AI-assisted oesophageal cancer MDT decision support: implications for real-world deployment of a validated machine learning tool

NHS University Hospital Southampton NHS Foundation Trust

NHS Fellowship in Clinical AI NHS Digital Academy Programme

University of Southampton

Mr Rory Ormiston, Ms Riannach Semple, Dr Nav Thavanesan, Prof Tim Underwood, Dr Ganesh Vigneswaran

Introduction

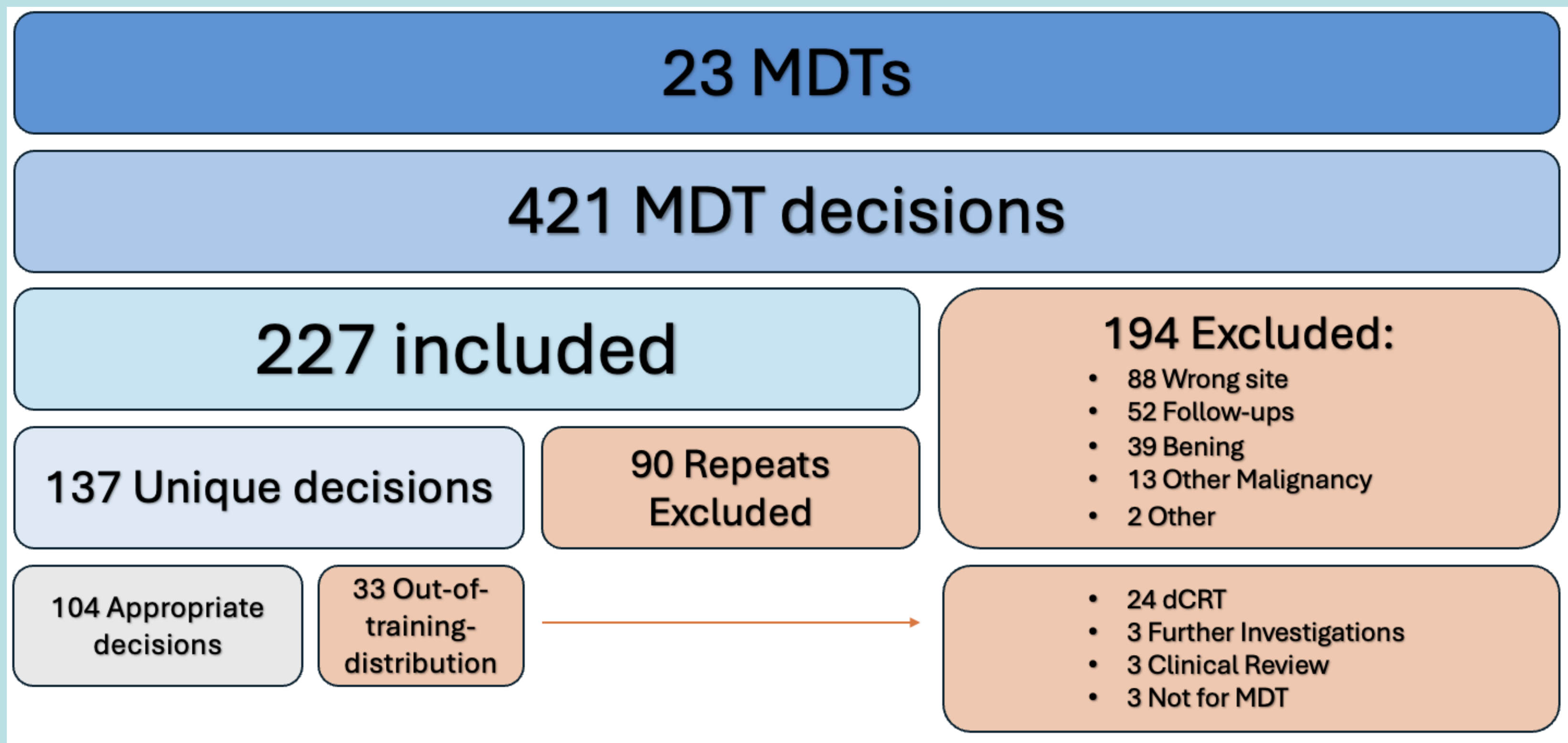
Machine learning (ML) can support multidisciplinary team (MDT) decision-making in oesophageal cancer, and we have previously developed and externally validated two ML clinical decision support tools (v1, v2). The next challenge is successful implementation into routine clinical practice. This study evaluates prospective real-world deployment of the tool in shadow mode to identify barriers to clinical implementation

Methods

The externally validated ML decision support tool (v2) was deployed in shadow mode across six months of routine oesophageal cancer MDT meetings. Eligible cases were entered prospectively, and implementation metrics including data completeness, ambiguous inputs and out-of-distribution MDT decisions were recorded. Findings were used to identify barriers to clinical deployment

Figure 1: Consort Diagram

A diagram showing the exact reasons for exclusion, the number of repeat decisions excluded and the out-of-training-distribution cases.



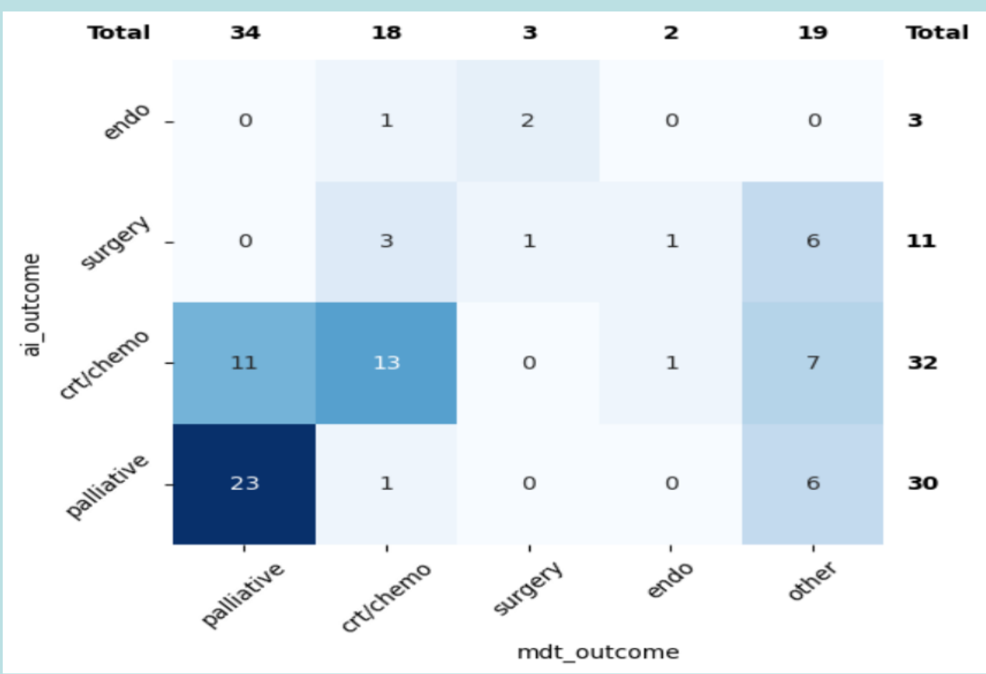
Results:

Twenty-three MDTs generated 421 discussions, with 137 unique eligible oesophageal cancer treatment decisions analysed. Twenty-four percent (33/137) of decisions were out-of-training-distribution, predominantly definitive chemoradiotherapy, leaving 104 decisions suitable for algorithm evaluation. Only 76/104 had complete input data. Within these 76 cases containing complete input data, the model achieved an overall accuracy of 49%.

Model v2

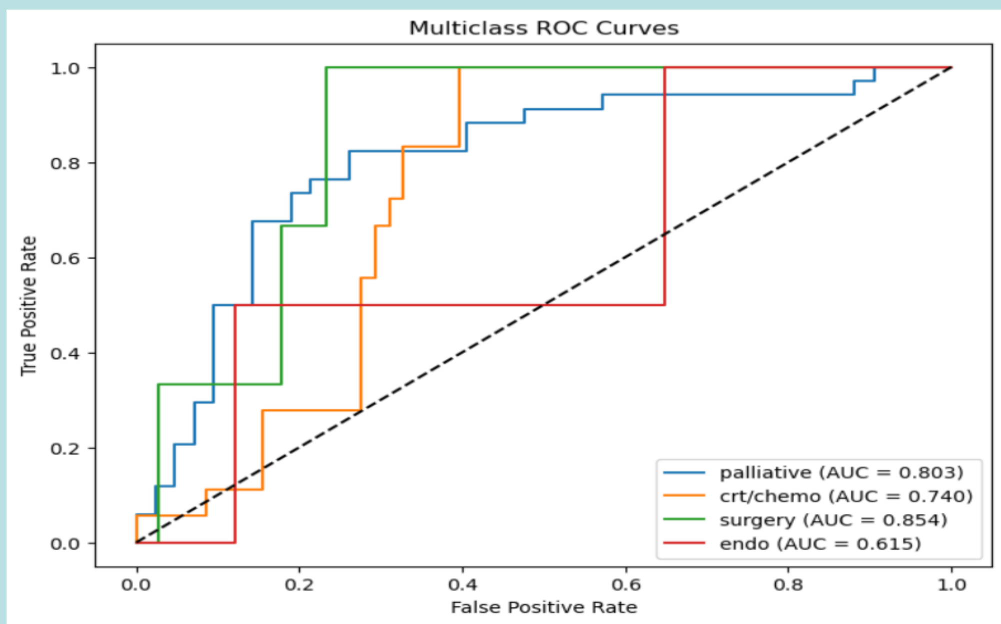
Model v2
Inputs:
Staging (T/N/M)
Histology (SCC/adeno)
Sex
Age
Performance status ECOG 0 - 4
Co-morbidities (7)
Outputs
Endoscopic resection
Surgery only
Surgery + chemotherapy
Surgery + chemoradiotherapy
Palliative Care

Figure 2: Confusion Matrix



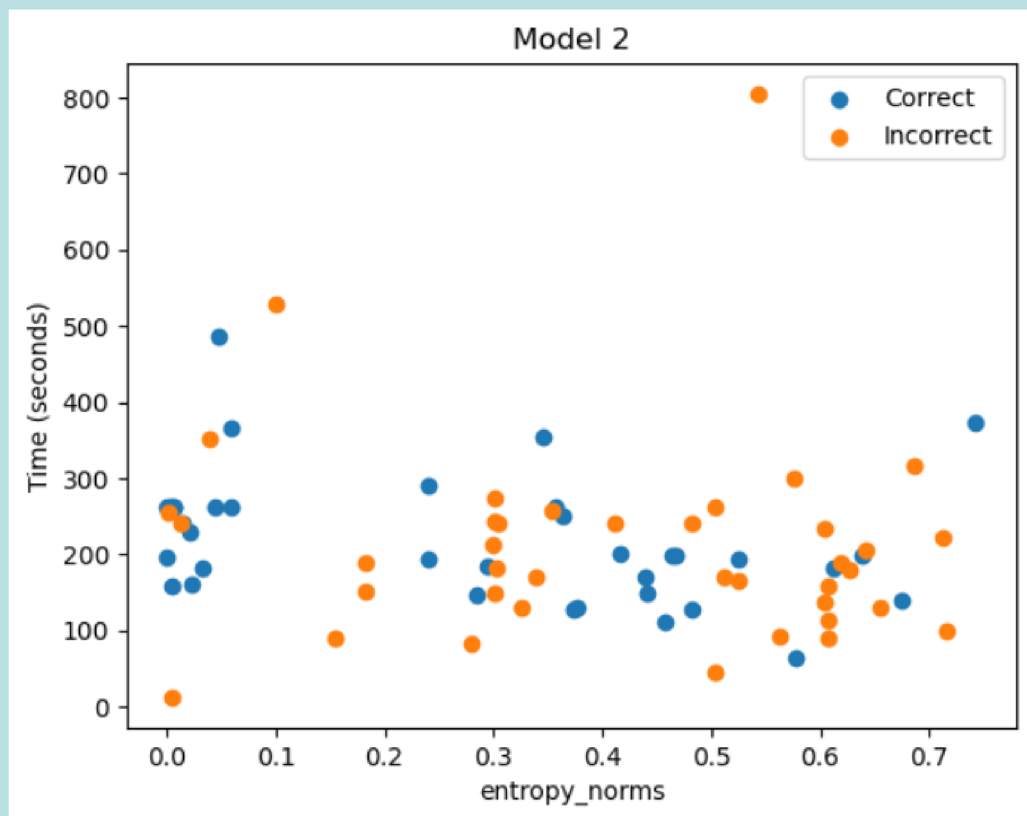
This confusion matrix shows the results for model V2 when input data was complete (73%, 76/104) and CRT and chemotherapy are combined.

Figure 3: ROC Curves



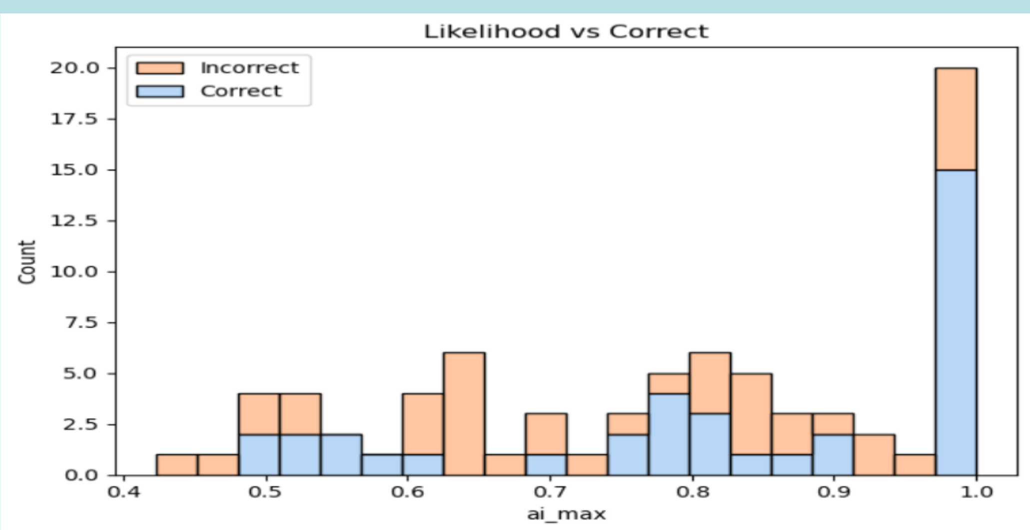
This diagram shows the ROC curves for each outcome. Only cases with complete data were used. CRT and chemo were combined.

Figure 4: Time - Entropy



This diagram shows correlation between MDT decision time (a proxy for complexity) and algorithm entropy. There is poor correlation.

Figure 5: Model Confidence and Prediction Accuracy



This graph shows that when the model (v2) is confident the accuracy is higher.

Conclusions:

Although the decision support tool demonstrated reasonable performance on recognised treatment pathways, prospective deployment exposed important real-world implementation challenges including out-of-training-distribution decisions, incomplete input data and evolving clinical practice. Model confidence was associated with prediction accuracy, but entropy did not reflect MDT decision complexity. These findings demonstrate that successful clinical AI depends as much on representative training data, robust routine clinical data capture and workflow integration as predictive performance.

Successful clinical AI is an implementation challenge as much as a machine learning challenge.

Lessons Learned:

- **High model performance ≠ successful implementation**
- **Registry datasets capture outcomes, not clinical reasoning**
- **Clinical practice evolves, models must evolve too**
- **Real MDT decisions are iterative, not point-in-time classifications**
- **Missing and ambiguous data challenge AI more than clinicians**
- **Decision support requires workflow integration, not just prediction**

References:

1. Thavanesan N, Vigneswaran G, Bodala I, Underwood TJ. The oesophageal cancer multidisciplinary team: can machine learning assist decision-making? J Gastrointest Surg. 2023;27(4):807-822. doi:10.1007/s11605-022-05575-8.
2. Thavanesan N, Bodala I, Walters Z, Ramchurn S, Underwood TJ, Vigneswaran G. Machine learning to predict curative multidisciplinary team treatment decisions in oesophageal cancer. Eur J Surg Oncol. 2023;49:106986. doi:10.1016/j.ejso.2023.106986.
3. Thavanesan N, Naiseh M, Terol M, Rahman SA, Hill SL, Parfitt C, et al. Oesophageal cancer multi-disciplinary tool: a co-designed, externally validated, machine learning tool for oesophageal cancer decision making. EClinicalMedicine. 2025;89:103527. doi:10.1016/j.eclinm.2025.103527.

Author contact details

Email: ro2e10@soton.ac.uk

FCAI Graduation Event 20th July 2027