Received: 31 July 2020 | Revised: 11 December 2020 | Accepted: 23 December 2020

DOI: 10.1002/emp2.12363

JACEP OPEN
JOURNAL OF THE AMERICAN COLLEGE OF EMERGENCY PHYSICIANS OPEN
A GLOBAL JOURNAL OF EMERGENCY MEDICINE
Open Access
WILEY

ORIGINAL RESEARCH

Cardiology

Machine learning compared with rule-in/rule-out algorithms and logistic regression to predict acute myocardial infarction based on troponin T concentrations

Anders Björkelund PhD[1] | **Mattias Ohlsson PhD**[1] | **Jakob Lundager Forberg MD, PhD**[2] | **Arash Mokhtari MD, PhD**[3,4] | **Pontus Olsson de Capretz MD**[4,5] | **Ulf Ekelund MD, PhD**[4,5] | **Jonas Björk PhD**[6,7]

[1] Department of Astronomy and Theoretical Physics, Lund University, Lund, Sweden

[2] Department of Emergency Medicine, Helsingborg Hospital, Helsingborg, Sweden

[3] Department of Cardiology, Skåne University Hospital, Lund, Sweden

[4] Department of Clinical Sciences at Lund, Lund University, Lund, Sweden

[5] Department of Emergency Medicine, Skåne University Hospital, Lund, Sweden

[6] Division of Occupational and Environmental Medicine, Lund University, Lund, Sweden

[7] Clinical Studies Sweden, Forum South, Skåne University Hospital, Lund, Sweden

Correspondence
Jonas Björk, PhD, Division of Occupational and Environmental Medicine, Lund University, Box 117, SE-221 00 Lund, Sweden.
Email: jonas.bjork@med.lu.se

Funding and support: This study was part of the AIR Lund (Artificially Intelligent use of Registers at Lund University) research environment, and received funding from the Swedish Research Council (VR; grant no. 2019-00198). The study also received funding from the Swedish Heart-Lung Foundation (2018-0173).

Abstract

Objective: Computerized decision-support tools may improve diagnosis of acute myocardial infarction (AMI) among patients presenting with chest pain at the emergency department (ED). The primary aim was to assess the predictive accuracy of machine learning algorithms based on paired high-sensitivity cardiac troponin T (hs-cTnT) concentrations with varying sampling times, age, and sex in order to rule in or out AMI.

Methods: In this register-based, cross-sectional diagnostic study conducted retrospectively based on 5695 chest pain patients at 2 hospitals in Sweden 2013–2014 we used 5-fold cross-validation 200 times in order to compare the performance of an artificial neural network (ANN) with European guideline-recommended 0/1- and 0/3-hour algorithms for hs-cTnT and with logistic regression without interaction terms. Primary outcome was the size of the intermediate risk group where AMI could not be ruled in or out, while holding the sensitivity (rule-out) and specificity (rule-in) constant across models.

Results: ANN and logistic regression had similar (95%) areas under the receiver operating characteristics curve. In patients (n = 4171) where the timing requirements (0/1 or 0/3 hour) for the sampling were met, using ANN led to a relative decrease of 9.2% (95% confidence interval 4.4% to 13.8%; from 24.5% to 22.2% of all tested patients) in the size of the intermediate group compared to the recommended algorithms. By

Supervising Editor: Nathan R. Hoot, MD, PhD.

contrast, using logistic regression did not substantially decrease the size of the intermediate group.

Conclusion: Machine learning algorithms allow for flexibility in sampling and have the potential to improve risk assessment among chest pain patients at the ED.

KEYWORDS

AI (Artificial Intelligence), cardiovascular epidemiology, computer assisted diagnostic techniques, diagnosis epidemiology, medical decision making, statistics and numerical data machine intelligence

1 | INTRODUCTION

1.1 | Background

Chest pain is one of the most common chief complaints among patients at the emergency department[1] (ED) and constitutes a major burden on the health care system.[2] In chest pain patients, the European Society of Cardiology (ESC) recommends the use of high-sensitivity cardiac troponin T (hs-cTnT) tests applied in either a 0/3-hour or a 0/1-hour algorithm.[3] According to the 0/1-hour algorithm, acute myocardial infarction (AMI) is ruled out (low risk) either if the first hs-cTnT is below 5 ng/L or if the first hs-cTnT is <12 ng/L and the change between the 0-hour and 1-hour samples is <3 ng/L. Conversely, AMI is ruled in (high risk) if the first sample is >52 ng/L or if the change to the

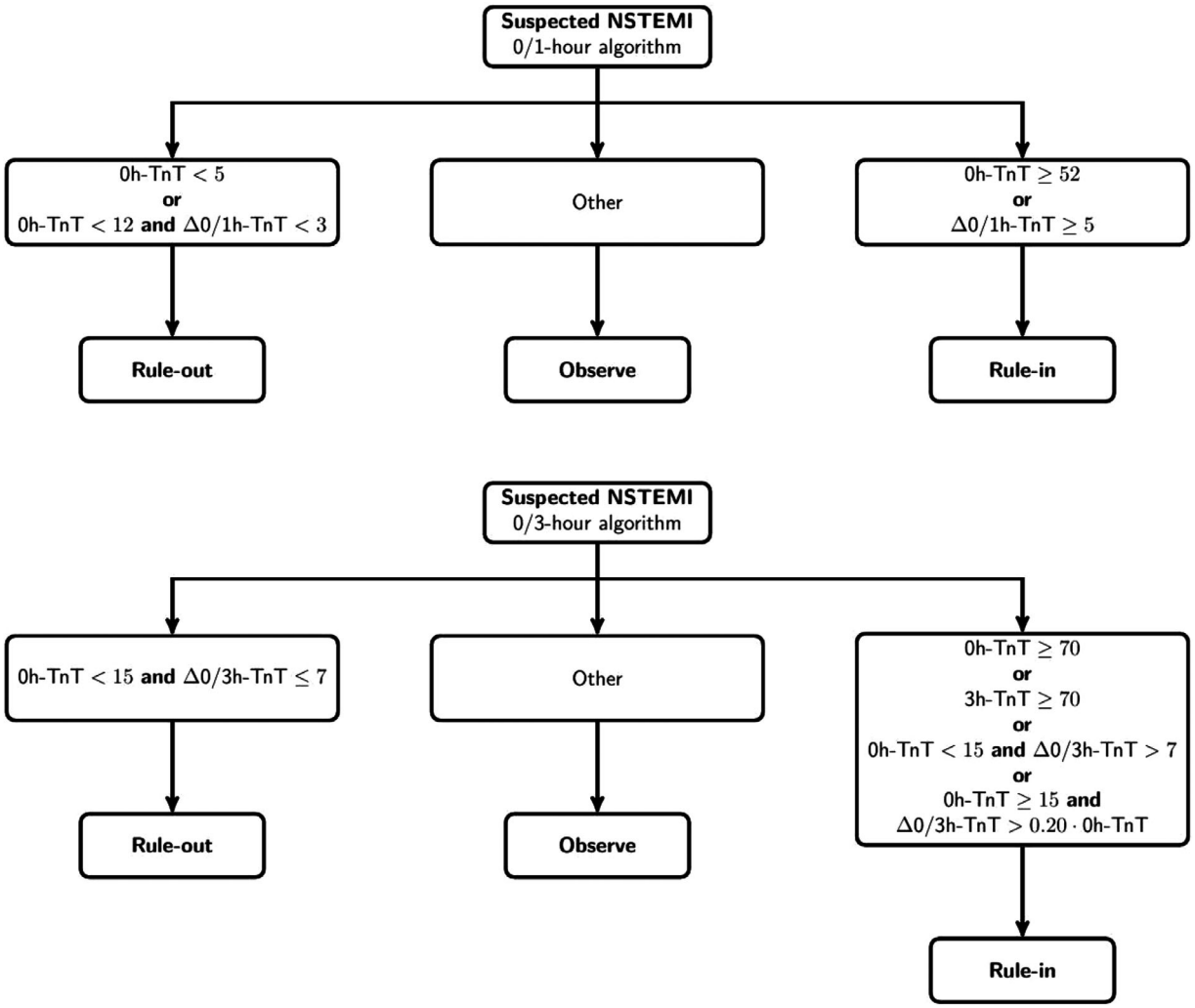

FIGURE 1 European Society of Cardiology (ESC) 0/1-hour and 0/3-hour algorithms for ruling in or out acute myocardial infarction based on high-sensitivity cardiac troponin T measured in ng/L. NSTEMI, non–ST-segment elevation myocardial infarction; TnT, troponin T

second sample is at least 5 ng/L (Figure 1). The ESC algorithms are widely used across EDs in Europe, and have recently been studied in the United States.[3] However, the ESC algorithms have 3 specific limitations. First, the algorithms are based on fixed thresholds and predefined time intervals between first and second sample, which may not be trivially achievable in a stressful ED environment. Second, they do not provide any probabilistic assessment of the risk. Third, they still leave a substantial proportion in an intermediate group that cannot be accurately classified as low or high risk.[4,5]

1.2 | Importance

Recent work has suggested that machine learning algorithms can be used to interpret serial troponin samples in a more flexible manner that does not depend on a specific time interval between samples and with improved prediction accuracy compared with rule-based algorithms.[6] However, the decision support provided by complex models is generally less explainable than rule-based algorithms and statistical models. There is thus a trade-off between performance and explainability in clinical prediction modeling, and added complexity is justified only if it leads to clinically meaningful improvement in prediction accuracy.[7] How much increased model complexity improves prediction accuracy is likely to differ across applications and must therefore be judged on a case-by-case basis.

1.3 | Goals of this investigation

The primary aim of the present study was to assess the predictive accuracy of a machine learning algorithm compared with the ESC 0/1 and 0/3 hour algorithms in ED chest pain patients, where the clinical task is to rule in or out AMI based on paired cardiac hs-cTnT tests, age, and sex. As a secondary aim, we compared the performance with a simpler and more explainable statistical model, logistic regression. As a primary outcome measure, we used the size of the intermediate group while keeping the sensitivity in the rule-out decision and specificity in the rule-in decision constant across the evaluated methods.

The Bottom Line

Diagnosis of myocardial infarction can be challenging. This study applied machine learning techniques (neural networks) to a sample of 5695 patients in Sweden with acute chest pain, using information on demographics and serial cardiac biomarkers to predict acute myocardial infarction. Using this method, the proportion of patients with diagnostic uncertainty after 2 biomarkers decreased (from 24.5% to 22.2%), compared with a clinical practice guideline.

2 | METHODS

2.1 | Study design and setting

The present study is a register-based, cross-sectional diagnostic study conducted retrospectively in ED patients visiting any of 2 hospitals in Region Skåne, Sweden during 2013 and 2014: (1) Skåne University Hospital at Lund (catchment area 310 000 individuals), (2) Helsingborg General Hospital (catchment area 250 000 individuals). The study was approved by the regional ethics review board in Lund, Sweden (Dnr 2018–708 and 2019–03523).

2.2 | Data collection

Patient data on chief complaint, age, and sex and laboratory data on hs-cTnT and discharge diagnosis originate from health care registers in Region Skåne and were extracted from the existing EXPECT (Evaluation of Unknown Predictors of Electrocardiographic Changes – a Transnational study) database.[8,9] The hs-cTnT tests were analyzed using the Roche Cobas e602 (Roche Diagnostics, Basel, Switzerland) at the time of the ED visit. This assay has a limit of blank of 3 ng/L, limit of detection 5 ng/L, and the 99th percentile is 14 ng/L.[10]

2.3 | Selection of participants

The inclusion of study participants is charted in Figure 2. Potentially eligible patients included all adults (≥18 years) who presented to the EDs with chest pain as chief complaint. Subsequent visits by the same patient during the study period were excluded from the database. ST-segment elevation myocardial infarction cases were not included, as this diagnosis is not based on troponin results. As study cohort for the present investigation, we selected a subset of patients who had a first hs-cTnT sample drawn within 4 hours of arrival and a second hs-cTnT taken at least 30 minutes but no >8 hours after the first.

2.4 | Reference standard

As reference standard for AMI at the index visit we used the diagnosis made by the attending physician at the hospital ward or the responsible emergency physician in the event of discharge from the ED. AMI diagnoses were made in accordance with the universal AMI definition 2013–2014[11,12] on the basis of a rise and/or fall of hs-cTnT with at least 1 value above 14 ng/L, together with symptoms, ECG changes, or imaging evidence of infarction. The reference standard was validated against a reference of expert adjudication for a subset of patients (n = 838) included in a separate study.[13,14] The overall agreement was 97%, with 78% sensitivity and 99% specificity (Supplementary Table E1).

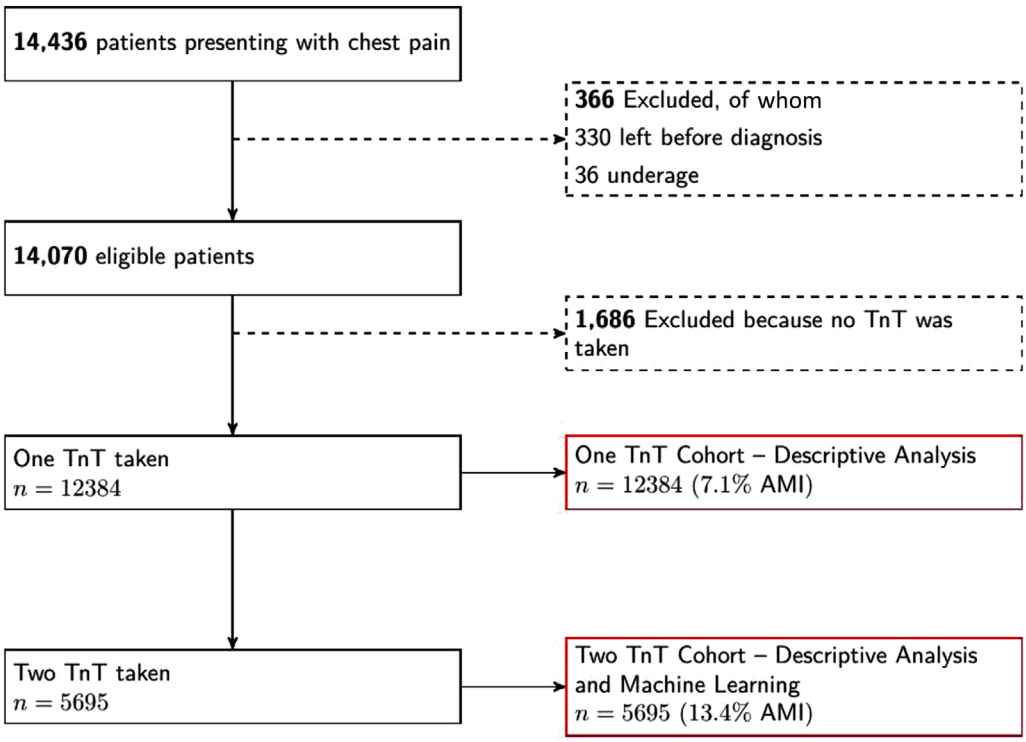

FIGURE 2 Flow chart for enrolment in the study cohort. AMI, acute myocardial infarction; TnT, troponin T

2.5 | Index test-Development of machine-learning and statistical models

We developed 2 different models to predict AMI, an artificial neural network (ANN) and a statistical logistic regression (LogReg) model, using Python 3.7 with Tensorflow version 1.14. Both models were developed on the same data using the same inputs and outputs. The inputs to the models were age, sex, and 2 serial hs-cTnT measurements together with the rate of change between them that were all log-transformed and normalized (mean zero, standard deviation one). The output was index visit AMI (yes/no). Both models can be used for the standard 0/1- and 0/3-hour sampling schemes but also allow for flexibility in the timing between the 2 hs-cTnT measurements.

Logistic regression can be regarded as a special case of neural network without any hidden layers and was thus implemented as a trivial neural network in the Tensorflow package. The ANN was a multilayer perceptron modeled with 1 hidden layer of 10 neurons between the inputs and the output node. To reduce the stochastic nature of the ANN and smooth the probability distribution we trained 10 identical networks using bagging, that is, the training data were bootstrap-resampled for each network and their outputs were combined by taking the mean of their probabilities. Each network used Relu activations on the hidden layer and was trained using the Adam optimizer and a dropout with a rate of 0.1.[15]

2.6 | Validation procedure

ANN and LogReg both give a number between 0 and 1 as output that reflects the estimated probability of AMI conditioned on the input data. To reliably estimate the predictive accuracy of the 2 models we carried out stratified 5-fold cross-validation 200 times each for ANN and LogReg, respectively. Each patient thus received 200 estimated probabilities that were then averaged to get a final probability estimate based on each model separately. The cross-validation setup effectively emulated a situation with a held-out test set, however, allowing the entire data set to be used for testing. Importantly, each patient was assessed by a model that was trained on data not including that very patient. In additional evaluation with calibration of rule-in and rule-out thresholds, we computed the corresponding probabilities averaged over all models when evaluated on their own training data, enabling a comparison between the influence of calibrating thresholds on derivation versus validation data.

2.7 | Data analysis

Initial evaluation of ANN and LogReg was carried out using the area under the receiver operating characteristic curve (AUC), calculated overall and stratified by age, sex, hospital, and time of hs-cTnT sampling.

The ANN and LogReg models were further evaluated against existing ESC rule-in/rule-out algorithms for hs-cTnT–a 0/1-hour algorithm and a 0/3-hour algorithm in the subset of patients with the tests required for these algorithms (Figure 1). These algorithms rely on 2 serial hs-cTnT samples taken 1 or 3 hours apart. Based on the hs-cTnT values and the algorithms, we assigned patients into one of the risk groups (Rule-in, Intermediate, Rule-out). For the 0/1h algorithm we followed the ESC guidelines[16] and considered cases where the second sample was taken between 30 minutes and 1.5 hours after the first. In the 0/3h case the ESC does not prescribe specific levels of hs-cTnT change nor does it assign patients to an intermediate group. Instead, we followed Thygesen et al[17] who defined such thresholds. We considered cases where the second sample was taken at least 2.5 hours after the first to qualify for this algorithm.

As primary outcome measure we used the size of the intermediate group that was obtained when the sensitivity (rule-out) and specificity (rule-in) were kept constant across models and algorithms. We determined rule-out and rule-in thresholds for the probabilities obtained from ANN and LogReg by calibrating them against the performance of the ESC algorithms. Specifically, we selected the probability threshold for rule-out that yielded a negative predictive value (NPV) and sensitivity at least as high as the corresponding metrics for the ESC algorithms. Similarly, we selected the probability threshold for rule-in that yielded a positive predictive value (PPV) and specificity at least as high as those of the ESC algorithms. As an additional analysis, we instead used the probabilities predicted on the training folds to calibrate the ANN and LogReg thresholds for rule-in and rule-out. We used bootstrap resampling (20 000 replications) on the data set with the averaged final probability estimates to obtain 95% confidence intervals (CIs) both for AUC and the primary outcome measure (size of the intermediate group). Model calibration was assessed by using the Hosmer-Lemeshow test calculated for the final probability estimates grouped in deciles.

TABLE 1 Baseline characteristics of (1) all adult ED patients with chest pain with at least 1 sample of high-sensitivity cardiac troponin T (hs-TnT), (2) present study cohort with 2 samples taken

	One sample (n = 12 384)	Study cohort - Two samples (n = 5695)
Age, years, mean (SD)	58.9 (18.8)	65.6 (16.0)
≤65, n (%)	7177 (58.0)	2517 (44.2)
Sex, n (%) female	5882 (47.5)	2496 (43.8)
Hospital, n (%) Lund	7097 (57.3)	3346 (58.8)
Time between hs-TnT samples, n (%)		
One sample	6548 (52.9)	0 (0.0) 0 (0.0)
0.5 – 1.5 h	944 (7.6)	944 (16.6)
1.5 – 2.5 h	1524 (12.3)	1524 (26.8)
≥ 2.5h	3227 (26.0)	3227 (56.7)
Disease history and treatments, n (%)		
AMI	1431 (11.6)	974 (17.1)
Unstable angina	483 (3.9)	344 (6.0)
Coronary artery bypass grafting	1231 (9.9)	849 (14.9)
Percutaneous coronary intervention	726 (5.9)	525 (9.2)
Heart failure	1143 (9.2)	714 (12.5)
Hypertension	4110 (33.2)	2449 (43.0)
Chronic obstructive pulmonary disease	706 (5.7)	426 (7.5)
Diabetes	1549 (12.5)	989 (17.4)
Renal failure	446 (3.6)	309 (5.4)
Peripheral artery disease	588 (4.7)	373 (6.5)
AMI at index visit, n (%)	880 (7.1)	779 (13.7)

AMI, acute myocardial infarction; ED, emergency department.

3 | RESULTS

3.1 | Characteristics of study subjects

A total of 12 384 patients had at least 1 hs-cTnT sample analyzed. The mean age of these patients was 59 years and 7% had AMI at the index visit (Table 1). The 5 695 patients with 2 serial samples hs-cTnT taken were older (mean age 66 years) and had more comorbidities and a higher risk of AMI (14%; Table 1). Figure 3 shows the association between the 2 hs-cTnT samples among patients with or without AMI at the index visit.

3.2 | Main results

The ANN model obtained an AUC of 95.1%, compared with 94.5% for the LogReg model (Figure 4). AUCs were for both models higher among younger than among older patients (Supplementary Table E2). No marked differences in AUC with respect to sex, hospital, or time between hs-cTnT samples were observed. The requirements of the ESC algorithms for the timing between the 2 hs-cTnT samples were met for 4 171 of the 5 695 patients. The 0/1h- and 0/3h-algorithms, each applied to patients meeting their respective time criteria, together ruled out AMI in 2307 (55.3%) of these patients with 96.9% sensitivity and 99.3% NPV (Table 2). A total of 842 patients (20.2%) were ruled in with 89.7% specificity and 55.9% PPV. The remaining intermediate risk patients (1022; 24.5%) had a 6.8% AMI risk. For ANN, the derived probability thresholds were ≤ 0.02164 for rule-out and ≥ 0.1278 for rule-in. The intermediate group decreased to 928 patients (22.2%; relative decrease 9.2%, 95% CI 4.4%–13.8%), while maintaining similar NPV and PPV as for the ESC algorithms. The LogReg model decreased the intermediate group only marginally to 1 007 patients (24.1%; relative decrease 1.5%, 95% CI -3.2% to 5.8%), with corresponding probability thresholds of ≤ 0.03159 and ≥ 0.1363 for rule-out and rule-in, respectively. ANN led to a more substantial decrease in the size of the intermediate group than LogReg also in the complete testing set with all sampling times included (Supplementary Table E3).

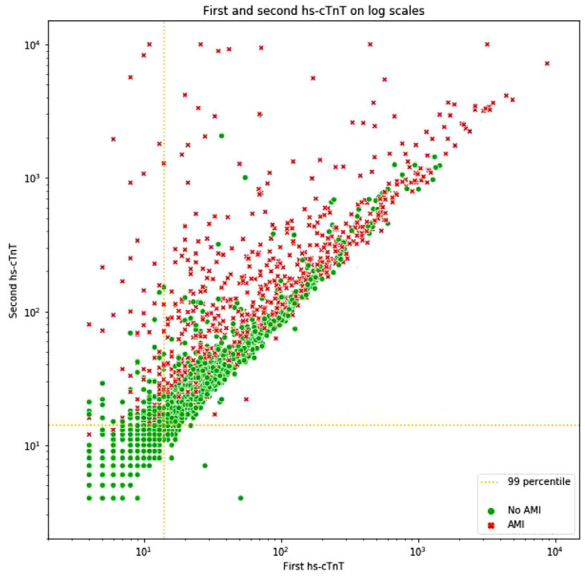

FIGURE 3 Plot of the first and second hs-cTnT sample (n = 5 695) using log-scaled axes. The dashed orange lines indicate the 99th percentile (14 ng/L). ANN, artificial neural network; hs-cTnT, high-sensitivity cardiac troponin T

Table 3 shows the agreement between the individual classifications obtained from ANN and LogReg in the complete testing set. The largest difference off the left-to-right agreement diagonal were 229 patients without AMI that ANN ruled out but that LogReg placed in the intermediate group. Corresponding comparisons with ESC are presented in Supplementary Table E4 and E5. The predicted probabilities of AMI obtained from ANN and LogReg were compared in Figure 5. Two large clusters of likely and unlikely cases emerged where both models agree (bottom left and top right of Figure 5). The most striking outliers were 15 patients without AMI where the LogReg model assigned a high (> 0.5) estimated probability of AMI, whereas the neural network assigned a very low probability (below the rule-in threshold). These patients were all young (aged 20–30) with very high but typically stable levels of hs-cTnT. Neither ANN nor LogReg had completely satisfactory calibration according to the Hosmer-Lemeshow test ($P < 0.001$ for both models). In particular, the 20% lowest probability estimates from ANN (below 0.001) tended to be falsely too low. The LogReg model, on the other hand, failed to capture accurately the increase in AMI risk in the range 0.15 – 0.60, as also suggested by the curved-shaped association between the estimates from ANN and LogReg in this range (Figure 5).

TABLE 2 Number of patients (among the 4 171 qualifying for either of the 0/1h or 0/3h algorithm) ruled in and out by the ESC algorithms compared with the artificial neural network (ANN) and logistic regression (LogReg) models. Sensitivity and specificity for ANN and LogReg were calibrated against the ESC algorithms

	ESC algorithms			ANN			LogReg		
	No AMI	AMI	Total n (%)	No AMI	AMI	Total n (%)	No AMI	AMI	Total n (%)
Rule-out, n (%)	2290	17	2307 (55.3)	2363	17	2380 (57.1)	2291	17	2308 (55.3)
Sensitivity, %	96.9			96.9			96.9		
NPV, %	99.3			99.3			99.3		
Intermediate, n (%)	953	69	1022 (24.5)	880	48	928 (22.2)	952	55	1007 (24.1)
Rule-in, n (%)	371	471	842 (20.2)	371	492	863 (20.7)	371	485	856 (20.5)
Specificity, %	89.7			89.7			89.7		
PPV, %	55.9			57.0			56.7		

AMI, acute myocardial infarction; ESC, European Society of Cardiology; NPV, negative predictive value; PPV, positive predictive value.

TABLE 3 Agreement in individual classifications from the artificial neural network (ANN) and logistic regression (LogReg) models, stratified on outcome (AMI vs no AMI; n = 5695)

	Patients with AMI (n = 779)		
	LogReg rule out	LogReg intermediate	LogReg rule-in
ANN rule-out	18	6	0
ANN intermediate	4	48	12
ANN rule-in	0	16	675
	Patients without AMI (n = 4916)		
	LogReg rule out	**LogReg intermediate**	**LogReg rule-in**
ANN rule-out	2929	229	5
ANN intermediate	133	966	108
ANN rule-in	0	121	425

AMI, acute myocardial infarction.

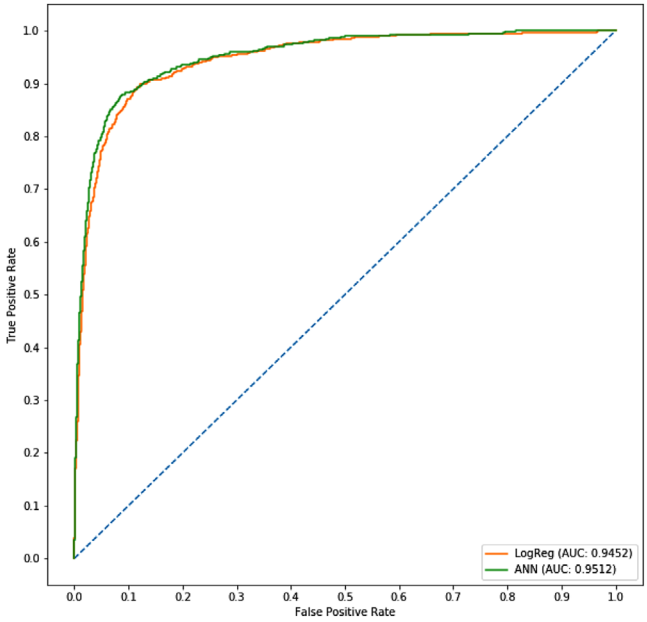

FIGURE 4 Receiver operating characteristic curves for the artificial neural network (ANN) and the logistic regression (LogReg) models (n = 5 695). AUC, area under the curve

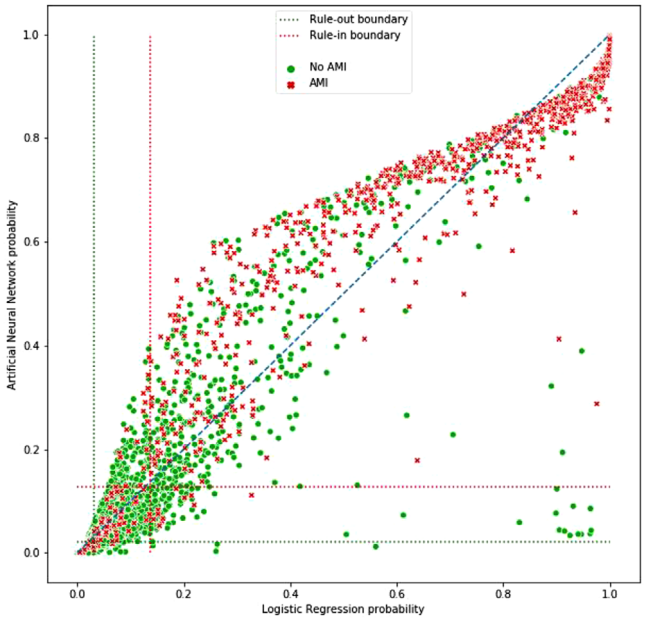

FIGURE 5 Comparison of probabilities of AMI from the logistic regression and artificial neural network models (n = 5 695). The dotted lines denote the probability thresholds calibrated against the ESC algorithms for ruling patients in or out. The closer to the diagonal, the more the models agree. AMI, acute myocardial infarction; ESC, European Society of Cardiology

3.3 | Additional results

When the ANN probability thresholds were calibrated based on the training set the rule-out threshold increased (from 0.02164 to 0.02477) whereas the rule-in threshold remained similar (test: 0.1278, training: 0.1247). This led to decreased rule-out sensitivity and decreased size of the intermediate group (Supplementary Table E6). For LogReg, calibration on the training set also increased the rule-out threshold (from 0.03159 to 0.03229) and the rule-in threshold was similar (test: 0.1363, training: 0.1358). This led to a decreased rule-out sensitivity, which also decreased the size of the intermediate group.

4 | LIMITATIONS

The algorithms were trained and tested on a selected subsample with 2 hs-cTnT analyzed and are therefore not necessarily generalizable to settings where indications for ordering a second sample are different. Additionally, the results may not generalize to settings where the prevalence of AMI differs at entry to the ED presentation.[18] Another limitation was that we used routine care diagnoses. However, although misclassifications in these diagnoses may bias the estimated magnitude of the performance difference toward the null, it should not affect the ranking of the compared algorithms. A further limitation was observed in the sensitivity analyses where both ANN and logistic regression were highly dependent on the choice of probability thresholds. Careful prospective validation, including detailed assessment of model calibration, is therefore warranted before implementation in clinical practice,[7] in order to ensure that safety requirements are met. Another challenge with the clinical implementation is that the ANN would generally require retraining on new populations whenever deployed across hospitals in order to ensure that unbiased probability estimates are maintained. The ANN may also require continuous learning and recalibration if population drift occurs within a particular setting. Finally, it should be noted that we did not extend the LogReg model to incorporate interaction terms between the input variables, for example, between age and troponin levels. Doing so would most likely have decreased the performance gap versus ANN while at the same time increasing the LogReg model complexity.

5 | DISCUSSION

Two salient findings emerged from the present study: (1) our decision-support models using age, sex, and two non-specific time sampling troponins had an improved performance compared to the ESC algorithms; and (2) ANN with hidden layers in our setting led to improved classification of AMI compared with a more explainable method based on logistic regression.

As the primary performance metric, we used the size of the intermediate group where AMI was not ruled in or out. The improvement in performance observed using this metric may seem modest but implies that almost 1 out of the 10 patients kept in the intermediate group by the ESC algorithms can be safely discharged if ANN is used. Similar performance improvement was observed irrespectively of the timing of the sampling. A practical advantage with both ANN and LogReg is that a numeric estimate of the AMI risk is obtained that may guide the

clinical decision further and not just a classification as is the case with the ESC algorithms. This is a particular advantage as health care systems in, for example, Europe and the United States may have different regulatory requirements and tolerances of risk and thus may have different preferred thresholds for ruling out AMI.[6]

There are only a few previous attempts to use machine learning within this field.[6,19,20] Our results are in line with the most recent of these studies, Than et al[6], where improved prediction accuracy of their machine algorithm relative to the ESC 0/1 and 0/3 hour algorithms were reported using hs-cTnI. We extend their results by using hs-cTnT and by showing that the improvement goes beyond what can be obtained with more straightforward logistic regression approaches that do not incorporate dynamic interactions between input variables. A major reason for the improvement is most likely not so much the "black box" power of ANN as such, but rather the more elaborate handling by the ANN of how level and changes in cardiac troponin concentration and demographic variables interact with respect to the likelihood of AMI.

Age is a strong risk factor for AMI and males have elevated risks at younger ages than females.[21] A previous study showed that the NPV of a hs-cTnT < 5 ng/L at presentation was similarly high in men and women and when stratified by age or cardiovascular risk factors.[22] However, the independence of the prognostic value of troponin does not necessarily extend to situations with paired hs-cTnT and varying time sampling schemes. We saw that the risk predictions and classifications by the ANN for specific troponin patterns differed with age. By contrast, classifications did not seem to vary with sex. Our expectation is that the performance advantage of using ANN versus more explainable models that do not allow for dynamic interactions will increase further if additional input features such as pain history, ECG, and comorbidities are used.[13] Another potentially interesting usage of the ANN methodology is when only the first hs-cTnT sample is available together with other clinical data as a support tool to guide the decision to rule in, rule out, or order a second sample.

In summary, versatile algorithms that are not bound to specific sampling schemes or fixed tolerances of risk have the potential to improve risk assessment among patients presenting with chest pain at the ED. Compared to the explainable models (the ESC algorithm and the logistic regression model), the machine learning model with hidden layers can detect interactions between for example, troponin levels and demographic variables that could potentially aid the physicians in reducing the intermediate risk group of patients. However, the intermediate group was still considerable. Adding other relevant features from the current ED visit, as well as from previous health care records, is therefore a topic for additional investigations.

AUTHOR CONTRIBUTIONS

AB, MO, AM, UE, and JB conceived the study, and it was developed further together with JLF and POC. AB was responsible for model development in collaboration with MO and JB. AB and JB drafted the manuscript. All authors reviewed the manuscript for important intellectual content and approved the final version.

REFERENCES

1. Pitts SR, Niska RW, Xu J, Burt CW. National Hospital Ambulatory Medical Care Survey: 2006 emergency department summary. *Natl Health Stat Report*. 2008;(7):1-38.
2. Makam AN, Nguyen OK. Use of cardiac biomarker testing in the emergency department. *JAMA Intern Med*. 2015;175:67-75.
3. Nowak RM, Christenson RH, Jacobsen G, et al. Performance of novel high-sensitivity cardiac Troponin I assays for 0/1-hour and 0/2- to 3-hour evaluations for acute myocardial infarction: results from the high-US study. *Ann Emerg Med*. 2020;76:1-13.
4. Mokhtari A, Borna C, Gilje P, et al. A 1-h combination algorithm allows fast rule-out and rule-in of major adverse cardiac events. *J Am Coll Cardiol*. 2016;67:1531-1540.
5. Reichlin T, Schindler C, Drexler B, 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.
6. Than MP, Pickering JW, Sandoval Y, et al. Machine learning to predict the likelihood of acute myocardial infarction. *Circulation*. 2019;40:899-909.
7. Kelly CJ, Karthikesalingam A, Suleyman M, Corrado G, King D. Key challenges for delivering clinical impact with artificial intelligence. *BMC medicine*. 2019;17:195.
8. Hansen TG, Pottegard A, Brandes A, et al. New-onset atrial fibrillation among patients with infection in the emergency department: a multicenter cohort study of 1-year stroke risk. *Am J Med*. 2020;133:352-359.e3.
9. Schade Hansen C, Pottegard A, Ekelund U, et al. Association between QTc prolongation and mortality in patients with suspected poisoning in the emergency department: a transnational propensity score matched cohort study. *BMJ open*. 2018;8:e020036.
10. Giannitsis E, Kurz K, Hallermayer K, Jarausch J, Jaffe AS, Katus HA. Analytical validation of a high-sensitivity cardiac troponin T assay. *Clin Chem*. 2010;56:254-261.
11. Thygesen K, Alpert JS, Jaffe AS, et al. Third universal definition of myocardial infarction. *Eur Heart J*. 2012;33:2551-2567.
12. Thygesen K, Alpert JS, White HD. Universal definition of myocardial infarction. *J Am Coll Cardiol*. 2007;50:2173-2195.
13. Mokhtari A, Lindahl B, Smith JG, Holzmann MJ, Khoshnood A, Ekelund U. Diagnostic accuracy of high-sensitivity cardiac troponin t at presentation combined with history and ECG for ruling out major adverse cardiac events. *Ann Emerg Med*. 2016;68:649-658.e3.
14. Mokhtari A, Lindahl B, Schiopu A, et al. A 0-hour/1-hour protocol for safe, early discharge of chest pain patients. *Acad Emerg Med*. 2017;24:983-992.
15. Goodfellow I, Bengio Y, Courville A. *Deep Learning*. Cambridge, Massachusetts: MIT Press; 2016.
16. Roffi M, Patrono C, Collet JP, 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.
17. Thygesen K, Mair J, Giannitsis E, et al. How to use high-sensitivity cardiac troponins in acute cardiac care. *Eur Heart J*. 2012;33:2252-2257.
18. Kohn MA, Kwan E, Gupta M, Tabas JA. Prevalence of acute myocardial infarction and other serious diagnoses in patients presenting to an urban emergency department with chest pain. *J Emerg Med*. 2005;29:383-390.
19. Green M, Björk J, Forberg J, Ekelund U, Edenbrandt L, Ohlsson M. Comparison between neural networks and multiple logistic regression to predict acute coronary syndrome in the emergency room. *Artif Intell Med*. 2006;38:305-318.
20. Eggers KM, Ellenius J, Dellborg M, et al. 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.
21. Anand SS, Islam S, Rosengren A, et al. Risk factors for myocardial infarction in women and men: insights from the INTERHEART study. *Eur Heart J*. 2008;29:932-940.
22. Shah AS, Anand A, Sandoval Y, et al. High-sensitivity cardiac troponin I at presentation in patients with suspected acute coronary syndrome: a cohort study. *Lancet (London, England)*. 2015;386:2481-2488.

AUTHOR BIOGRAPHY

Jonas Björk is professor in Epidemiology at Lund University, Sweden, principal investigator for the interdisciplinary research programme AIR Lund, Artificially Intelligent use of Registers at Lund and Halmstad Universities.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

How to cite this article: Björkelund A, Ohlsson M, Forberg JL, et al. Machine learning compared with rule-in/rule-out algorithms and logistic regression to predict acute myocardial infarction based on troponin T concentrations. *JACEP Open*. 2021;2:e12363. https://doi.org/10.1002/emp2.12363