Peer Review Activity – Unit 3 Seminar

Paper 1: Soenksen *et al.*, 2022

Purpose / contribution: This study presents the HAIM framework, integrating multiple data modalities (text, imaging, tabular, time-series) to enhance clinical predictions across healthcare tasks. It demonstrates that multimodal AI significantly improves diagnostic and prognostic accuracy compared to single-modality models, contributing to precision medicine and generalisable clinical decision-support systems.

Method appropriateness: Machine learning–based predictive modelling is suitable for the goal of improving clinical prediction. The inclusion of multimodal architectures and Shapley (SHAP) interpretability ensures the model captures complex relationships while remaining transparent for medical use.

Data collection & analysis: Using the large-scale MIMIC dataset, the authors trained and tested over 14,000 models across diverse prediction tasks. Cross-validation, ablation studies, and SHAP analyses robustly assess modality contributions and predictive performance, providing a strong analytical foundation aligned with the research objectives.

Evidence support: Empirical findings show 6–33% performance improvements with multimodal integration. Detailed ablation and interpretability analyses validate claims of increased generalisability and explainability. The main limitation is the lack of external validation beyond MIMIC, but the internal evidence is statistically rigorous and reproducible.

Enhancement suggestion: Future work should include external validation from independent hospitals, evaluate deployment feasibility in clinical workflows, and incorporate clinician feedback. Adding uncertainty quantification and testing robustness to missing data would further strengthen translational impact.

Paper 2: Amann *et al.*, 2023

Purpose / contribution: This qualitative study explores how stroke survivors, families, and healthcare professionals perceive AI in stroke medicine. It identifies expectations, trust issues, and ethical concerns shaping acceptance, contributing valuable human-centred insight into social and emotional aspects of AI adoption in clinical practice.

Method appropriateness: Semi-structured interviews provide the depth and flexibility required to capture nuanced perceptions. Grounded in the 'sociology of expectations,' the qualitative approach appropriately aligns with the aim of understanding lived experiences rather than measuring technical performance.

Data collection & analysis: Participants from Germany and Switzerland were interviewed and analysed using thematic coding. Both inductive and deductive methods

identified major themes such as trust, responsibility, and perceived benefits. This approach effectively reveals stakeholders' attitudes and ethical reflections about AI in stroke care.

Evidence support: Findings are well supported by direct participant quotations and coherent thematic structures. Transparency in coding, reflexivity, and cross-group comparison strengthen credibility. The study's small sample size limits generalisability but allows for deep, contextual interpretation.

Enhancement suggestion: Broader participant recruitment across healthcare systems would enhance transferability. A mixed-methods follow-up could quantify identified attitudes, and longitudinal interviews would reveal evolving perceptions as AI becomes more integrated in clinical settings.

References:

Amann, J. *et al.* (2023) 'Expectations and attitudes towards medical artificial intelligence: A qualitative study in the field of stroke', *PLOS ONE*, 18(1), p. e0279088. Available at: https://doi.org/10.1371/journal.pone.0279088.

Soenksen, L.R. *et al.* (2022) 'Integrated multimodal artificial intelligence framework for healthcare applications', *npj Digital Medicine*, 5(1), p. 149. Available at: https://doi.org/10.1038/s41746-022-00689-4.