Can Artificial Intelligence Produce Safe Infection-Prevention Leaflets for Maternity and Neonatal Patients?

Authors: Shideh Kiafar[1&2], Susan Knowles[1&2]
[1] University College Dublin ; [2] National Maternity Hospital, Dublin

Background

Artificial Intelligence (AI) can rapidly produce patient information leaflets. However, accuracy, clinical safety, and ethical standard remain a concern. Nurses and midwives need practical methods to evaluate whether AI generated materials could be trusted to meet the required patient safety, trust, and dignity.

Aim and Objectives

To evaluate and compare AI generated infection prevention leaflets with National Maternity Hospital (NMH) leaflets using a scoring tool across 9 domains: accuracy, safety, completeness, tone, readability, transparency, fairness, actionability, and overall suitability.

Methods

- Six AI leaflets generated using the same predefined prompt template, without access to NMH originals.
- Six NMH leaflets selected.
- Four independent reviewers scored all leaflets across nine domains using a 5-point scale (1 = poor, 5 = excellent).

Leaflet Topics (n=6)

- Neonatal MRSA
- Group B Streptococcus
- Mumps
- Zika virus
- Neonatal Multi-drug resistant organisms (MDROs)
- Postnatal wound care

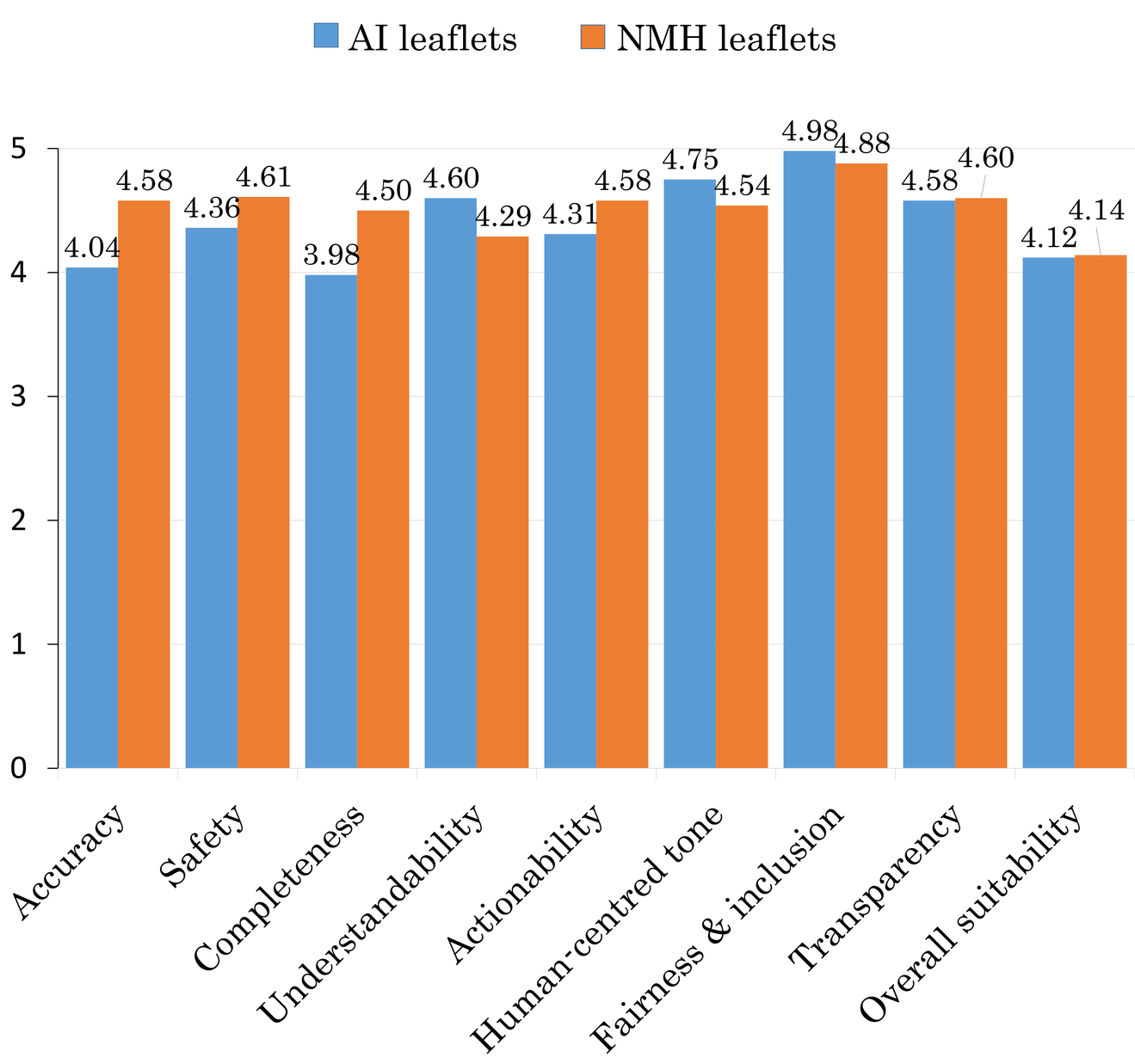
Reviewers (n = 4)

- Microbiology registrar
- Patient representative
- Midwife
- Information officer

422 total valid ratings
213 AI ratings | 209 NMH ratings

Key Findings

Mean Scores by Domain
(AI vs National Maternity Hospital Leaflets)



Summary of Selected Domains

Domain	AI Mean Score	NMH Mean Score	Key Findings
Accuracy	4.04	**4.58**	NMH scored higher
Safety	4.36	**4.61**	NMH scored higher
Completeness	3.98	**4.50**	NMH scored higher
Understandability	**4.60**	4.29	AI scored higher
Human-centred tone	**4.75**	4.54	AI scored higher
Overall	4.42/5	4.52/5	NMH scored higher

Scores: 1 = poor, 5 = excellent. No domain remained statistically significant after adjustment for multiple comparisons

Qualitative Reviewer Comments

AI leaflet strengths

- ✓ Plain language
- ✓ Human tone
- ✓ Transparency

Concerns identified in AI leaflets

- ✓ Accuracy concerns
- ✓ Oversimplification
- ✓ Clinically significant omissions required for decision-making

Implications

Applicability & Transferability

The scoring approach is repeatable, scalable, and suitable for other hospitals. Useful for governance, digital health evaluation, and patient education committees.

Practice Implications

Approach is repeatable, Ethical and clinical oversight and multidisciplinary approval needed for AI generated materials, staff require knowledge and clear guidance to evaluate AI generated materials.

Limitations

- Subjective scoring
- Partly unblinded reviews
- Single setting (six topics from one hospital).
- Rapidly evolving AI models

Conclusion

AI can produce clear and user-friendly information leaflets. However, AI generated materials cannot replace clinical expertise.

They require substantial clinical input, ethical review, and multidisciplinary oversight before use in maternity or neonatal settings. Variation in reviewer scoring highlights the need for staff AI literacy and multidisciplinary governance.

Full project summary

Scan the QR code to access the full project summary, including methodology and detailed results.

Scan me