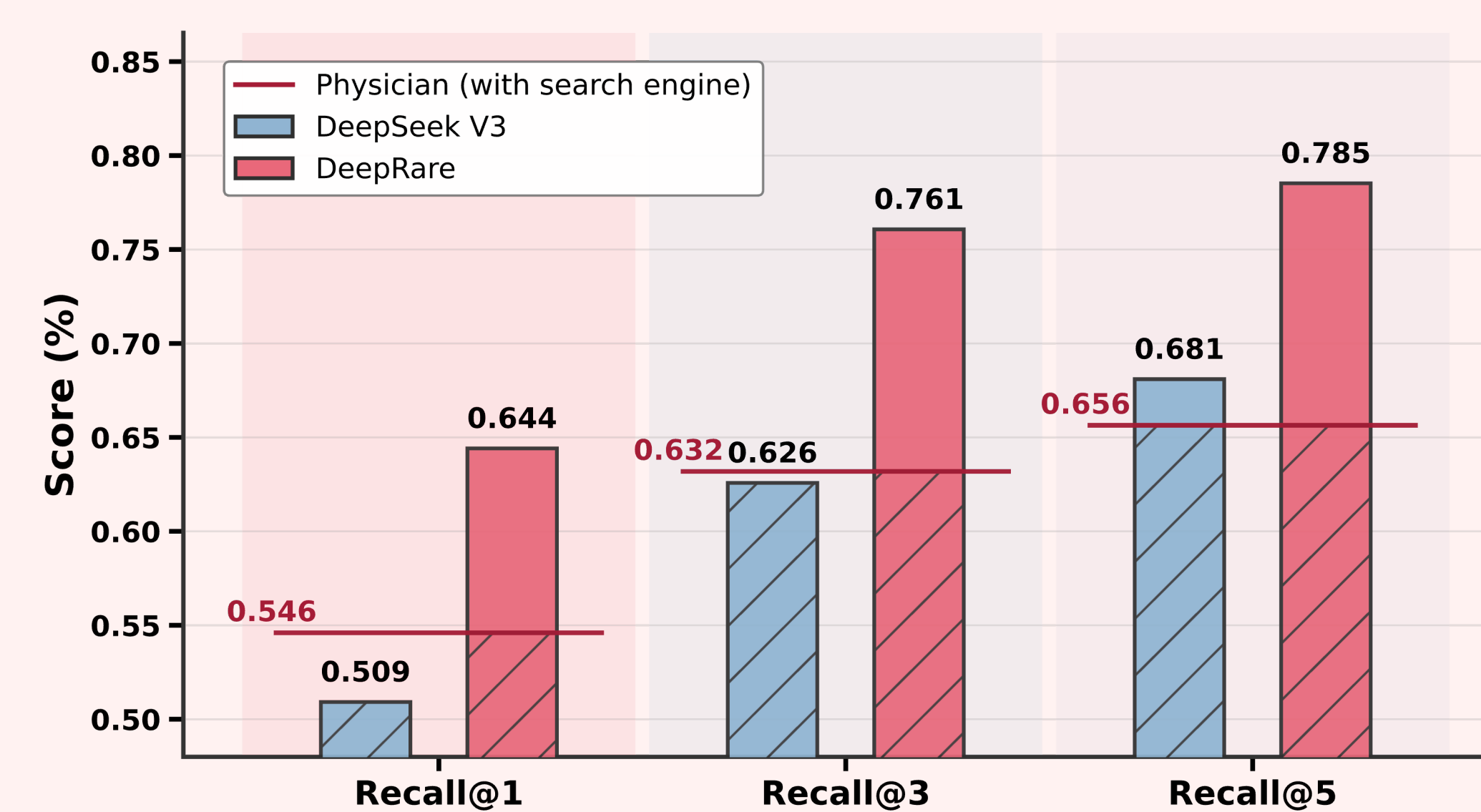


## RESULTS

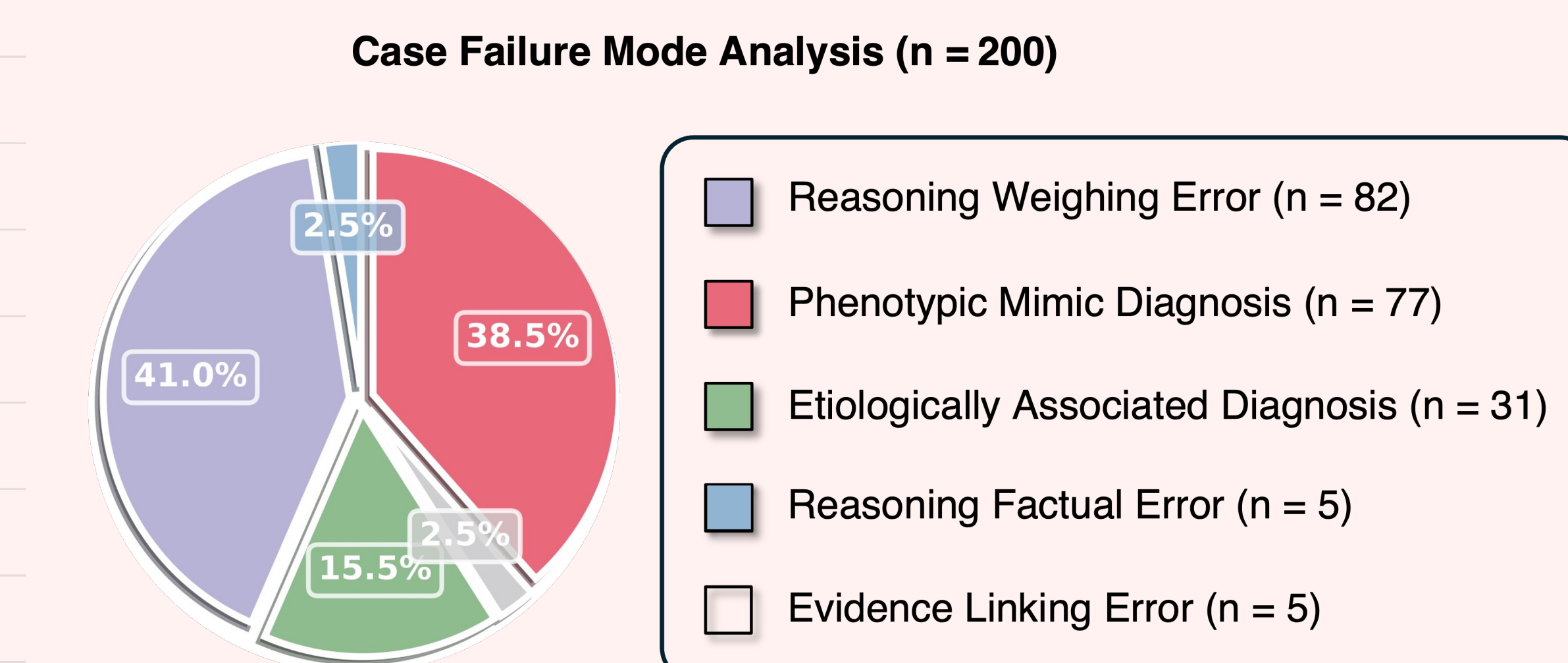
- **Phenotype-Based Diagnosis:** DeepRare achieved an average Recall@1 of 57.18% across 8 datasets, outperforming the next-best reasoning LLM by 23.79%.



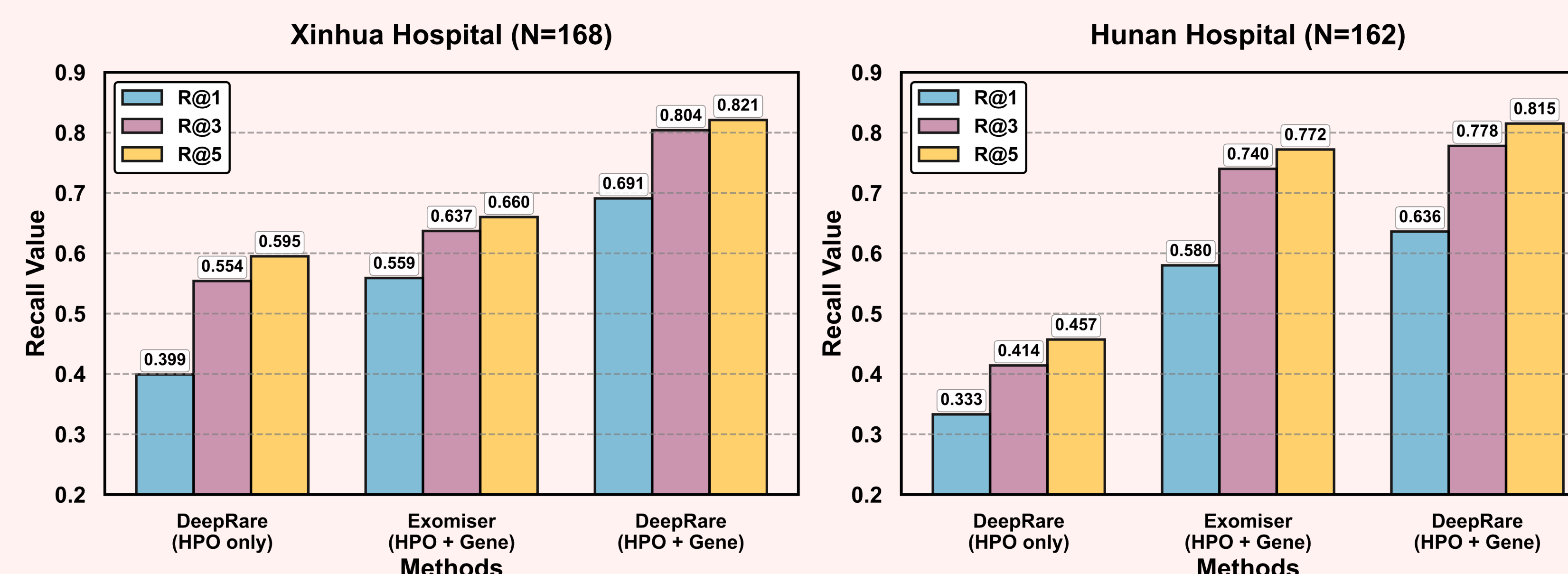
### DeepRare vs. Doctors



### Failure Case Mode Analysis



### Multi-Modal Analysis (HPO + VCF):



Significantly outperforming the standard bioinformatics tool (Exomiser).

## METHODS

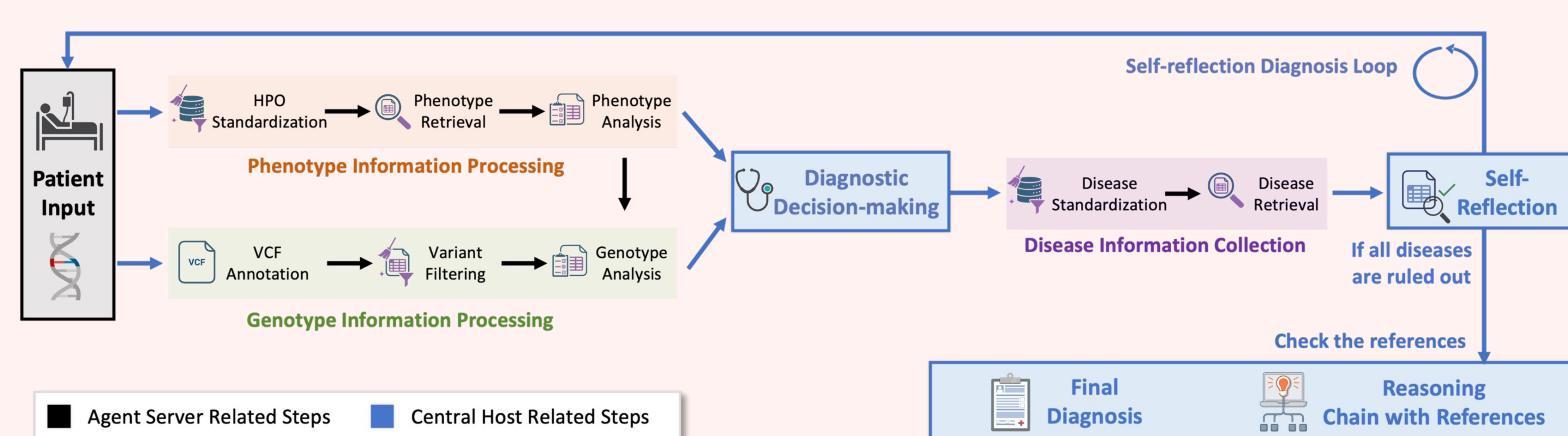
**System Inputs:** Free-text, HPO terms, & VCF files.



### Three-Tier Architecture:

- Central Host:** Orchestrates workflow via memory bank.
- Agent Servers:** Execute phenotype/genotype analysis & search.
- External Sources:** Integrates literature, Orphanet/OMIM, & case banks.

**Diagnostic Workflow:** Begins with Information Collection (HPO standardization and multi-source retrieval), followed by Self-Reflective Diagnosis to generate ranked predictions with traceable medical evidence.



## ONGOING WORK – DeepRare2

Shifting focus from differential diagnosis to precise Genomic Variant Interpretation.

- **Advanced Pathogenicity:** Integrating deep functional annotations & inheritance modeling.
- **Upgraded AI Agents:** More sophisticated design for deeper literature mining.
- **Enhanced Trust:** Improved attribution and traceability to shorten the diagnostic path.

