SIIM CAIMI25
Conference on Artificial Intelligence in Medical Imaging
OCT 20-21 | UNIVERSITY OF CALIFORNIA, IRVINE

Predicting Major Adverse Cardiovascular Events Using Multimodal Models

Frank Li, PhD, Emory University; Kéana Aitcheson, MBBS; Theo Dapamede, MD, PhD; Hari Trivedi, MD; Judy Gichoya, MD, MS, FSIIM

Introduction/Background
Traditional 10-year Major Adverse Cardiovascular Events (MACE) risk tools like ASCVD risk scores and PREVENT rely on lab inputs that may be unavailable. This study focuses on evaluating CXR-based multimodal models as an accessible alternative for MACE prediction, comparing their performance to image-based (ConvNext CNN), embedding-based, and ASCVD-derived logistic regression models, as well as the standard ASCVD risk score.

Methods/Intervention
This retrospective study included 3,494 patients (2,828 training; 666 testing, aged 40–79) with 55.3% male demographic and a racial make-up of 45.0% Black, 46.8% White and 8.2% Other. Frontal chest radiographs (varied types) from 2009–2021 were used. All models were trained on the same dataset (50:50 MACE vs. no MACE, 80:20 train-validation split). The embedding model used CXR-derived embeddings from a pretrained RAD-DINO model. The multimodal model combined embeddings with LVEF values (averaged if ranges), while the ASCVD model applied logistic regression. The image-based model used full CXRs inputted to a ConvNext CNN. Time-dependent AUC over 10 years, Cox models (hazard ratios, concordance), and classification metrics (sensitivity, specificity, AUC), partial hazard, ROC, and Youden's J thresholds were used for evaluation.

Results/Outcome
Time-dependent AUC showed the multimodal model performed best (AUC: 0.699 at year 1, 0.692 at year 10), followed by image (0.688) and embedding (0.685) models. ASCVD-based models performed lower (10-year AUC = 0.583). Cox models showed all predictors were significantly associated with MACE ($p < 0.005$), with the image model showing the strongest effect (HR = 13.13). In multivariable analysis, only the image model remained independently predictive (HR = 7.72). Kaplan-Meier curves showed good risk separation, particularly for image, embedding, and multimodal models. Threshold-based classification showed the multimodal model had the best balance of sensitivity (0.66), specificity (0.53), and AUC (0.614), while ASCVD models had high sensitivity (0.83) but low specificity (0.30).

Conclusion
Multimodal DL approaches integrating CXRs, echocardiographic and clinical data provide superior long-term MACE prediction compared to traditional risk scores. Image-only and embedding models demonstrate standalone prognostic potential.

Statement of Impact
CXR-based models provide accessible, opportunistic MACE risk assessment, valuable for asymptomatic individuals or settings with limited lab access.

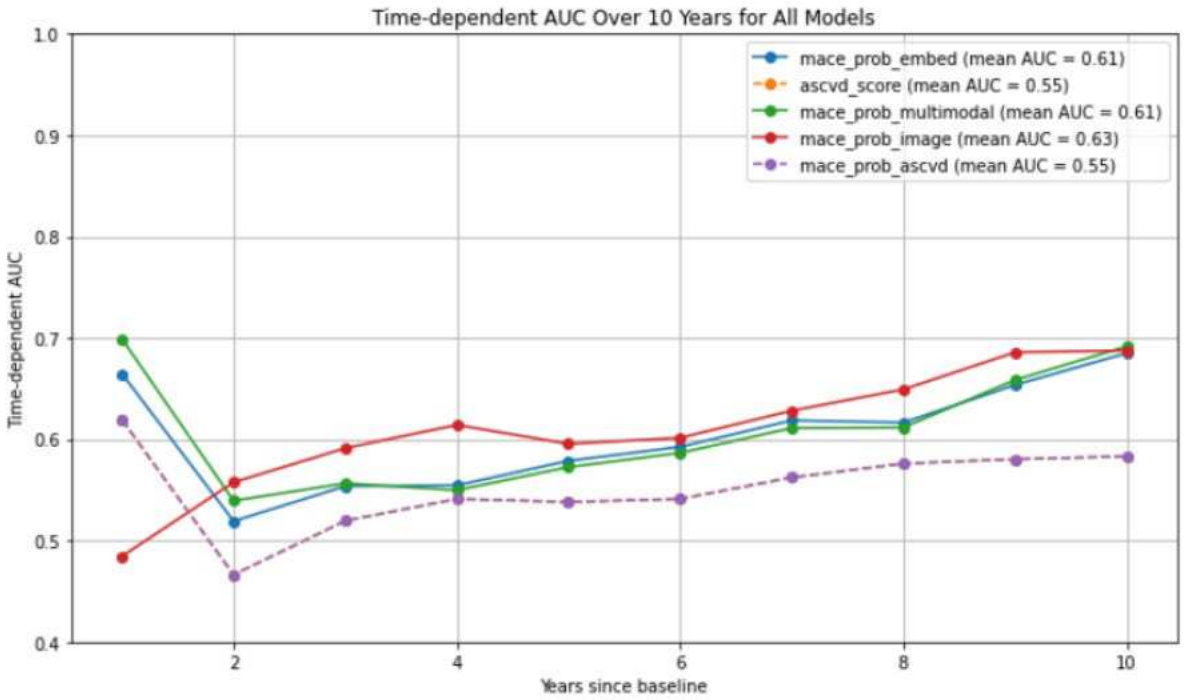

Figure 1: Time-dependent AUC curves comparing the discriminative performance of four models in predicting major adverse cardiovascular events (MACE) over a 10-year period, against the ASCVD risk scores.

Keywords

Multimodal Models; Foundation Models; Prognostic Models; CXR; ASCVD; MACE