NAPCRG 52nd Annual Meeting — Abstracts of Completed Research 2024.

Submission Id: 6939

Title

Machine learning models for atrial fibrillation detection in primary care using electronic health records: systematic review

Priority 1 (Research Category)

Cardiovascular disease

Presenters

Mhd Diaa Chalati, MD, Chetan Shirvankar, MBBS, Samira Rahimi, PhD, Eng

Abstract

Context: Atrial fibrillation (AFib) significantly impacts patient morbidity and mortality, despite existing screening practices. Machine learning (ML) models offer potential for improved detection of AFib from electronic health records (EHR). However, uncertainties remain regarding the performance of these models compared to standard care, their generalizability and applicability in primary care settings.

Objective: To synthesize data on the effectiveness, generalizability, and clinical relevance of ML models in detecting AFib cases using EHR in primary care settings.

Study Design and Analysis: This systematic review followed PRISMA guidelines. We searched seven databases from inception to May 2023. Data extraction and appraisal followed CHARMS checklist. Risk of bias was assessed with PROBAST tool and clinical applicability with MI-CLAIM checklist. Two independent reviewers screened and extracted data using Covidence® with disagreements resolved by a third reviewer.

Setting or Dataset: Studies conducted in primary care settings were included. Studies using only ECG or wearable devices for AFib detection were excluded.

Population Studied: Adults without a prior AFib diagnosis.

Intervention/Instrument: Application of ML models to EHR.

Outcome Measures: Detection of new AFib cases, comparing ML models to standard care.

Results: Of 4,536 references, 16 studies met inclusion criteria. Most were from North America (n=10) and used retrospective cohort designs (n=12). Only four models were externally validated. Risk of bias was uncertain due to insufficient reporting on handling missing data. Most studies did not share their code. Due to heterogeneity in performance measures, meta-analysis was not feasible. AUROC ranged from 0.71 to 0.948, with seven ML models outperforming control groups. Models combining ML with other clinical tools showed improved discrimination. ML identified novel AFib detection variables, including recent heart failure, cancer, and MRI/echo-based metrics.

Conclusion: Several ML models outperformed standard care in AFib detection from EHR, especially when combined with other clinical tools. However, Uncertain risk of bias, limited external validation, and ambiguity regarding code sharing and missing data hinder ML generalizability in primary care. Future research should standardize reporting, validation methods, and calibration to improve adoption and reproducibility.