Critical Review Form
Diagnostic Test

Herman R, Meyers HP, Smith SW, et al. International evaluation of an artificial intelligence–powered electrocardiogram model detecting acute coronary occlusion myocardial infarction. Eur Heart J Digit Health. 2024;5(2):123-133.

Objectives: “introduce an international validation of an automated deep learning artificial intelligence (AI) model detecting acute OMI [occlusive myocardial infarction] using only a single-standard 12-lead ECG as input and hypothesize that it would outperform the existing state-of-the-art ECG criteria for the detection of acute OMI and match the performance of interpreters with special expertise in ECG OMI diagnosis.” (p. 125)

Methods: This retrospective study evaluated an automated deep learning AI model for detecting acute occlusive myocardial infarction (OMI) using only single, standard 12-lead ECGs. It proceeded in four stages: model development in a derivation cohort, blinded validation in geographically distinct European and US test sets, comparison with state-of-the-art STEMI-based ECG criteria and expert interpreters, and subgroup performance analyses. The study drew data from 9,764 suspected ACS patients at the Cardiovascular Centre Aalst (2011–2021) in Belgium and an additional international ACS image database of 2,368 patients.

High-resolution waveform data were exported from the MUSE system or reconstructed from ECG images using CE-certified PMcardio digitization. ECGs recorded more than 24 hours before or after coronary angiography, or with poor signal quality, were excluded. Included patients were randomly assigned to a derivation or an internal European validation set. The derivation set included ECGs adjudicated as OMI or not OMI by expert ECG interpreters and angiographic outcomes, using a reference standard based on ischemic symptoms, troponin elevation, and angiographic evidence of an acute culprit lesion with TIMI flow 0–1 or TIMI 2–3 requiring emergent or urgent PCI; patients without dynamic biomarker changes were ruled out for OMI even if they underwent angiography.

In the internal EU test set, independent reviewers adjudicated culprit vessels, stenosis, TIMI flow, collateral circulation, and treatment strategy, including timing from first ECG to revascularization. The US external test set came from the DOMI ARIGATO database, which includes clinically verified angiograms and expert-annotated ECGs from two US centers. Benchmarking compared the AI model against blinded STEMI-criteria assessment and OMI labels from two ECG experts with high interobserver agreement.

The derivation cohort comprised 18,616 ECGs from 10,543 patients (mean age 66 ± 14 years, 65.9% male, 22.9% OMI), while the overall test cohort included 3,254 ECGs from 2,222 patients (mean age 62 ± 14 years, 67% male, 21.6% OMI), split into an internal EU cohort (2,016 ECGs from 1,630 contacts, 16% OMI) and a US cohort (1,238 ECGs from 633 contacts, 36.2% OMI). US patients were younger, had more ECGs before angiography, and more frequently had STEMI-positive ECGs, whereas sex distribution, peak troponin, and culprit-vessel TIMI flow were similar between cohorts.

Critical Review Form: Diagnostic Test

<table>
<tr><th>Guide</th><th>Comments</th></tr>
<tr><td colspan="2">Are the results valid?</td></tr>
<tr><td>Did clinicians face diagnostic uncertainty?</td><td>Yes. The datasets included patients with suspected ACS, but without coronary angiography results available.</td></tr>
<tr><td>Was there a blind comparison with an independent gold standard applied similarly to all patients?
(Confirmation Bias)</td><td>No. Not all patients underwent coronary angiography (62.2% in the overall test sets, 58.2% in the internal EU test set, and 72.7% in the US test set underwent coronary angiography). There is risk of partial and differential verification bias. It is also likely that those performing and interpreting the gold standard (angiography) were aware of ECG findings (incorporation bias).</td></tr>
<tr><td>Did the results of the test being evaluated influence the decision to perform the gold standard?
(Ascertainment Bias)</td><td>Yes. While not explicitly states, it is clear that ECG findings influenced the decision to perform coronary angiography. Specifically, those patients with ECG findings consistent with STEMI criteria would be much more likely to undergo coronary angiography (and presumably to do so sooner) than those not meeting STEMI criteria.</td></tr>
<tr><td colspan="2">What are the results?</td></tr>
<tr><td>What likelihood ratios were associated with the range of possible test results?</td><td>At the "optimal threshold" (overall test set), likelihood ratios for diagnosing OMI were as follows:
STEMI criteria: LR+ 14.1, LR- 0.69
AI Model: LR+ 12.8, LR- 0.21
ECG experts: LR+ 17.0, LR- 0.28

• The OMI AI model achieved an AUC of 0.94 (95% CI 0.92–0.95) overall, with similar performance on EU internal (AUC 0.95) and US external (AUC 0.90) test sets.
• Sensitivity across demographic and ECG subgroups was generally stable (≈72–78%), with slightly higher specificity in patients aged <45 and 45–65 years.
• Sensitivity was markedly higher in patients presenting with STEMI ECGs (93.3% vs 67.6% without STEMI), while specificity was higher in those without ST elevation (94.2% vs 68.7% with ST elevation).
• Specificity was somewhat lower for left anterior descending and right coronary artery culprit lesions (~84% and ~81%, respectively), but sensitivity did not differ significantly by culprit artery.
• Performance remained comparable when alternative OMI definitions (different TIMI flow/troponin cutoffs or PCI-based definitions) were used.</td></tr>
</table>

	• Compared with STEMI criteria, the AI model had much higher sensitivity for OMI (80.6% vs 32.5%) and similar sensitivity to ECG experts (73.0%), with accuracy equal to experts and higher than STEMI criteria. • STEMI criteria had the highest specificity (97.7%), with slightly lower specificity for experts (95.7%) and the AI model (93.7%). • Mean time to OMI diagnosis was shorter with the AI model than with STEMI criteria (2.3 vs 5.3 hours) and similar to ECG experts (2.9 hours). • Rates of intervention (e.g., PCI) in OMI patients were similar regardless of presence of STEMI criteria or which OMI definition was applied. • AI model "false positives" had a higher rate of myocardial injury than STEMI-based false positives (14.4% vs 2.4%), similar to expert false positives (~14%). • Of OMI cases missed by STEMI criteria (330 patients), 61% of those with delayed revascularization (>2 hours) were correctly identified by the AI model on the first ECG, with a median revascularization time of 9.3 hours. • The AI model correctly identified 42% of OMI cases missed by ECG experts, who had a median time to angiography of 7.2 hours and were frequently inferior/posterior territory lesions.
How can I apply the results to patient care?	
Will the reproducibility of the test result and its interpretation be satisfactory in my clinical setting?	In theory yes. The authors studied automated deep learning artificial intelligence (AI) model for detecting acute OMI based on ECG interpretation. This model could theoretically be used for any ECG conducted on any patients with suspect ACS. The model was evaluated in both US and EU patient populations and should be generalizable to patients in our institution (external validity).
Are the results applicable to the patients in my practice?	Yes. We frequently encounter patients with symptoms concerning for ACS and use ECGs in all of these patients to evaluate for STEMI or other occlusive MI, along with risk stratification tools and cardiac enzymes. The use of AI to assist in ECG interpretation would be valuable, particularly in those patients who do not meet STEMI criteria.
Will the results change my management strategy?	No. At this time we do not have access to AI-assisted ECG interpretation. Further research to understand the ability of AI to detect OMI, and further improvements in AI technology, will be needed before most institutions will be willing to use this technology.
Will patients be better off as a result of the test?	Perhaps. While it is unlikely that AI-assisted ECG interpretation will replace physician interpretation, as further research and improvements in technology occur it is likely that the use of AI will help augment ECG interpretation and lead to improved accuracy in questionable cases. While this study did not demonstrate superiority to "ECG experts," there may still be benefit in the assistance of emergency physicians.

Limitations:

1. **This was a retrospective, observational study subject to several potential sources of bias (confounding, recall bias).**

2. **Not all patients underwent coronary angiography, raising the risk of partial and differential verification bias.**

3. **It is likely that those performing and interpreting the gold standard (angiography) were aware of ECG findings (incorporation bias).**

4. **Significant differences in clinical practice between the US and EU were suggested by these results, potentially limiting the generalizability of the overall findings.**

Bottom Line:

This retrospective study evaluating an automated deep learning AI model for detecting acute occlusive myocardial infarction (OMI) based on 12-lead ECG in both US and EU patient populations found that while AI outperformed STEMI criteria, it provided similar accuracy to ECG experts. Further refinement of technology and further prospective research on the impact of AI-assisted ECG interpretation are needed.