

From Accuracy to Equity: Addressing Hallucination Risks and Bias in AI-Driven Clinical Decision Support Systems

Reena Ughreja

School of Pharmaceutical Sciences, Atmiya University

INTRODUCTION & AIM



Study focuses on methods to reduce hallucinations and ensure fairness in AI use.

METHOD



RESULTS & DISCUSSION

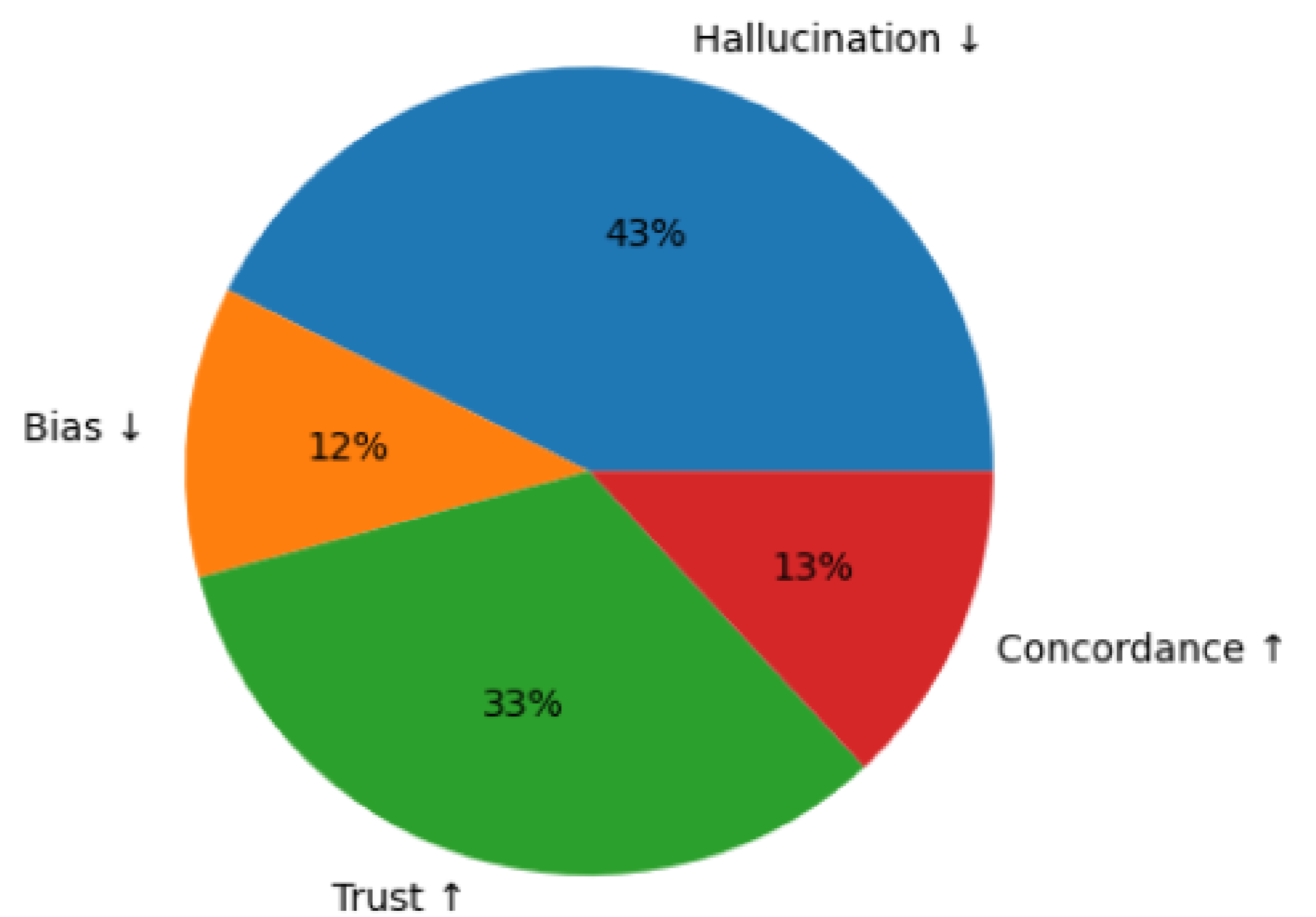
RAG + clinician feedback reduced **AI hallucinations by 41.7%** ($p < 0.001$)

Bias mitigation improved fairness, disparity reduced **18.3% → 6.9%** ($p = 0.004$)

Explainable AI increased **clinician trust by 32%**

Human-in-the-loop validation improved **clinical concordance from 76.4% to 89.2%**

RESULTS & DISCUSSION



CONCLUSION

- Multilayered approach reduces AI hallucinations in clinical systems
- Combination of model safeguards, data diversity, and human supervision improves performance
- Strategies help increase fairness and reliability in AI decisions
- Supports ethical and safe use of AI in healthcare
- Necessary for equitable clinical implementation of AI

FUTURE WORK

- Validate model on larger multi-centre clinical datasets
- Include real-time clinical decision support testing
- Improve bias detection across diverse populations
- Develop stronger explainable AI frameworks
- Study regulatory and ethical guidelines for clinical AI use

REFERENCES

References

- [1] Zang et al, *Mathematics* 2025, 13(5), 856
- [2] WHO regulation www.who.int/news
- [3] Hasanzadeh et al, *NPJ Digit Med.* 2025, 11(8), 154