Table 1 | The MI-CLAIM checklist

Before paper submission		
Study design (Part 1)	**Completed: page number**	**Notes if not completed**
The clinical problem in which the model will be employed is clearly detailed in the paper.	☑ Page 3 (Abstract), Page 5 (Introduction)	
The research question is clearly stated.	☑ Page 3 (Abstract), Page 5 (Introduction)	
The characteristics of the cohorts (training and test sets) are detailed in the text.	☑ Extended data table3 (Online Supplemental Tables)	
The cohorts (training and test sets) are shown to be representative of real-world clinical settings.	☑ Page 1. Dataset (Online Supplemental Methods)	
The state-of-the-art solution used as a baseline for comparison has been identified and detailed.	☑ Page 5 (Introduction), Page13-15 (Results)	
Data and optimization (Parts 2, 3)	**Completed: page number**	**Notes if not completed**
The origin of the data is described and the original format is detailed in the paper.	☑ Page 1. Dataset (Online Supplemental Methods)	
Transformations of the data before it is applied to the proposed model are described.	☑ Page 3-6 (Online Supplemental Methods)	
The independence between training and test sets has been proven in the paper.	☑ Page 3 (Online Supplemental Methods)	
Details on the models that were evaluated and the code developed to select the best model are provided.	☑ Page 8 (Online Supplemental Methods)	
Is the input data type structured or unstructured?	☑ Structured ☐ Unstructured	
Model performance (Part 4)	**Completed: page number**	**Notes if not completed**
The primary metric selected to evaluate algorithm performance (e.g., AUC, F-score, etc.), including the justification for selection, has been clearly stated.	☑ Page 8 (Online Supplemental Methods)	
The primary metric selected to evaluate the clinical utility of the model (e.g., PPV, NNT, etc.), including the justification for selection, has been clearly stated.	☑ Page 9 (Online Supplemental Methods)	
The performance comparison between baseline and proposed model is presented with the appropriate statistical significance.	☑ Page 5 (Introduction), Page13-15 (Results)	
Model examination (Part 5)	**Completed: page number**	**Notes if not completed**
Examination technique 1[a]	☑ Page 10 (Results)	
Examination technique 2[a]	☑ Page 12 (Results)	
A discussion of the relevance of the examination results with respect to model/algorithm performance is presented.	☑ Page 9 (Results)	
A discussion of the feasibility and significance of model interpretability at the case level if examination methods are uninterpretable is presented.	☑ Page 8 (Online Supplemental Methods)	
A discussion of the reliability and robustness of the model as the underlying data distribution shifts is included.	☒ Not completed	
Reproducibility (Part 6): choose appropriate tier of transparency		**Notes**
Tier 1: complete sharing of the code	☑	
Tier 2: allow a third party to evaluate the code for accuracy/fairness; share the results of this evaluation	☐	
Tier 3: release of a virtual machine (binary) for running the code on new data without sharing its details	☐	
Tier 4: no sharing	☐	

PPV, positive predictive value; NNT, numbers needed to treat. [a]Common examination approaches based on study type: for studies involving exclusively structured data, coefficients and sensitivity analysis are often appropriate; for studies involving unstructured data in the domains of image analysis or natural language processing, saliency maps (or equivalents) and sensitivity analyses are often appropriate.