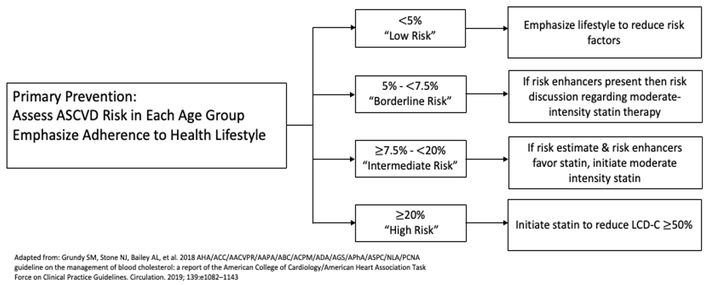
An EHR-Based Machine Learning Model Predicts Myocardial Infarction Better than an ECG-based Machine Learning Model and the Pooled Cohort Equations

Linyuan Jing, PhD[1], John M. Pfeifer, MD, MPH[2,3], Martin Kang, MD[2], Dustin Hartzel[1], Sushravya Raghunath, PhD[2], Brandon K. Fornwalt, MD, PhD[2], Christopher M. Haggerty, PhD[1,4]

[1]Department of Translational Data Science and Informatics, Geisinger, Danville, PA, USA, [2]Tempus Labs Inc, Chicago, IL, USA, [3]Heart and Vascular Center, Evangelical Hospital, Lewisburg, PA, USA, [4]Heart Institute, Geisinger, Danville, PA, USA

INTRODUCTION

- Patients at high risk for myocardial infarction (MI) benefit from treatments designed for primary prevention, especially cholesterol lowering therapy.
- The pooled cohort equations (PCE) are the most commonly used risk predictor for future atherosclerotic cardiovascular disease (ASCVD), but show only modest performance.



Adapted from: Grundy SM, Stone NJ, Bailey AL, et al. 2018 AHA/ACC/AACVPR/AAPA/ABC/ACPM/ADA/AGS/APhA/ASPC/NLA/PCNA guideline on the management of blood cholesterol: a report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines. Circulation. 2019; 139:e1082–1143

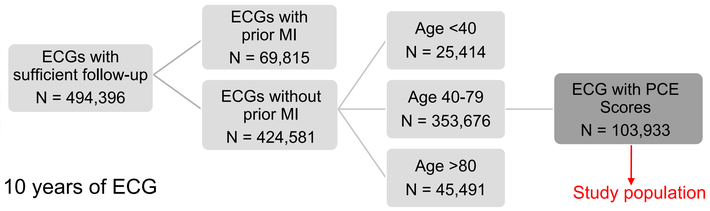
Hypothesis:

Electronic health record (EHR)-based, and ECG-based machine learning models are better at predicting MI as compared to PCE.

METHODS

Study Population

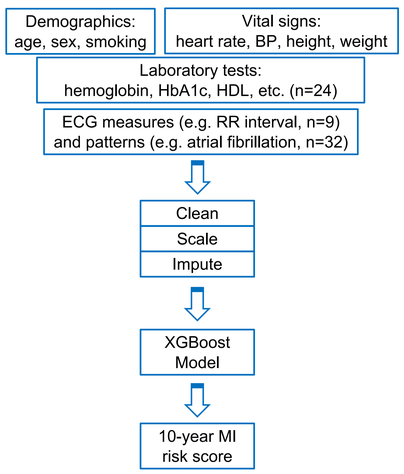
All Geisinger patients ages 40-79 who had 1) at least 1 clinically acquired ECG 2) no history of MI and 3) had PCE scores calculated at the time of ECG



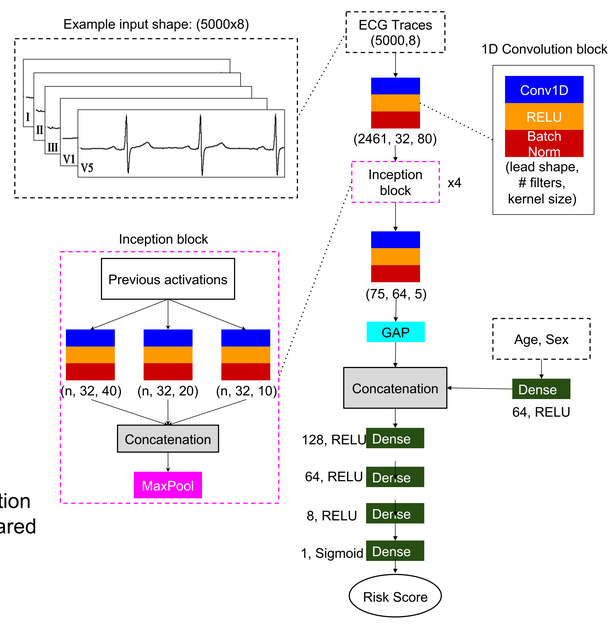
Primary Endpoint: MI event within 10 years of ECG

EHR Data and Machine Learning Models

1. **XGBoost** model with structured EHR data as input features



2. **Deep neural network (DNN)** that used ECG voltage data (10 second, 8 independent leads), age, and sex as inputs



- Models were evaluated by 5-fold cross-validation
- Performances of different models were compared using area under the receiver operating characteristic curve (AUROC)

SUMMARY

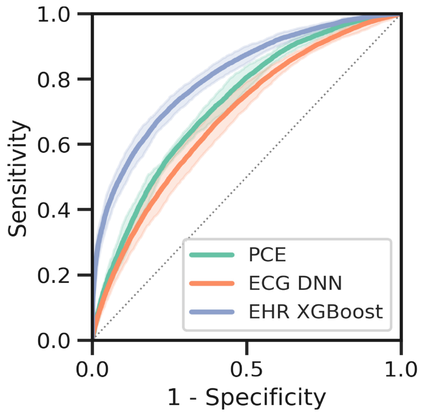
- An EHR-based XGBoost model, but not an ECG-based DNN, is superior to the PCE in predicting future MI.
- Patients identified as high risk by the EHR-based model, but low risk by the PCE, have a high rate of future MI.
- Statin use in that group is low, suggesting ample opportunity for intervention.

RESULTS

A total of **103,933** ECGs from **34,932** patients had sufficient follow-up (occurrence of MI or 10 years follow-up in EHR), **21%** of ECGs were followed by an MI event within 10 years.

Age, yr	**62 (10)**	**Heart Failure**	**5%**
Sex, male	51%	Hypertension	68%
BMI, kg/m2	32 (9)	Diabetes	33%
Smoking	61%	Atrial Fibrillation	17%

The EHR-based XGBoost model had the best performance



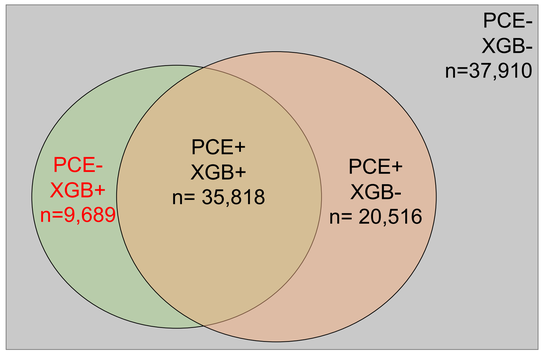
%	PCE	**EHR XGB**	ECG DNN
AUROC	72	**81**	68
Sensitivity	76	**76**	76
Specificity	59	**66**	54
PPV	20	**24**	19
NPV	95	**95**	94

Note: operating points for XGB and DNN models were selected by matching PCE sensitivity

9% (n=9,689) of the total encounters were predicted to be 'high risk' by the EHR-based model and not by the pooled cohort equations.

➢ 10-year MI event rate in this group was 26%

➢ Only 40% of patients in this group were on a statin



	PCE+ XGB+	PCE+ XGB-	PCE- XGB+	PCE- XGB-
n	35,818	20,516	9,689	37,910
MI events	14454	2495	2486	2127
Event rate	40%	12%	26%	6%
% on statin	50%	42%	40%	32%

20% (n=20,516) of encounters were predicted to be 'high risk' by the PCE and not by the EHR-based model. The event rate in that subgroup was 12%.

Limitations

- Retrospective data only
- PCE scores not available on all patients

Disclosures: Geisinger receives funding from Tempus for ongoing development of predictive modeling technology and commercialization. None of the Geisinger authors have ownership interest in any of the intellectual property resulting from the partnership. John Pfeifer, Brandon Fornwalt and Sushravya Raghunath are Tempus employees.

Acknowledgements: We thank Amrita A. Iyer, Ph.D, for poster preparation & review.

Correspondence: John M. Pfeifer, john.pfeifer@tempus.com