



THE UNIVERSITY *of* EDINBURGH

Edinburgh Research Explorer

Machine learning to predict the likelihood of acute myocardial infarction

Citation for published version:

MI3 collaborative, Than , MP, Pickering, JW, Sandoval, Y, Shah, ASV, Tsanas, A, Apple, FS, Blankenberg, S, Cullen , L, Mueller , C, Neumann , JT, Twerenbold , R, Westermann , D, Beshiri, A & Mills, NL 2019, 'Machine learning to predict the likelihood of acute myocardial infarction', *Circulation*, vol. 140, no. 11, pp. 899-909. <https://doi.org/10.1161/CIRCULATIONAHA.119.041980>

Digital Object Identifier (DOI):

[10.1161/CIRCULATIONAHA.119.041980](https://doi.org/10.1161/CIRCULATIONAHA.119.041980)

Link:

[Link to publication record in Edinburgh Research Explorer](#)

Document Version:

Peer reviewed version

Published In:

Circulation

General rights

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.



Machine Learning to Predict the Likelihood of Acute Myocardial Infarction

Running Title: *Than et al.; Machine Learning in Myocardial Infarction*

Martin P. Than, MBBS^{1*}; John W. Pickering, PhD^{1,2*}; Yader Sandoval, MD³;
Anoop S.V. Shah, MBChB, PhD^{4,5}; Athanasios Tsanas, PhD⁵; Fred S. Apple, PhD⁶;
Stefan Blankenberg, MD⁷; Louise Cullen, PhD⁸; Christian Mueller, MD⁹;
Johannes T. Neumann, MD⁷; Raphael Twerenbold, MD⁹; Dirk Westermann, MD⁷;
Agim Beshiri, MD¹⁰; Nicholas L. Mills, MBChB, PhD^{4,5}; on behalf of the MI³ collaborative.

* Joint first authors

¹Emergency Department, Christchurch Hospital, Christchurch, New Zealand; ²Christchurch Heart Institute, Department of Medicine, University of Otago, Christchurch, New Zealand; ³Department of Cardiovascular Diseases, Mayo Clinic, Rochester, MN; ⁴BHF Centre for Cardiovascular Sciences, University of Edinburgh, Edinburgh, United Kingdom; ⁵Usher Institute of Population Health Sciences and Informatics, University of Edinburgh, Edinburgh, United Kingdom; ⁶Department of Laboratory Medicine and Pathology, Hennepin County Medical Center, and University of Minnesota, Department of Laboratory Medicine & Pathology Minneapolis, MN; ⁷Department of General and Interventional Cardiology, University Heart Center, Hamburg, Germany; ⁸Emergency Department, Royal Brisbane and Women's Hospital, Brisbane, Australia; ⁹Universitätsspital Basel, Basel, Switzerland; ¹⁰Abbott Diagnostics, Abbott Laboratories, Lake Forest, IL

Address for Correspondence:

Nicholas L. Mills, MBChB, PhD⁴
BHF/University Centre for Cardiovascular Science
University of Edinburgh
Room SU.226
Chancellor's Building
Edinburgh EH16 4SB
United Kingdom
Tel: +44 (0) 131 242 6378
Email: nick.mills@ed.ac.uk

Abstract

Background: Variations in cardiac troponin concentrations by age, sex and time between samples in patients with suspected myocardial infarction are not currently accounted for in diagnostic approaches. We aimed to combine these variables through machine learning to improve the assessment of risk for individual patients.

Methods: A machine learning algorithm (myocardial-ischemic-injury-index [MI^3]) incorporating age, sex, and paired high-sensitivity cardiac troponin I concentrations, was trained on 3,013 patients and tested on 7,998 patients with suspected myocardial infarction. MI^3 uses gradient boosting to compute a value (0-100) reflecting an individual's likelihood of a diagnosis of type 1 myocardial infarction and estimates the sensitivity, negative predictive value (NPV), specificity and positive predictive value (PPV) for that individual. Assessment was by calibration and area under the receiver-operating-characteristic curve (AUC). Secondary analysis evaluated example MI^3 thresholds from the training set that identified patients as low-risk (99% sensitivity) and high-risk (75% PPV), and performance at these thresholds was compared in the test set to the 99th percentile and European Society of Cardiology (ESC) rule-out pathways.

Results: Myocardial infarction occurred in 404 (13.4%) patients in the training set and 849 (10.6%) patients in the test set. MI^3 was well calibrated with a very high AUC of 0.963 [0.956-0.971] in the test set and similar performance in early and late presenters. Example MI^3 thresholds identifying low-risk and high-risk patients in the training set were 1.6 and 49.7 respectively. In the test set, MI^3 values were <1.6 in 69.5% with a NPV of 99.7% (99.5%-99.8%) and sensitivity of 97.8% (96.7-98.7%), and were ≥ 49.7 in 10.6% with a PPV of 71.8% (68.9-75.0%) and specificity of 96.7% (96.3-97.1%). Using these thresholds, MI^3 performed better than the ESC 0/3-hour pathway (sensitivity 82.5% [74.5-88.8%], specificity 92.2% [90.7-93.5%]) and the 99th percentile at any time-point (sensitivity 89.6% [87.4-91.6%]), specificity 89.3% [88.6-90.0%]).

Conclusions: Using machine learning, MI^3 provides an individualized and objective assessment of the likelihood of myocardial infarction, which can be used to identify low-risk and high-risk patients who may benefit from earlier clinical decisions.

Clinical Trial Registration: Unique Identifier: Australian New Zealand Clinical Trials Registry: ACTRN12616001441404. URL: <https://www.anzctr.org.au>

Key Words: Myocardial infarction; acute coronary syndrome; machine learning; troponin

Non-standard Abbreviations and Acronyms

AUC- area under the receiver-operating-characteristic curve

CI- confidence intervals

MI^3 - myocardial-ischemic-injury-index

NPV- negative predictive value

PPV- positive predictive value

STEMI- ST-segment elevation myocardial infarction

URL- upper reference limit

Clinical Perspective

What is new?

- In an international collaboration involving 11,011 patients from nine countries, we used machine learning to train and test a novel algorithm that estimates the probability of myocardial infarction for an individual patient.
- The MI³ algorithm generates a value that takes into consideration age, sex, paired cardiac troponin I concentrations, and rate of change in troponin concentration, to estimate the negative and positive predictive value for each patient value associated with these measures.
- This represents one of the first effective demonstrations of how machine learning could be used to guide clinical decision making in patients with suspected acute coronary syndrome.



What are the clinical implications?

- The MI³ algorithm is more versatile than existing algorithms as it is not dependent on fixed cardiac troponin thresholds, does not require serial testing to be performed at specific time points, and recognises that different healthcare systems have different priorities and tolerances of risk.
- Prospective studies are now required to evaluate patient outcomes and resource utilisation following implementation of the MI³ algorithm into clinical practice.

Introduction

The use of cardiac troponin testing in clinical practice is evolving rapidly.(1-3) Advances in assay analytical precision now permit quantification of cardiac troponin concentrations in the normal reference range with novel applications for early diagnostics and risk stratification in patients being assessed for possible acute coronary syndromes.(4) Historically, international guidelines have recommended the use of serial cardiac troponin measurements over 6 to 12 hours to define myocardial infarction in patients with a rise and/or fall in cardiac troponin concentration where at least one of the serial measured concentrations is above the 99th percentile.(5,6) However, some studies have challenged this approach suggesting lower thresholds can risk stratify patients to low, intermediate, or high risk of myocardial infarction when using high-sensitivity cardiac troponin assays.(7-15) These strategies have been incorporated into accelerated diagnostic pathways that advocate earlier troponin measurement at presentation and 1 to 3 hours later to facilitate prompt diagnosis and treatment in those with myocardial infarction or to expedite discharge in those without.(16,17)

The performance of these pathways varies across different populations, reflecting variation in cardiac troponin concentrations with age and sex.(13,18-21) This heterogeneity is not reflected in strategies which advocate the use of fixed thresholds for all patients, which only allow classification of patients as low, intermediate, or high risk and that do not reflect more subtle variations in risk.(6) Machine learning has been advocated as an objective, replicable, approach to integrate multiple quantitative variables to improve diagnostic accuracy.(22,23) Our aim was to test an algorithm, the myocardial-ischemic-injury-index (MI³), which had been trained by machine learning to estimate an individual patient's likelihood of myocardial infarction.

Methods

Transparency and Openness Promotion

The analysis code for this study is available on request. The algorithm is proprietary and subject to a patent application, however we are able to share this with researchers who agree to using the algorithm only for research purposes with a data sharing agreement.

Study design

This was an analysis of prospectively collected data from multiple centers to train and test the MI³ algorithm to predict the diagnosis of type 1 myocardial infarction. The training set comprised data from two cohorts (14,24) and the test set comprised data from seven cohorts of patients attending the Emergency Department with suspected myocardial infarction.(15,25-29) Training and testing are the nomenclature of machine learning, and are analogous to derivation and validation in studies of new diagnostic biomarkers.

MI³ incorporates age, sex, paired cardiac troponin I concentrations at presentation and at another early, yet flexible, time point, and rate of change of cardiac troponin I concentration. These variables (features) were selected *a priori* because they are a) objective and automatically captured from electronic hospital records, b) include serial measurements as recommended by guidelines, and c) associated with the diagnosis of myocardial infarction. MI³ computes a value from 0-100 (the MI³ value), which reflects the likelihood of a diagnosis of type 1 myocardial infarction for each patient during hospitalization (higher values indicate greater likelihood). The algorithm uses an embedded reference table to report for each individual patient estimates of sensitivity, negative predictive value (NPV), specificity and positive predictive value (PPV) of the diagnosis for a given MI³ value. MI³ was developed on the training data set by Abbott Diagnostics using a machine learning technique called gradient boosting. This technique iteratively trains a set of sequential weak learners (here decision trees) using the provided features to map onto the outcome (whether the patient was

or was not diagnosed with myocardial infarction). For further details regarding the gradient boosting method see also *Supplement*.(30) This is analogous to, but more complex than, the beta-coefficient weightings of a logistic regression model. The algorithm was provided to an independent statistician [JP] who evaluated its performance in the test set. JP had full access to all the test set data and takes responsibility for its integrity and the analysis.

We report according to relevant sections of the Transparent Reporting of multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD) statement.(31) The study was registered on the Australian New Zealand Clinical Trials Registry; ACTRN12616001441404.

Participants and cohorts

Patients presenting with symptoms suggestive of myocardial infarction in whom serial cardiac troponin measurements were obtained were included. Patients with ST-segment elevation myocardial infarction (STEMI) at presentation were excluded. Cohorts were identified for inclusion if they were prospective, had cardiac troponin I concentrations measured with the Abbott ARCHITECT_{STAT} high-sensitivity assay (Abbott Diagnostics, Chicago, Illinois) at presentation and at a second time point approximately 1-3 hours later (details in the *Supplement Materials*), the final diagnosis was adjudicated according to the Universal Definition of Myocardial Infarction (5), and ethical approval permitted sharing of individual patient-level data (*Supplemental Table 1*). All cohort studies were conducted in accordance with the Declaration of Helsinki and approved by the local research ethics committee or institutional review board. Written informed consent was obtained where this was required. All adjudication was completed prior to development of the MI³ algorithm.

Outcome definitions and adjudication

The primary outcome was the adjudicated diagnosis of type 1 myocardial infarction during the index admission. Although high-sensitivity cardiac troponin I was measured in all

patients, other cardiac troponin assays were used for adjudication in some cohorts

(Supplemental Table 1).

Algorithm development

A gradient boosting model was developed using pre-defined features (age, sex, paired cardiac troponin I concentrations at presentation and at another early, yet flexible, time point, and rate of change of cardiac troponin I concentration) to estimate the likelihood of a diagnosis of type 1 myocardial infarction. Once the model was trained, it was used to generate MI^3 values for each patient in the test sets.

Statistical analysis

Primary analysis

We describe algorithm performance in the test set by 1) visual inspection of a calibration curve to show how accurately MI^3 values estimate the likelihood of myocardial infarction, and 2) by the area under the receiver-operating-characteristic curve (AUC) to quantify how well the MI^3 values discriminated between those with and without myocardial infarction. Additionally, we compared diagnostic metric outputs from the algorithm (sensitivity, NPV, specificity, and PPV) for each patient with the metrics determined in the test set using each individual's MI^3 value as a threshold.

Secondary analyses

MI^3 is designed to be used as a continuous measure. However, we recognize that in this field most tools rely on thresholds to guide clinical decisions. Therefore, as illustrative examples of how an individual hospital may choose to use MI^3 , we demonstrate its diagnostic performance at two exemplar MI^3 value thresholds. First, we determined the MI^3 values with their 95% confidence interval that gave a sensitivity $\geq 99.0\%$ or NPV $\geq 99.5\%$ in the training set and assessed the accuracy of these example threshold values in the test set. These diagnostic criteria were pre-specified and based on an international survey of acceptable risk

by Emergency Department physicians (32), and prior prospective studies defining risk stratification thresholds for high-sensitivity cardiac troponin (14). Second, we determined the MI³ values that gave a specificity $\geq 90\%$ and a PPV $\geq 75\%$ in the training set based on consensus of the project steering committee, and assessed their performance in the test set. We used 1,000 bootstrapped samples to determine these MI³ thresholds and their 95% confidence intervals. All analyses were performed in R (Version 3.2.4: The R Foundation for Statistical Computing).

Additional analyses

Pre-specified subgroup analyses were performed by age, sex, co-morbidities (coronary artery disease, diabetes mellitus, hypertension, current smoking), time from symptom onset to first sample draw, time between tests, and the presence or absence of myocardial ischemia on the electrocardiogram. Performance of the algorithm was also evaluated for the outcomes of type 1 myocardial infarction within the next 30 days and for type 1 or 2 myocardial infarction on index admission. Finally, we compared the performance of MI³ using the example thresholds derived from our training set with the 99th percentile at any time-point, and the European Society of Cardiology (ESC) 0/1-hour and 0/3-hour pathways (6) for a diagnosis of type 1 myocardial infarction in our test set.

Results

The training set comprised 3,013 patients of whom 404 (13.4%) had a diagnosis of type 1 myocardial infarction. The set was predominantly male (63%) with a mean age of 62.4 years (*Table 1*). The test set comprised 7,998 patients, 62% male with a mean age of 58.8 years and mean time between samples of 2.5 hrs (standard deviation 1.2 hrs). Of these patients, 849 (10.6%) had a diagnosis of type 1 myocardial infarction. There were no missing data for any of the variables used in the training and testing sets. Patients in the testing set were younger,

less likely to have known coronary artery disease, but more likely to smoke cigarettes, have diabetes mellitus, dyslipidemia, or a family history of coronary artery disease than those in the training set.

Primary analysis: calibration and discrimination

The MI³ algorithm was well calibrated and discriminated between those with and without type 1 myocardial infarction (AUC 0.963 [95% CI 0.957-0.968], **Figure 1**). Compared to the MI³ output estimated metrics, the sensitivity was similar, specificity and NPV were slightly higher, and the PPV marginally lower in the test set (**Supplemental Figure 1**).

There was no difference in AUCs for those presenting within 3 hours (0.966 [0.959-0.973]) compared to later presenters (0.965 [0.959-0.972]) and no difference when stratifying by sex (men 0.962 [0.955-0.969] and women 0.962 [0.952-0.973], **Supplemental Figure 2**). The AUCs were higher in patients with no prior history of coronary artery disease, diabetes mellitus or hypertension compared to patients with these co-morbidities. The AUC was higher in younger compared to older patients and in those with no myocardial ischemia on the ECG compared to those with ischemia.

Secondary analysis: example diagnostic thresholds

The MI³ threshold values from the training set that corresponded to our pre-specified diagnostic performance metrics were 1.6 (0.9-3.0; sensitivity $\geq 99.0\%$), 3.1 (1.7-4.7; NPV $\geq 99.5\%$), 17.2 (13.8-21; specificity $\geq 90.0\%$), and 49.7 (36.6-60.0; PPV $\geq 75\%$, **Table 2**). In the test set, MI³ values of 1.6 and 3.1 gave a sensitivity of 97.8% (96.7-98.7%) and a NPV of 99.4% (99.2-99.6%), respectively. MI³ values of 17.2 and 49.7 gave a specificity of 91.7% (91.1-92.3%) and a PPV of 71.8% (68.9-75.0%) (**Table 2**).

If, for example, patients with MI³ values < 1.6 were to be classified as low-risk of myocardial infarction then this threshold would identify 69.4% (68.4%-70.4%) as low-risk of whom 0.5% (0.3%-0.7%) would be false negatives. If patients with MI³ values ≥ 49.7 were

classified as high-risk then this threshold would identify 10.6% (10.0%-11.2%) as high risk of whom 28.1% (25.1% to 31%) would be false positives (*Supplemental Table 2*). These two exemplar thresholds were used for all subsequent analysis. The MI³ threshold value of 1.6 performed similarly across all subgroups including those who presented within 3 hours of symptom onset (sensitivity 98.7% [97.3%-100%] and NPV 99.8% [99.7%-100%], *Supplemental Figure 3*). The MI³ threshold value of 49.7 also performed similarly across most groups with the exception of sex and time from symptom onset, where the PPV was lower in women than men, and in those presenting within 3 hours compared-those presenting more than 3 hours from symptom onset (*Supplemental Figure 4*).

Secondary analysis: type 1 myocardial infarction within 30 days

In addition to the 849 (10.6%) patients with type 1 myocardial infarction during the initial hospitalization there were 23 (2.9%) with myocardial infarction following discharge within the next 30 days. The MI³ value discriminated between those patients with and without type 1 myocardial infarction within the next 30 days with an AUC of 0.957 (0.951-0.963). Threshold values of 1.6 and 49.7 gave a sensitivity of 96.6% (95.3-97.8%) and PPV of 71.9% (69.0-74.9%), respectively (*Supplemental Table 2*).

Secondary analysis: type 1 or 2 myocardial infarction on presentation

In addition to the 849 (10.6%) patients with type 1 myocardial infarction there were 216 (2.7%) with type 2 myocardial infarction during the initial hospitalization. The MI³ value discriminated between those with and without type 1 or type 2 myocardial infarction with an AUC of 0.963 (0.957-0.968, *Supplemental Figure 5*). The example low-risk MI³ threshold, 1.6, had a sensitivity of 97.4% (96.3-99.5%), and an NPV of 99.5% (99.3-99.7%), identifying 69.5% of patients as low-risk. The example high-risk threshold, 49.7, had a specificity of 97.7% (97.3-98.0%), and a PPV of 80.8% (78.1-83.5%), identifying 10.6% of patients as high-risk (*Supplemental Table 3*).

Secondary analysis: comparison with other recommended diagnostic strategies

In all 7,998 patients in the test set, the 99th percentile upper reference limit at any time-point identified 6,473 (80.9%) patients as low-risk (NPV 98.6% [98.3-98.9%], sensitivity 89.6% [87.4-91.6%]) and 1,525 (19.1%) as high-risk (PPV 49.9% [47.4-52.4%], specificity 89.3% [88.6-90.0%]). A total of 1,652 patients (21%) from the test set were eligible for inclusion in the analysis of the ESC 0/3-hour pathway with ≥ 2.5 hours between serial samples (**Figure 2**). This pathway identified 86.7% (1,433/1,652) of patients as low-risk (NPV 98.5% [97.8-99.1%], sensitivity 82.5% [74.5-88.8%]) with 21 missed events, and 13.3% (219/1,652) of patients as high-risk (PPV 45.2% [38.5-52.1%], specificity 92.2% [90.7-93.5%]). In the same patient group, $MI^3 < 1.6$ or ≥ 49.7 , identified 70.7% (1,168/1,652) of patients as low-risk (NPV 99.9% [99.5-100%], sensitivity 99.2% [95.4-100%]) with one missed event, and 9.0% (149/1,652) of patients as high-risk (specificity 95.9% [94.8-96.8%], PPV 57.7 [49.4-65.8%]) respectively. Use of these thresholds identified 20.3% (335/1,652) of patients as intermediate risk (MI^3 1.6 to 49.6) of whom 33 patients had an event.

Only 336 patients (4%) from the test set were eligible for inclusion in the analysis of the ESC 1-hour pathway with > 0.5 hour but ≤ 1.5 hours between serial samples (**Figure 3**).

This pathway identified 54.3% (183/336) of patients as low-risk (NPV 100% [98.0-100%], sensitivity 100% [93.2-100%]) with no missed events, and 18.3% (61/336) of patients as high-risk (PPV 75.4% [62.7-85.5%], specificity 94.7% [91.4-97.0%]), and 27.4% (92/336) of patients as intermediate risk of whom 6 had an event. Here $MI^3 < 1.6$ or ≥ 49.7 , identified 64.3% (216/336) of patients as low-risk (NPV 100% [98.3-100%], sensitivity 100% [93.2-100%]) with no missed events, and 14.9% (50/336) of patients as high-risk (specificity 97.2% [94.5-98.8%], PPV 84.0 [70.9-92.8%]) respectively. Use of these thresholds identified 20.8% (70/336) of patients as intermediate risk (MI^3 1.6 to 49.6) of whom 10 patients had an event.

Discussion

In a large, international, multi-centre study of over 11,000 patients with suspected myocardial infarction, we used machine learning to train and test a novel decision tool that incorporates simple, objective variables known to be associated with the diagnosis of myocardial infarction to accurately predict the likelihood of a diagnosis of myocardial infarction. The algorithm was well calibrated, and the overall diagnostic performance was identical in both training and test data sets. This study has several unique and important characteristics.

First, this technique provides an *individualised* and precise assessment of risk by using age, sex, and paired high-sensitivity cardiac troponin I concentrations and allows for the complex and non-linear ways in which these variables may interact. This is in contrast to contemporary algorithms in clinical use which are based on fixed time-points for sampling, fixed troponin thresholds, and do not account for any interaction between input variables.

One exception is the Troponin only-Manchester Acute Coronary Syndrome rule which uses a logistic regression model to provide risk estimates for myocardial infarction, incorporating age, sex, multiple clinical variables, and a single high-sensitivity troponin T measurement.(33,34) They reported an AUC of 0.90. This model has a number of strengths, but does not take into account dynamic interaction between variables and has to-date only been assessed using thresholds allocating patients into one of four risk categories.(35,36)

Second, MI³ recognises the importance of both the magnitude and rate of change in cardiac troponin concentration for the diagnosis of myocardial infarction, without applying a fixed absolute or percentage change in concentration, or mandating specific time-points for serial testing. MI³ performed well across a range of time intervals between sampling. This enhances its transferability because variation in the sample timing is commonplace in busy Emergency Departments. Importantly, there was no difference in performance when stratified by time from symptom onset, which means that unlike some other approaches, MI³ can be

applied in those patients presenting early (i.e. within 3 hours of symptom onset). This is important because early presenters are a sizable subset of patients (34% of patients in the test cohort) and also because the time of symptom onset is often uncertain.

Third, the large cohort size and number of patients with type 1 myocardial infarction allowed for a robust analysis of sub-groups.

Fourth, at the example thresholds MI^3 was better than the 99th percentile alone or the ESC 0/3-hour pathway at identifying low-risk and high-risk patients. Consistent with previous reports, (18,20,21) both of these approaches gave a low diagnostic sensitivity and poor positive predictive value despite being widely used in clinical practice (37). MI^3 compared well to the ESC 0/1-hour pathway, with the primary advantage of MI^3 being flexibility in the timing of serial testing, and the simplicity of using probabilities rather than multiple thresholds to stratify risk in individual patients.

Fifth, the algorithm performed well even when type 2 myocardial infarction was included as an outcome event.

Finally, MI^3 provides guidance on high-risk patients reporting the PPV and specificity for type 1 myocardial infarction that could be used to initiate earlier treatment or expedite cardiology consultation. Previous attempts to optimise specificity for myocardial infarction have used absolute or relative changes in troponin concentrations to differentiate from chronic myocardial injury or have recommended thresholds well above the 99th percentile.(16,17,19,38) MI^3 avoids the need to decide on the use of relative or absolute changes or on a threshold for change *a priori*. In essence, the developed machine-learning model uses the rate of concentration change, patient age, and sex to decide on the weighting of relative and absolute troponin concentration changes.

These features are presented to the gradient boosting algorithm which outputs the MI^3 value which can provide a clinical decision support tool for the assessment of all patients

with suspected myocardial infarction. The clinical decision support tool also reports the diagnostic parameters associated with the calculated MI^3 value. These diagnostic parameters cannot be derived for an individual patient, and therefore the tool uses an embedded reference table to report estimates of sensitivity, specificity and negative and positive predictive values from our training set alongside the calculated MI^3 value. This tool is easy to implement in practice, is objective, because it does not rely on potentially inconsistent assessment of symptoms or patient history, and provides accurate estimates of the likelihood of myocardial infarction for each individual patient to aid clinical decision making.

We are aware of two attempts to use machine learning in this field, both using an artificial neural network. Green and colleagues in 2005 developed a model for predicting acute coronary syndrome based on demographics, prior history, symptom duration, and diastolic blood pressure.⁽³⁹⁾ This model had an AUC of 0.778. In 2007, a model based on serial measurements of myoglobin and contemporary troponin I concentrations, either alone or in combination with their rate of change was evaluated in 310 patients.⁽⁴⁰⁾ The output of this algorithm was dichotomised for diagnostic purposes, such that it had a sensitivity of 99%. These pioneering efforts were undertaken before the availability of high-sensitivity troponins and prior to the availability of multiple high-quality cohorts that we have been able to draw on.

Finally, most studies that have developed or assessed strategies to risk stratify patients with suspected myocardial infarction have enrolled a limited number of patients with a small number of events resulting in limited precision and therefore limited external generalizability. In contrast, MI^3 has been trained and tested in a population that includes patients from nine studies across multiple geographic regions with over 1,250 events, and significant variation in the prevalence of disease, suggesting this approach is generalizable and could be used in any healthcare setting worldwide. Furthermore, rather than being an inflexible diagnostic

strategy with multiple set thresholds, the MI³ algorithm is a dynamic tool that in the future could be retrained for individual health care settings depending on disease prevalence and diagnostic priorities to facilitate healthcare delivery.

From a clinical perspective, each patient will have an MI³ value that takes into consideration their own age and sex and measured cardiac troponin concentrations. This approach is distinct from previous algorithms. Whilst early diagnostic pathways, such as the ESC 0/1-hour pathway, identify groups of patients with high negative and moderate positive predictive value, they do not report these metrics or derive them for individual patients. This is a particularly important limitation for the one third of patients triaged to the observe zone, for whom the 0/1 hour pathway provides no guidance. The MI³ algorithm is also more versatile because it does not require serial testing to be performed at specific timepoints, and recognises that different healthcare systems have different priorities and tolerances of risk. The MI³ algorithm allows implementation to be tailored accordingly. For example, in some more conservative institutions triage to out-patient investigation might be acceptable only where the NPV is >99.8% (false negative rate of 1 in 1,000; see *Figure 4*). Similarly, a cardiology consultation in the Emergency Department might be triggered if the PPV of myocardial infarction is >60%, but direct transfer to the cardiac cath lab only considered where the PPV is >80%. Prospective studies are now required to evaluate patient outcomes and resource utilisation following implementation of this algorithm into clinical practice.

Limitations

Not all sites used the high-sensitivity cardiac troponin I assay to adjudicate the diagnosis of myocardial infarction, with several sites using either contemporary assays or high-sensitivity cardiac troponin T assays. This should reduce performance of the algorithm, yet we observed little heterogeneity across the cohorts. The choice of the high-risk example threshold was pre-specified on the basis of the PPV which is dependent on prevalence and may explain some of

the heterogeneity between cohorts. Institutions could choose their own threshold based on the local prevalence of myocardial infarction or clinical priorities. Factors such as acceptability of risk, availability of inpatient beds or outpatient services, or the need to transfer patients for coronary angiography may influence the adoption of local thresholds. In contrast to the ESC 0/1- and 0/3-hour pathways, the primary purpose of MI³ is to assess risk in an individual patient. There are some limitations to our comparison with both pathways due to the requirement for precise sample intervals, which reduced the proportion of patients eligible for inclusion to 21% and 4% of our test cohort respectively. Prospective studies are needed to better understand the advantages and disadvantages of pathways based on individual probabilities rather than fixed thresholds. Finally, it is conventional in machine learning to use a larger data set to train and smaller data set to test out of practical considerations to develop a robust algorithmic model. We have used the smaller data set to train because it was available before the other data sets and because it was already a large data set. Machine learning also has the capability to utilize a large number of features which could include aspects of patient history and clinical symptoms. These variables were not included because our intention was to develop a tool that only uses variables that are objective and always available to ensure our algorithm can be applied widely in clinical practice. Finally, we focus on the training and testing of the MI³ algorithm and do not report a comparison with other linear regression or machine learning methods here.

Conclusion

The MI³ clinical decision support tool incorporates simple and objective variables including age, sex and serial cardiac troponin I concentrations measured using a high-sensitivity assay to rapidly estimate risk of myocardial infarction. It can be used to individualise the risk assessment of patients with suspected myocardial infarction or to categorize patients into low- or high-risk groups.

Disclosures

Abbott Diagnostics applied the supervised machine learning technique ‘boosting’ to train the MI³ algorithm [AB, SD, JH] and provided a description of the training process which is supplied in the supplement. Abbott Diagnostics have filed for an international patent on the use of the MI³ algorithm (International Publication Number WO2017/173353A1). Abbott Diagnostics did not participate in data collection in any of the contributing studies, nor in the testing of algorithm performance. Abbott Diagnostics provided some funding to support two group meetings and data extraction. The Abbott employees listed [AB, SD, JH] contributed as co-authors to the drafting of the manuscript. NLM has acted as a consultant for Abbott Diagnostics, Siemens Healthineers, Singulex and LumiraDx.



Sources of Funding

MPT is supported by a Clinical Practitioners Research Fellowship by the Health Research Council of New Zealand. JWP is supported by a Senior Research Fellowship in Acute Care from the Emergency Care Foundation, Canterbury Medical Research Foundation, and Canterbury District Health Board. ASVS and NLM are supported by an Intermediate Clinical Research Fellowship (FS/19/17/34172) and the Butler Senior Clinical Research Fellowship (FS/16/14/32023) from the British Heart Foundation (BHF), respectively.

The study was also supported by a Research Excellence Award from the British Heart Foundation (RE/18/5/34216) and by Health Data Research UK which receives its funding from HDR UK Ltd (HDR-5012) funded by the UK Medical Research Council, Engineering and Physical Sciences Research Council, Economic and Social Research Council, Department of Health and Social Care (England), Chief Scientist Office of the Scottish Government Health and Social Care Directorates, Health and Social Care Research and Development Division (Welsh Government), Public Health Agency (Northern Ireland), BHF

and the Wellcome Trust. The funders had no role in the study and the decision to submit this work to be considered for publication.

References

1. Sandoval Y, Smith SW, Apple FS. Present and future of cardiac troponin in clinical practice: a paradigm shift to high-sensitivity assays. *Am J Med.* 2016;129:354–365.
2. Hollander JE, Than MP, Mueller C. State-of-the-Art Evaluation of Emergency Department Patients Presenting With Potential Acute Coronary Syndromes. *Circulation.* 2016;134:547–564.
3. Westermann D, Neumann JT, Sørensen NA, Blankenberg S. High-sensitivity assays for troponin in patients with cardiac disease. *Nat Rev Cardiol.* 2017;386:1.
4. Apple FS, Sandoval Y, Jaffe AS, Ordonez-Llanos J, for the IFCC Task Force on Clinical Applications of Cardiac Bio-Markers. Cardiac Troponin Assays: Guide to Understanding Analytical Characteristics and Their Impact on Clinical Care. *Clin Chem.* 2017;63:73–81.
5. Thygesen K, Alpert JS, Jaffe AS, Simoons ML, Chaitman BR. ESC/ACCF/AHA/WHF Expert Consensus Document: Third universal definition of myocardial infarction. *Circulation.* 2012;126:2020–2035.
6. Roffi M, Patrono C, Collet J-P, Mueller C, Valgimigli M, Andreotti F, Bax JJ, Borger MA, Brotons C, Chew DP, et al. 2015 ESC Guidelines for the management of acute coronary syndromes in patients presenting without persistent ST-segment elevation: Task Force for the Management of Acute Coronary Syndromes in Patients Presenting without Persistent ST-Segment Elevation of the European Society of Cardiology (ESC). *Eur Heart J.* 2016;37:267–315.
7. Body R, Carley S, McDowell G, Jaffe AS, France M, Cruickshank K, Wibberley C, Nuttall M, Mackway-Jones K. Rapid exclusion of acute myocardial infarction in patients with undetectable troponin using a high-sensitivity assay. *J Am Coll Cardiol.* 2011;58:1332–1339.
8. Rubini Gimenez M, Hoeller R, Reichlin T, Zellweger C, Twerenbold R, Reiter M, Moehring B, Wildi K, Mosimann T, Mueller M, et al. Rapid rule out of acute myocardial infarction using undetectable levels of high-sensitivity cardiac troponin. *Int J Cardiol.* 2013;168:3896–3901.
9. Bandstein N, Ljung R, Johansson M, Holzmann MJ. Undetectable High-Sensitivity Cardiac Troponin T Level in the Emergency Department and Risk of Myocardial Infarction. *J Am Coll Cardiol.* 2014;63:2569–2578.
10. Carlton E, Greenslade JH, Cullen LA, Body R, Than MP, Pickering JW, Aldous S, Carley S, Hammett C, Kendall J, et al. Evaluation of High-Sensitivity Cardiac Troponin I Levels in Patients With Suspected Acute Coronary Syndrome. *JAMA Cardiol.* 2016;1:405–412.
11. Thelin J, Melander O, Öhlin B. Early rule-out of acute coronary syndrome using undetectable levels of high sensitivity troponin T. *Eur Heart J Acute Cardiovasc Care.* 2015;4:403–409.
12. Body R, Mueller C, Giannitsis E, Christ M, Ordóñez-Llanos J, de Filippi CR, Nowak R, Panteghini M, Jernberg T, Plebani M, et al. The Use of Very Low Concentrations of High-sensitivity Troponin T to Rule Out Acute Myocardial Infarction Using a Single Blood Test. *Acad Emerg Med.* 2016;23:1004–1013.

13. Pickering JW, Than MP, Cullen LA, Aldous S, Avest ET, Body R, Carlton EW, Collinson P, Dupuy AM, Ekelund U, et al. Rapid Rule-out of Acute Myocardial Infarction With a Single High-Sensitivity Cardiac Troponin T Measurement Below the Limit of Detection: A Collaborative Meta-analysis. *Ann Intern Med.* 2017;166:715–724.
14. Shah ASV, Anand A, Sandoval Y, Lee KK, Smith SW, Adamson PD, Chapman AR, Langdon T, Sandeman D, Vaswani A, et al; on behalf of the High-STEACS investigators. High-sensitivity cardiac troponin I at presentation in patients with suspected acute coronary syndrome: a cohort study. *Lancet.* 2016;386:2481–2488.
15. Sandoval Y, Smith SW, Love SA, Sexter A, Schulz K, Apple FS. Single high-sensitivity cardiac troponin I to rule-out acute myocardial infarction. *Am J Med.* 2017;130:1076-1083.
16. Rubini Gimenez M, Twerenbold R, Jaeger C, Schindler C, Puelacher C, Wildi K, Reichlin T, Haaf P, Merk S, Honegger U, et al. One-hour rule-in and rule-out of acute myocardial infarction using high-sensitivity cardiac troponin I. *Am J Med.* 2015;128:861–864.
17. Reichlin T, Schindler C, Drexler B, Twerenbold R, Reiter M, Zellweger C, Moehring B, Ziller R, Hoeller R, Rubini Gimenez M, et al. One-hour rule-out and rule-in of acute myocardial infarction using high-sensitivity cardiac troponin T. *Arch Intern Med.* 2012;172:1211–1218.
18. Pickering JW, Greenslade JH, Cullen LA, Flaws D, Parsonage WA, George P, Worster A, Kavsak PA, Than MP. Validation of presentation and 3 h high-sensitivity troponin to rule-in and rule-out acute myocardial infarction. *Heart.* 2016;102:1270–1278.
19. Pickering JW, Greenslade JH, Cullen LA, Flaws D, Parsonage WA, Aldous S, Worster A, Kavsak PA, and Than MP. Assessment of the European Society of Cardiology 0 Hour/1 Hour Algorithm to Rule Out and Rule In Acute Myocardial Infarction. *Circulation.* 2016;134:1532–1541.
20. Parsonage WA, Mueller C, Greenslade JH, Wildi K, Pickering JW, Than MP, Aldous S, Boeddinghaus J, Hammett CJ, Hawkins T, et al. Validation of NICE diagnostic guidance for rule out of myocardial infarction using high-sensitivity troponin tests. *Heart.* 2016;102:1279–1286.
21. Chapman AR, Anand A, Boeddinghaus J, Ferry AV, Sandeman D, Adamson PD, Andrews J, Tan S, Cheng SF, D'Souza M, et al. Comparison of the Efficacy and Safety of Early Rule-Out Pathways for Acute Myocardial Infarction. *Circulation.* 2017;135:1586–1596.
22. Obermeyer Z, Emanuel EJ. Predicting the Future - Big Data, Machine Learning, and Clinical Medicine. *N Engl J Med.* 2016;375:1216–1219.
23. Krittanawong C, Zhang H, Wang Z, Aydar M, Kitai T. Artificial Intelligence in Precision Cardiovascular Medicine. *J Am Coll Cardiol.* 2017;69:2657–2664.
24. Neumann JT, Sörensen NA, Schwemer T, Ojeda F, Bourry R, Sciacca V, Schaefer S, Waldeyer C, Sinning C, Renné T, et al. Diagnosis of Myocardial Infarction Using a High-Sensitivity Troponin I 1-Hour Algorithm. *JAMA Cardiol.* 2016;1:397.
25. Reichlin T, Hochholzer W, Bassetti S, Steuer S, Stelzig C, Hartwiger S, Biedert, Nora Schaub S, Buerge C, et al. Early diagnosis of myocardial infarction with sensitive cardiac troponin assays. *N Engl J Med.* 2009;361:858–867.
26. Than MP, Cullen LA, Aldous S, Parsonage WA, Reid CM, Greenslade JH, Flaws D, Hammett CJ, Beam DM, Ardagh MW, et al. 2-Hour accelerated diagnostic protocol to assess patients with chest pain symptoms using contemporary troponins as the only biomarker: the ADAPT trial. *J Am Coll Cardiol.* 2012;59:2091–2098.
27. Than MP, Aldous S, Lord SJ, Goodacre S, Frampton CMA, Troughton R, George P,

- Florkowski CM, Ardagh M, Smyth D, et al. A 2-hour diagnostic protocol for possible cardiac chest pain in the emergency department: a randomized clinical trial. *JAMA Intern Med.* 2014;174:51–58.
28. Than MP, Pickering JW, Aldous SJ, Cullen LA, Frampton CMA, Peacock WF, Jaffe AS, Goodacre SW, Richards AM, Ardagh MW, et al. Effectiveness of EDACS Versus ADAPT Accelerated Diagnostic Pathways for Chest Pain: A Pragmatic Randomized Controlled Trial Embedded Within Practice. *Ann Emerg Med.* 2016;68:93–102.
 29. Cullen LA, Greenslade JH, Hawkins T, Hammett C, O’Kane S, Ryan K, Parker K, Schluter J, Dalton E, Brown AFT, et al. Improved Assessment of Chest pain Trial (IMPACT): assessing patients with possible acute coronary syndromes. *Med J Aust.* 2017;207:195–200.
 30. Friedman J.H. Greedy Function Approximation: A Gradient Boosting Machine. *Ann Stats.* 2001;29:1189–1232.
 31. Collins GS, Reitsma JB, Altman DG, Moons KGM. Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis Or Diagnosis (TRIPOD): the TRIPOD statement. *J Clin Epidemiol.* 2015;68:112–121.
 32. Than MP, Herbert M, Flaws D, Cullen LA, Hess E, Hollander JE, Diercks D, Ardagh MW, Kline JA, Munro Z, et al. What is an acceptable risk of major adverse cardiac event in chest pain patients soon after discharge from the Emergency Department? A clinical survey. *Int J Cardiol.* 2013;166:752–754.
 33. Body R, Carley S, McDowell G, Pemberton P, Burrows G, Cook G, Lewis PS, Smith A, and Mackway-Jones K. The Manchester Acute Coronary Syndromes (MACS) decision rule for suspected cardiac chest pain: derivation and external validation. *Heart.* 2014;100:1462–1468.
 34. Body R, Carlton E, Sperrin M, Lewis PS, Burrows G, Carley S, McDowell G, Buchan I, Greaves K, Mackway-Jones K. Troponin-only Manchester Acute Coronary Syndromes (T-MACS) decision aid: single biomarker re-derivation and external validation in three cohorts. *Emerg Med J.* 2017;34:349–356.
 35. Greenslade JH, Nayer R, Parsonage WA, Doig S, Young J, Pickering JW, Than MP, Hammett C, and Cullen LA. Validating the Manchester Acute Coronary Syndromes (MACS) and Troponin-only Manchester Acute Coronary Syndromes (T-MACS) rules for the prediction of acute myocardial infarction in patients presenting to the emergency department with chest pain. *Emerg Med J.* 2017;34:517–523.
 36. Body R, Boachie C, McConnachie A, Carley S, van den Berg P, Lecky FE. Feasibility of the Manchester Acute Coronary Syndromes (MACS) decision rule to safely reduce unnecessary hospital admissions: a pilot randomised controlled trial. *Emerg Med J.* 2017;34:586–592.
 37. Shah ASV, Anand A, Strachan FE, Ferry AV, Lee KK, Chapman AR, Sandeman D, Stables CL, Adamson PD, Andrews JPM, et al; High-STEACS Investigators. High-sensitivity troponin in the evaluation of patients with suspected acute coronary syndrome: a stepped-wedge, cluster-randomised controlled trial. *Lancet.* 2018;392:919–928.
 38. Mueller C, Giannitsis E, Christ M, Ordóñez-Llanos J, DeFilippi C, McCord J, Body R, Panteghini M, Jernberg T, Plebani M, et al. Multicenter Evaluation of a 0-Hour/1-Hour Algorithm in the Diagnosis of Myocardial Infarction With High-Sensitivity Cardiac Troponin T. *Ann Emerg Med.* 2016;68:76–87.
 39. Green M, Björk J, Lundager Hansen J, Ekelund U, Edenbrandt L, Ohlsson M. Detection of acute coronary syndromes in chest pain patients using neural network ensembles. In J. Fonseca (Ed.), *Second International Conference on Computational Intelligence in Medicine and Healthcare*. Published 2005; 182–187.

40. Eggers KM, Ellenius J, Dellborg M, Groth T, Oldgren J, Swahn E, Lindhal B. Artificial neural network algorithms for early diagnosis of acute myocardial infarction and prediction of infarct size in chest pain patients. *Int J Cardiol.* 2007;114:366–374.



Circulation

Table 1. Baseline characteristics of training and testing sets

Variable	Training (<i>n</i> =3,013)	Testing (<i>n</i> =7,998)
Age, years	62.4 ± 14.9	58.8 ± 15.1
Sex, female (%)	1,113 (36.9)	3,058 (38.2)
History of coronary artery disease (%)	1,143 (37.9)	2,143 (26.8)
History of myocardial infarction (%)	630 (21.1)	1,599 (20.0)
Diabetes mellitus (%)	436 (14.6)	1,494 (18.7)
Dyslipidemia (%)	1,232 (41.3)	3,835 (47.9)
Hypertension (%)	1,705 (57.2)	4,570 (57.1)
Current smoker (%)	648 (21.9)	1,957 (24.7)
Family history of coronary artery disease (%)	986 (33.9)	3,197 (40.6)
Symptom onset to blood draw ≤3h (%)	1,065 (33.0)	3,613 (38.5)
Time between blood draws, hours	1.2 [1.0-2.5] 2.0 ± 1.9	2.2[2.0-2.6] 2.5 ± 1.2

Values are mean ± standard deviation; median (lower quartile-upper quartile); n (%).



Circulation

Table 2. Performance in the testing set of example MI³ threshold values from the training set

TRAINING				TESTING						
Statistical thresholds	MI ³ threshold (95%CI)	Proportion low risk, %	Proportion high risk, %	MI ³ threshold	Sensitivity	NPV	Specificity	PPV	Proportion low risk, %	Proportion high risk, %
Sensitivity ≥99.0%	1.6 (0.9-3.0)	59.8% (57.9-61.5%)	-	1.6	97.8% (96.8-98.7%)	99.7% (99.5-99.8%)	77.4% (76.4-78.4%)	33.9% (32.0-35.8%)	69.4% (68.4-70.4%)	-
NPV ≥99.5%	3.1 (1.7-4.7)	68.6% (66.9-70.2%)	-	3.1	95.8% (94.3-97.0%)	99.4% (99.2-99.6%)	83.6% (82.7-84.4%)	40.9% (38.9-43.1%)	75.2% (74.2-76.1%)	-
Specificity ≥90%	17.2 (13.8-21.2)	-	20.7% (19.2-22.1%)	17.2	88.7% (86.7-90.8%)	98.6% (98.3-98.8%)	91.7% (91.1-92.3%)	56.0% (53.3-58.5%)	-	16.8% (16.0-17.7%)
PPV of ≥75%	49.7 (36.6-60.0)	-	12.8% (11.6-14.0%)	49.7	71.5% (68.4-74.3%)	96.6% (96.2-97.0%)	96.7% (96.3-97.1%)	71.7% (68.9-74.8%)	-	10.6% (9.9-11.2%)

NPV = negative predictive value, PPV = positive predictive value

Note: Sensitivity and NPV thresholds divide the population into low-risk and not low-risk groups (i.e. they do not determine a high-risk group). Similarly, specificity and PPV thresholds divide the population into high-risk and not high-risk groups (i.e. they do not determine a low-risk group).

Circulation

Figure Legends

Figure 1. Calibration and discrimination of the MI³ algorithm. (a) Calibration of the MI³ algorithm with the observed proportion of patients with type 1 myocardial infarction in the test data set. Each point represents 100 patients. The dashed lines represent perfect calibration. (b) Receiver-operating-characteristic curve showing discrimination of the MI³ algorithm in the test data set. Some MI³ values shown for illustrative purposes only.

Figure 2. Performance of MI³ algorithm compared to the ESC 3-hour algorithm

Performance of MI³ at example thresholds (a) and the European Society of Cardiology (ESC) 3-hour algorithm (b) for high-sensitivity cardiac troponin I (hs-cTnI) in 1,652 patients with ≥ 2.5 hours between serial samples in the test set. URL = upper reference limit

Figure 3. Performance of MI³ algorithm compared to the ESC 1-hour algorithm

Performance of MI³ at example thresholds (a) and the European Society of Cardiology (ESC) 1-hour algorithm (b) for high-sensitivity cardiac troponin I (hs-cTnI) in 336 patients with >0.5 hour but ≤ 1.5 hours between serial samples in the test set.

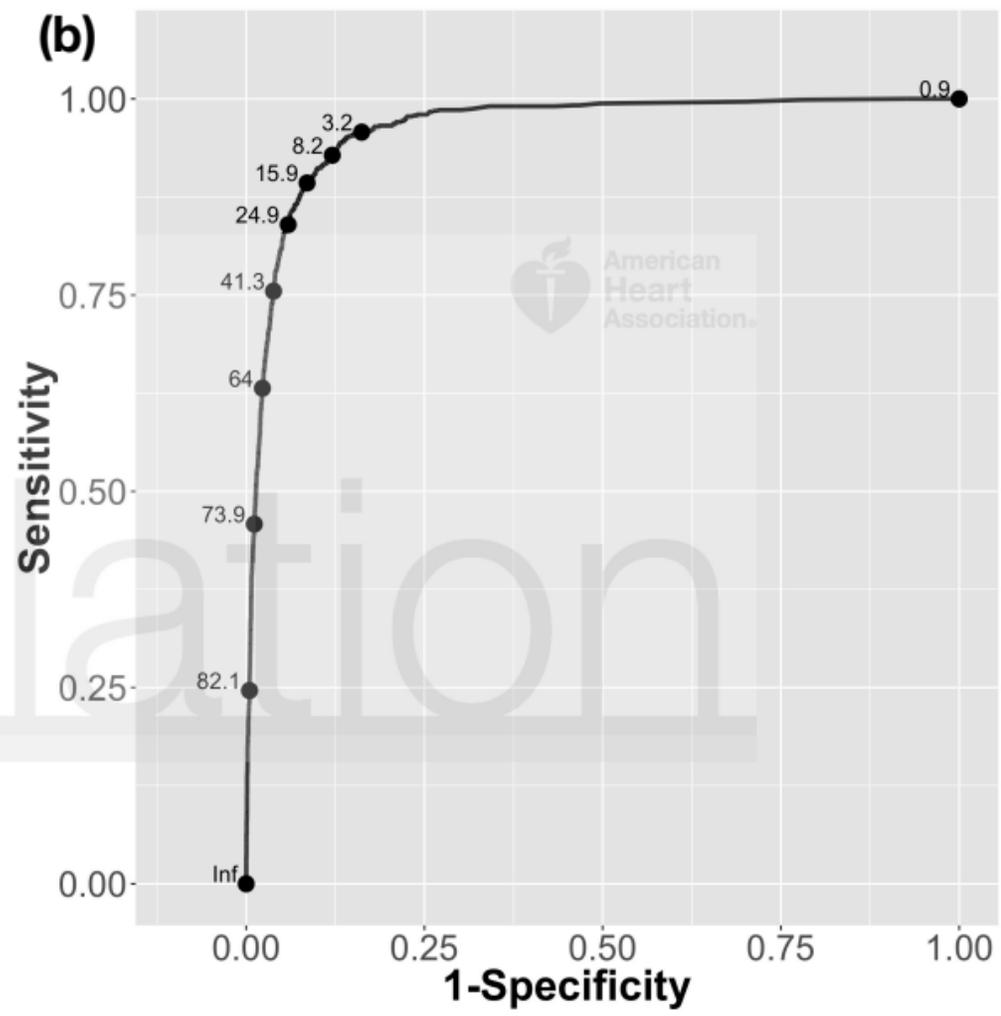
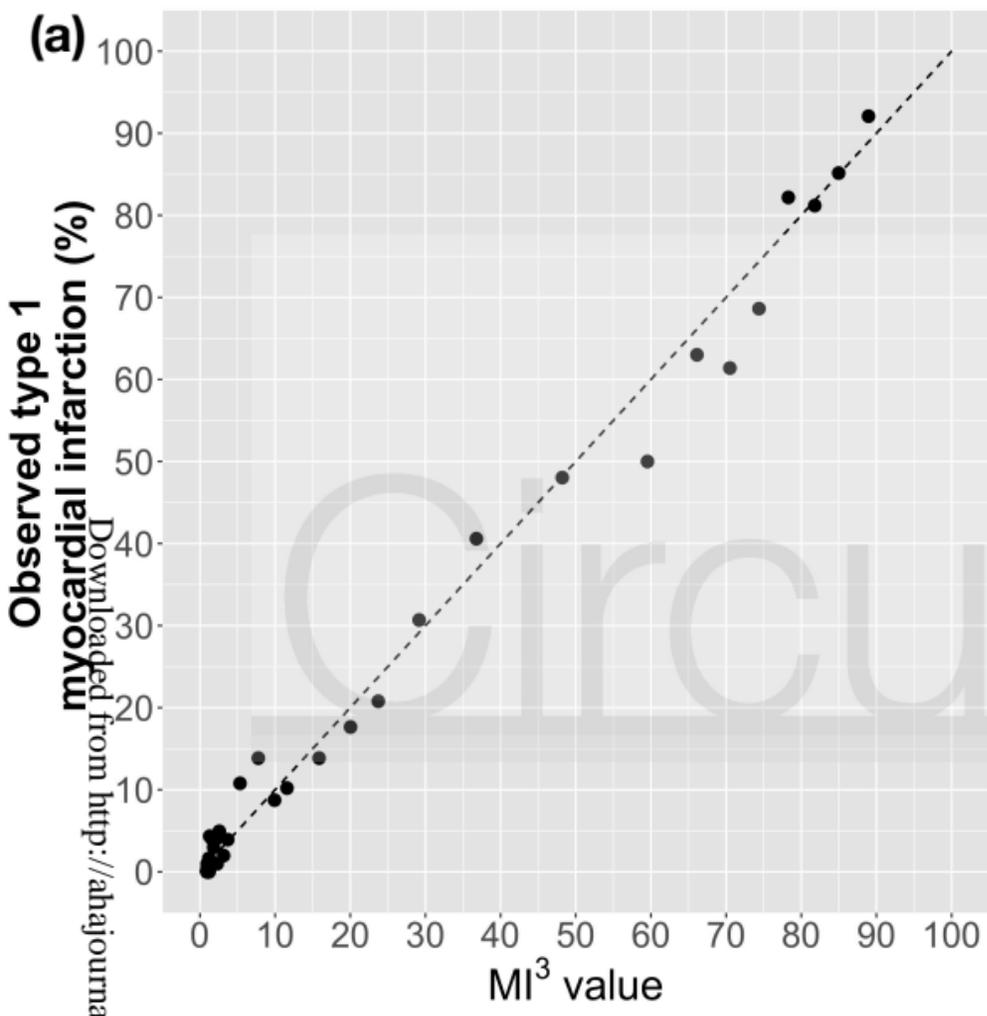
Figure 4. MI³ clinical decision support tool to estimate the likelihood of myocardial infarction for individual patients

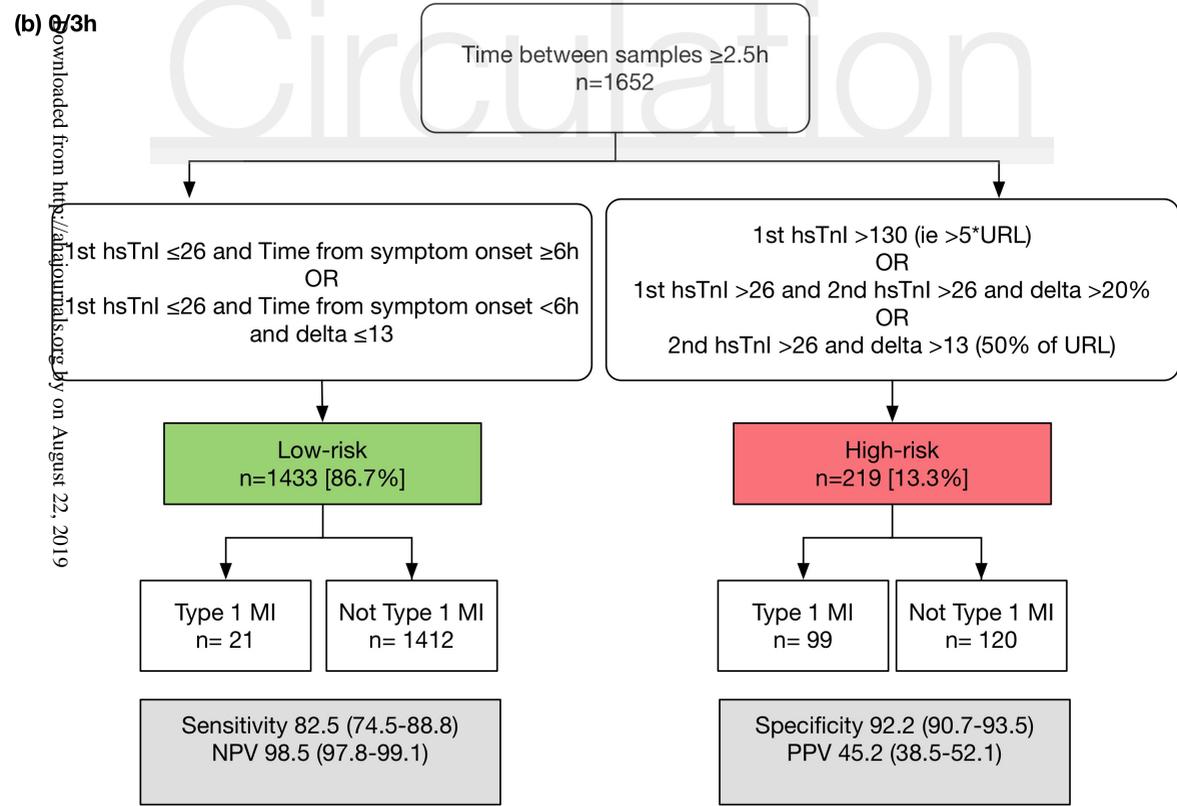
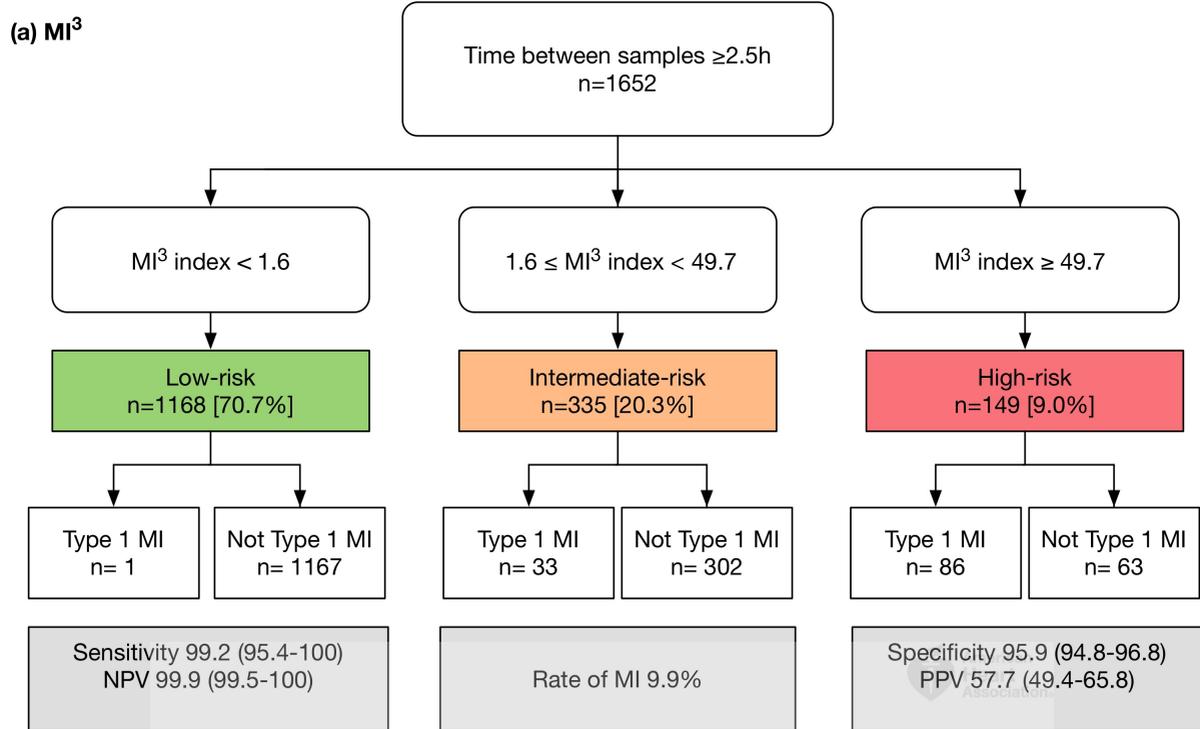
A mock-up of how the MI³ algorithm may be presented to physicians and patients. The top row illustrates a low-risk patient, the middle row an intermediate risk and the bottom row a high-risk patient based on the example MI³ values of 0.9, 8.2 and 68.2 respectively. The screens on the left are for data input and return the MI³ value and estimated diagnostic

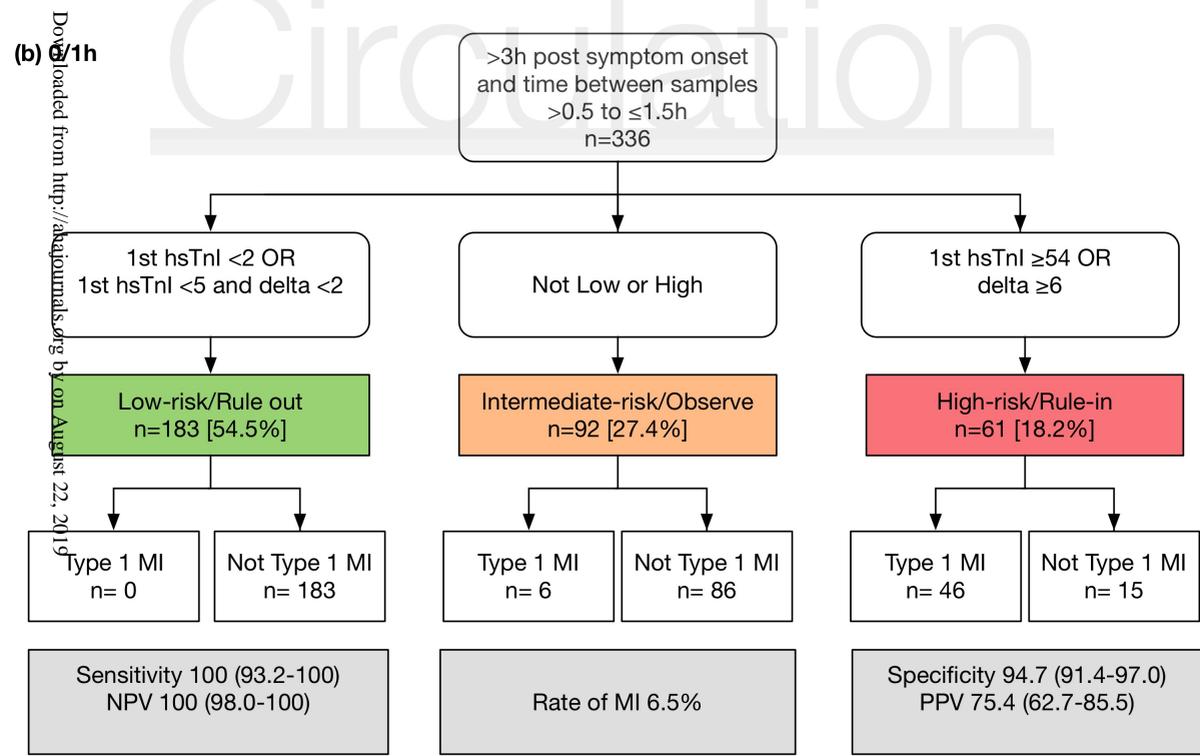
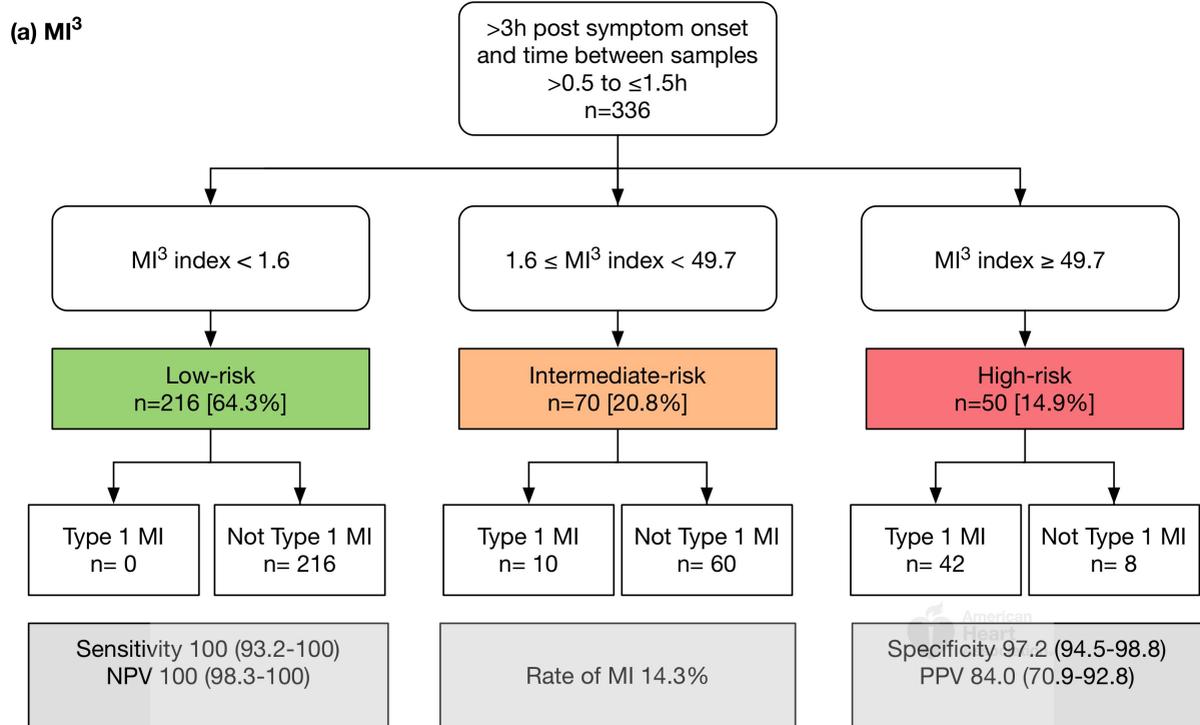
metrics for an individual patient. A screen swipe presents the data in a natural frequency number and graphical format for the patient.



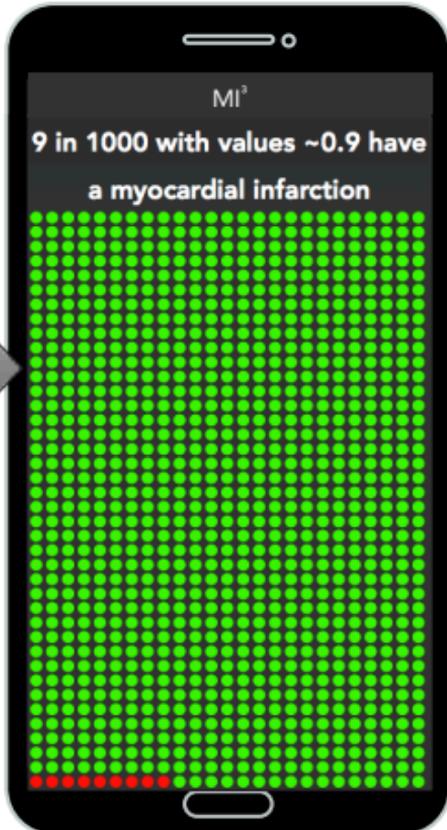
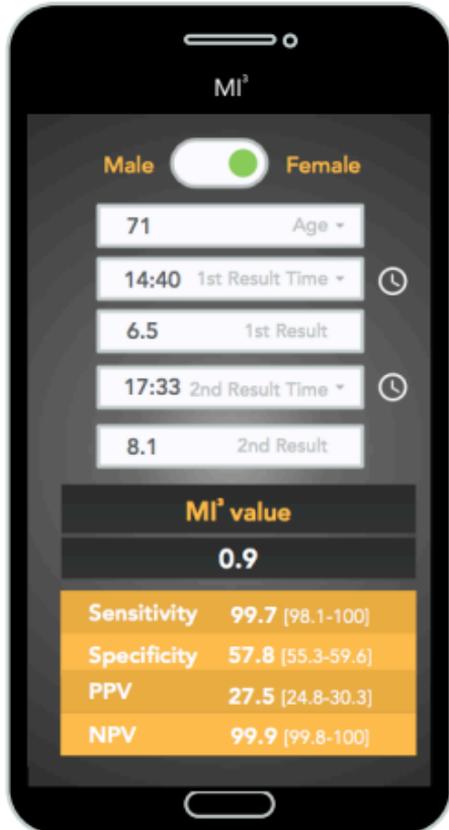
Circulation



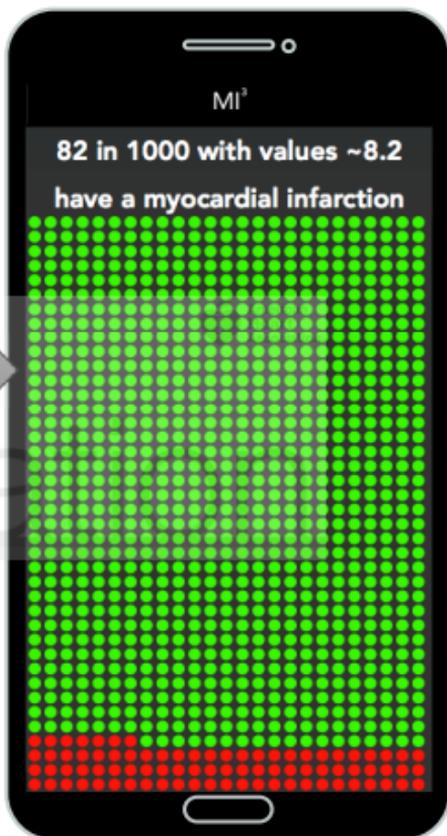
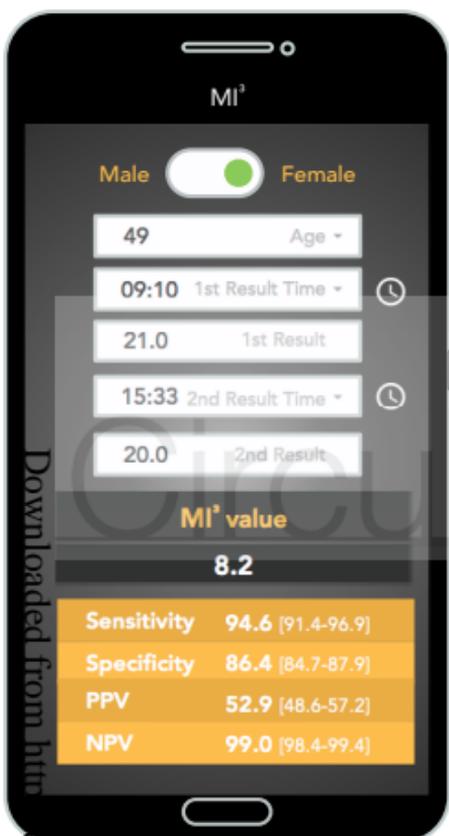




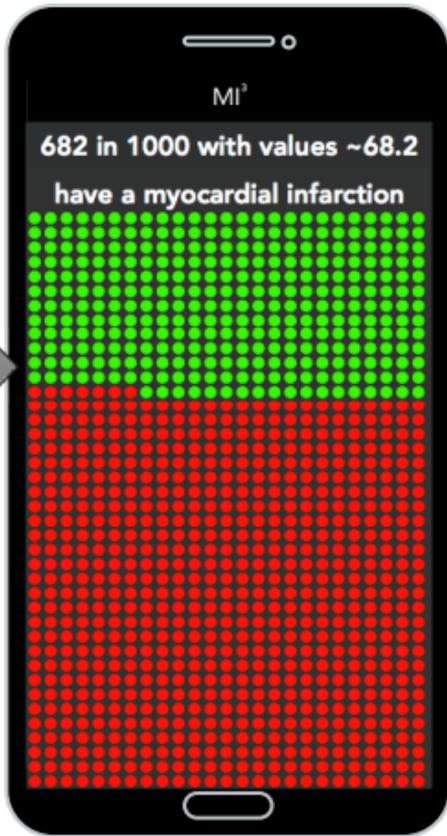
LOW RISK



INTERMEDIATE RISK



HIGH RISK



Downloaded from https://www.ahajournals.org/ by on August 22, 2019