EHR-based Machine Learning Model Predicts Drug-induced QT Prolongation With Superior Performance Compared To Clinical Risk Predictors

Linyuan Jing, PhD,[1,2] Thomas Morland, MD,[1] Christopher Kelsey,[1]
Sushravya Raghunath, PhD,[2] John Pfeifer, MD,[2] Jeffrey Ruhl,[1]
Brandon Fornwalt, MD, PhD,[2] Christopher Haggerty, PhD[1]

1. Geisinger, Danville, PA, USA

2. Tempus Labs Inc, Chicago, IL, USA

tmorland@geisinger.edu

Geisinger

TEMPUS

Disclosures

- Funding
  - Geisinger receives funding from Tempus for ongoing development of predictive modeling technology and commercialization
  - None of the Geisinger authors have ownership interests in any of the intellectual property resulting from the partnership. However, Dr. Fornwalt, Dr. Raghunath, and Dr. Pfeifer are now Tempus employees

- Individual Disclosures
  - L. Jing: Employee of Tempus Labs, Inc. T. Morland: Research Grant; Significant; Tempus Labs, Inc. Stock Shareholder; Significant; United Therapeutics Corporation. Other; Significant; United Therapeutics Corporation. C.R. Kelsey: None. S. Raghunath: Employee of Tempus Labs, Inc. J. Pfeifer: Employee of Tempus Labs, Inc. J. Ruhl: None. B. Fornwalt: Other; Significant; Tempus. C.M. Haggerty: Stock Shareholder; Modest; Merck & Co. Other; Significant; Tempus.

What is Drug-induced Long QTc?

- Drug-induced long QTc is an iatrogenic delay in ventricular repolarization which can lead to the life-threatening arrhythmia, *torsades de pointes* (TDP)

- >200 drugs are QTc prolongers
  - Antiarrhythmics
  - Antibiotics
  - Antipsychotics
  - Antidepressants
  - Cancer therapies

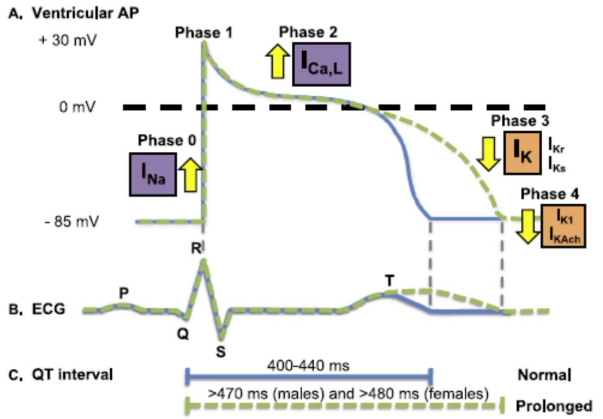


FIG 1. Electrophysiological basis of LQTS. (A) Tracings of the normal cardiac ventricular action potential (blue) observed in healthy cases and prolonged cardiac ventricular action potential (green) observed in long QT syndrome. (B) Schematic representation of a normal ECG (blue) and QT interval prolongation (green). (C) Schematic depiction of normal (blue) and prolonged (green) QT intervals.

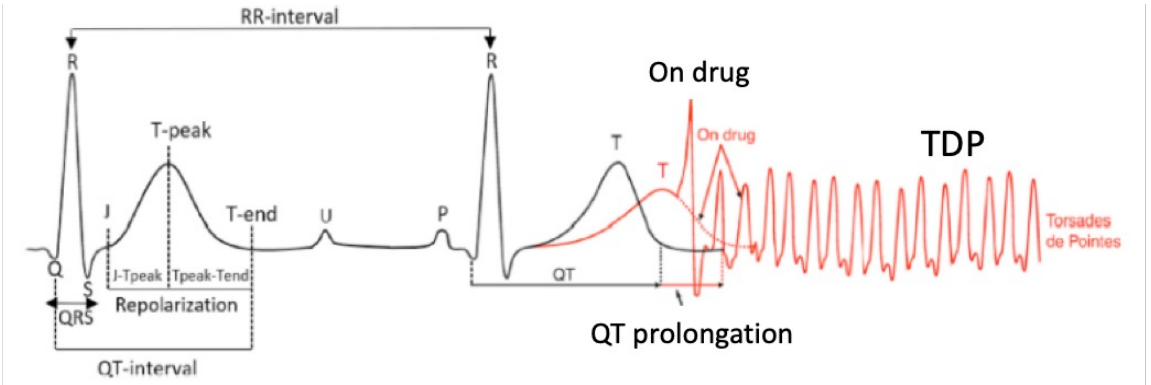


Fig. 1. An illustration of drug-induced QT-interval prolongation and the R-on-T phenomenon that initiates torsades de pointes (TdP) arrhythmia on the electrocardiogram (ECG).

Current Solutions

Tisdale Score

- Predicts QTc >500 ms or >60 ms compared to baseline
- Prospective, observational study of 1,200 patients in cardiac critical care unit
- Good performance: AUC=0.832

- Limitation: lack of generalizability due to inpatient focus

Optimized RISQ-PATH Score

- Broader application
  - Applicable to outpatient
- Moderate performance: AUC=0.772

- Limitation: data collected at target ECG, not reflecting baseline status at drug start

Tisdale *et al.* (2013): *Circulation: Cardio Quality/Outcomes*
Vandael *et al.* (2018): *British Journal of Clinical Pharmacology*

Potential for a Machine Learning Approach

Machine learning has demonstrated high accuracy (e.g., AUROC >0.900) on targets related to QTc but has not been used to produce a general clinical model with high accuracy

Eur Heart J. 2021 Oct 7;42(38):3948-3961. doi: 10.1093/eurheartj/ehab588.

Deep learning analysis of electrocardiogram for risk prediction of drug-induced arrhythmias and diagnosis of long QT syndrome

Edi Prifti [1 2], Ahmad Fall [1], Giovanni Davogustto [3], Alfredo Pulini [1 4], Isabelle Denjoy [5], Christian Funck-Brentano [6], Yasmin Khan [7], Alexandre Durand-Salmon [7], Fabio Badilini [8], Quinn S Wells [3 9], Antoine Leenhardt [5], Jean-Daniel Zucker [1 2], Dan M Roden [3 9 10], Fabrice Extramiana [5], Joe-Elie Salem [3 6 9]

Original Investigation

February 10, 2021

Use of Artificial Intelligence and Deep Neural Networks in Evaluation of Patients With Electrocardiographically Concealed Long QT Syndrome From the Surface 12-Lead Electrocardiogram

J. Martijn Bos, MD, PhD[1,2]; Zachi I. Attia, PhD[3]; David E. Albert, MD[4]; et al

Author Affiliations | Article Information

JAMA Cardiol. 2021;6(5):532-538. doi:10.1001/jamacardio.2020.7422

ELECTROPHYSIOLOGY AND ARRHYTHMIAS

SESSION TITLE: MACHINE LEARNING FOR ARRHYTHMIA PREDICTION

Abstract 15056: Machine Learning Prediction of Long QT Syndrome

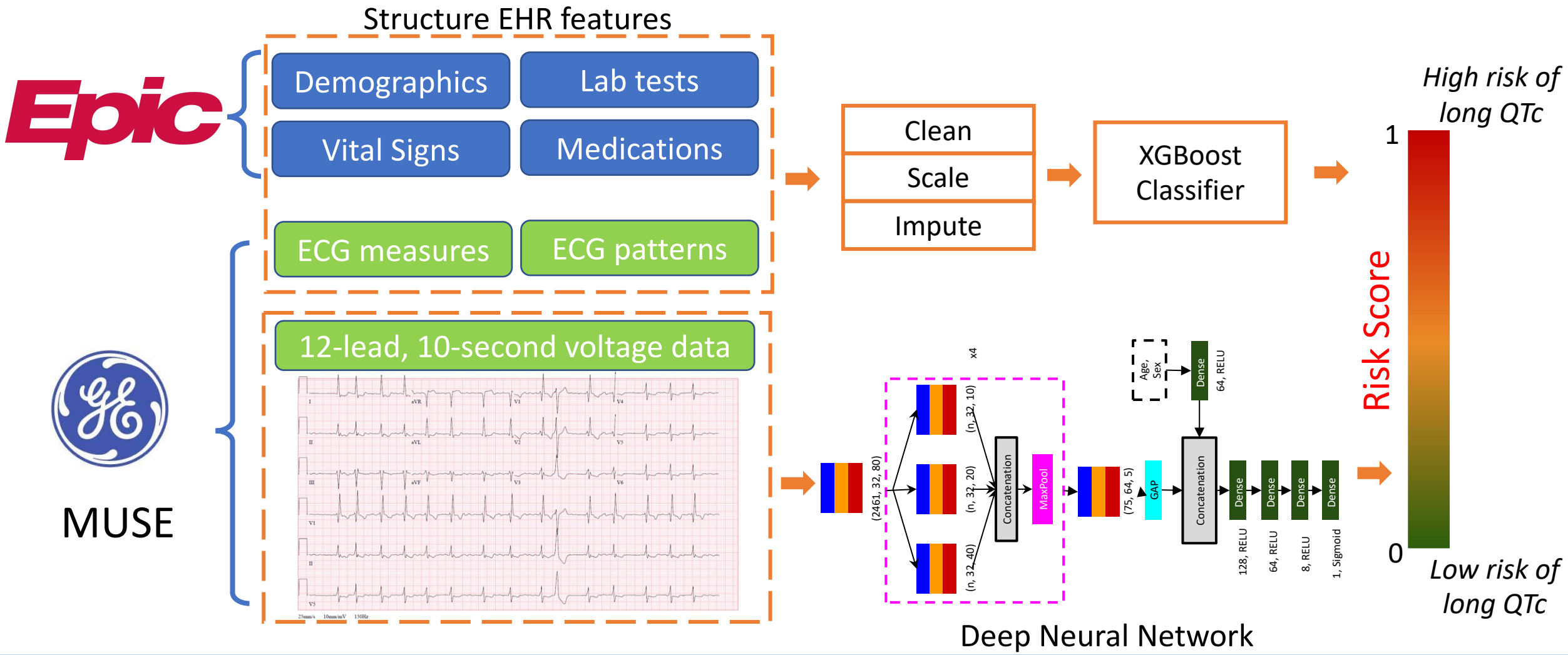
Steven Simon, Divneet Mandair, Michael A Rosenberg and **Premanand Tiwar**

Originally published 12 Nov 2020 | https://doi.org/10.1161/circ.142.suppl_3.15056 | Circulation. 2020;142:A15056

Design Criteria

- Highly accurate across a variety of clinical settings and scenarios
- Derived and validated utilizing data commonly available in the electronic health record (EHR) *at the time of medical decision-making*
- Assessment of drug-induced long QTc risk in general population as opposed to a congenital LQTS population

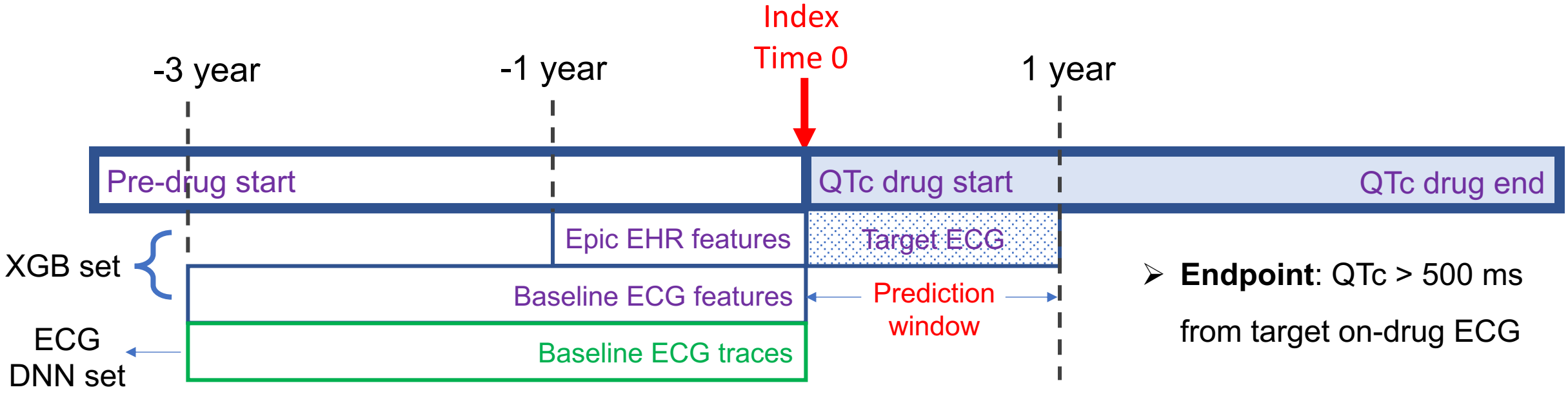
Study Design



Geisinger

TEMPUS

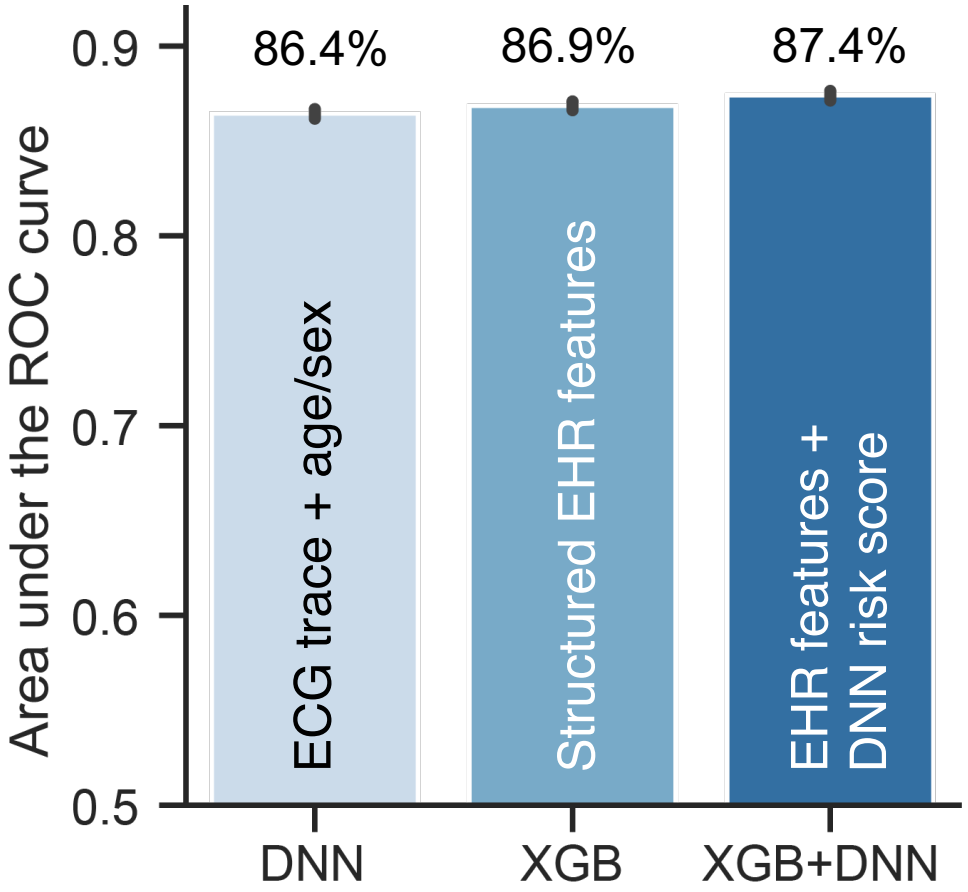
Study Population



- **Endpoint**: QTc > 500 ms from target on-drug ECG

	XGB dataset	ECG DNN dataset
N patients	345,371	182,448
% Long QTc	5.7%	7.7%
Median Age, yrs	62 [48, 74]	66 [54, 78]
% Male sex	45%	47%
Total med count	2 [1, 3]	2 [1, 3]

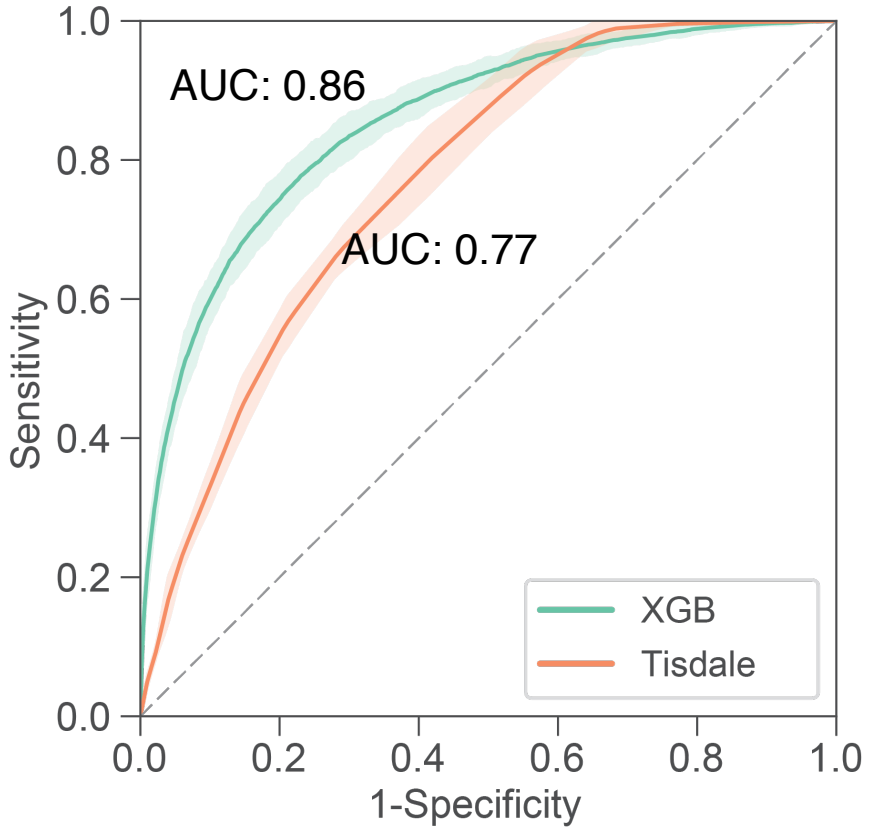
Both XGBoost and DNN Trace Models Show Comparable Performance



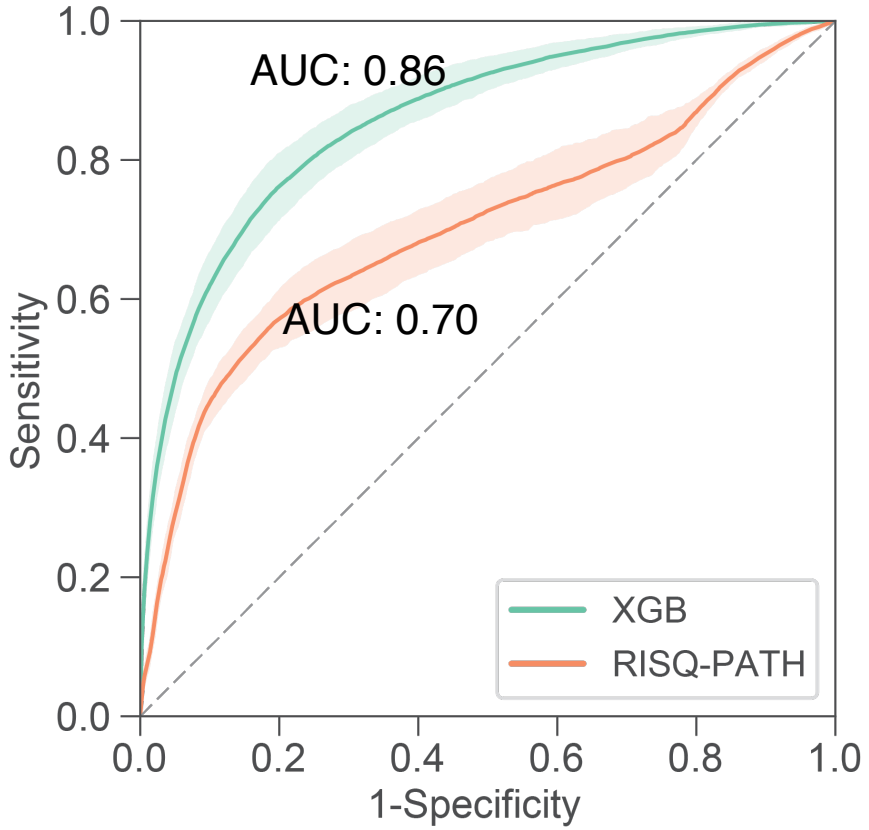
- Dataset:
  - N = 182,448,
  - 7.7% long QTc (>500 ms)
- No significant improvement in composite model

All models were evaluated by 5-fold cross-validation (reported as mean[95% CI])

Machine Learning Shows Superior Predicting Power to Tisdale and RISQ-PATH



N=110,588, 8.8% long QTc



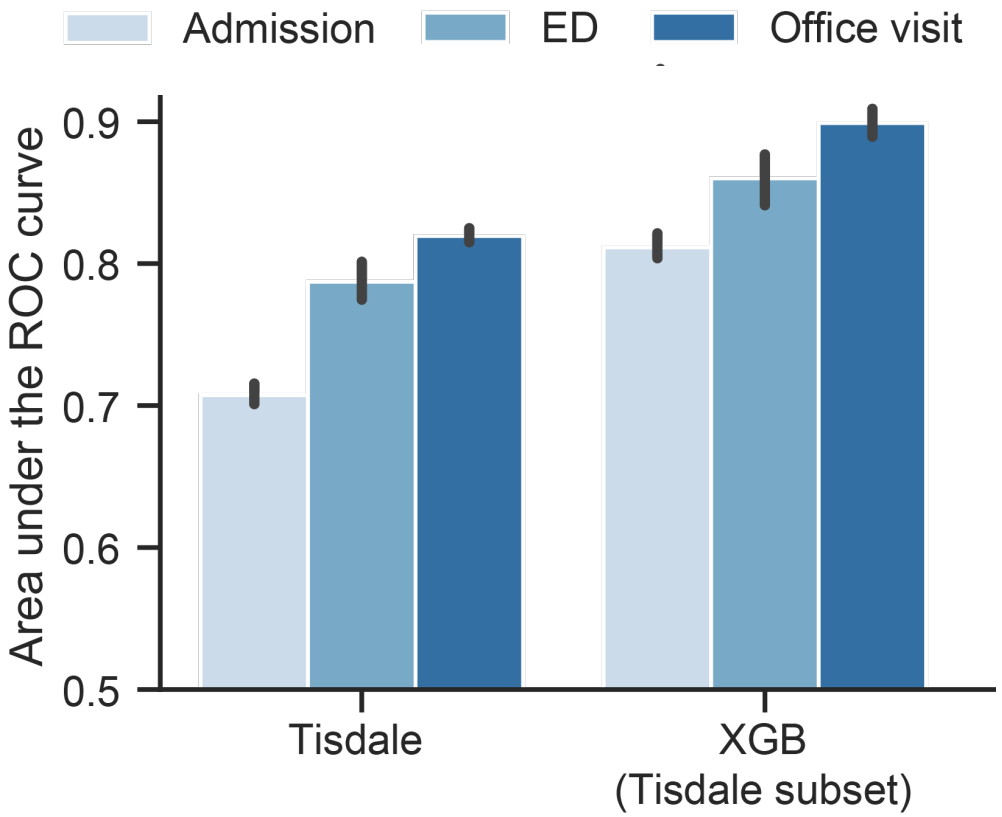
N=345,371, 5.7% long QTc

Geisinger

Tisdale scores were only available in patients with ECGs within 24h of drug start

TEMPUS

Machine Learning Shows Superior Performance in All Clinical Settings



Setting	Data size	% Long QTc
Total	110,588	8.8
Office Visit	32,271	3.3
Admission	62,528	13
ED	15,759	3.6

- Tisdale scores were predominantly calculated in hospital admissions
- In contrast, 75% of the overall dataset were from outpatient records

Machine Learning Model Has Better PPV

- When matching number of patients predicted as high-risk for long QTc, XGBoost model shows superior PPV (>50%) as compared to Tisdale and RISQ-PATH

	% Long QTc	% Predicted High-risk	Sensitivity	Specificity	PPV	NPV
Tisdale	8.8 [8.0, 9.7]	3.1 [1.7, 4.5]	9.8 [6.5, 13.1]	97.6 [96.3, 98.8]	28.3 [25.1, 31.5]	91.8 [90.9, 92.6]
XGBoost		3.1 [2.4, 3.8]	28.9 [22.2, 35.5]	98.5 [98.0, 99.0]	54.0 [49.8, 58.3]	95.8 [95.1, 96.5]
RISQ-PATH	5.7 [5.2, 6.2]	2.0 [1.3, 2.7]	9.8 [6.7, 12.9]	98.5 [97.9, 99.0]	28.3 [25.8, 30.7]	94.7 [94.1, 95.3]
XGBoost		2.0 [1.5, 2.5]	21.4 [16.5, 26.3]	99.2 [98.9, 99.5]	61.5 [56.8, 66.2]	95.4 [94.8, 96.0]

Results are reported as mean [95% confidence interval] across 5 test folds

Conclusions

- Machine learning models, either using structured EHR features or using ECG traces alone, can predict drug-induced QTc prolongation at medication initiation in a general clinical population with high accuracy
  - ➢ Composite model using both EHR features and DNN risk score shows marginal improvement in model performance
- Machine learning model is superior to clinical risk predictors, Tisdale and RISQ-PATH at different clinical settings and operating points

Limitations

- Model trained on all QTc drugs: variable QTc risks and different mechanisms of action
  - Model was evaluated on commonly used, individual drugs but performance was unreliable due to small sample size for single medication
- Model endpoint QTc >500 ms
  - Rational:
    - QTc > 500 ms can be harmful with or without TDP, e.g., interrupted or adjusted pharmacotherapy
    - QTc > 500ms is a reasonable surrogate for risk of death from TDP
  - Limitation: TDP events may not be captured
    - Death from TDP rarely captured on ECG
- Retrospective data from a single health system
  - Could validate initial findings with external data and/or prospective study

Acknowledgements

- Geisinger and Tempus teams
- Tempus Laboratories funding
- Thank you!

tmorland@geisinger.edu

Appendix

Machine Learning Model Shows Better Metrics than Original Tisdale/RISQ-PATH Studies

- When matching sensitivities originally reported by Tisdale and RISQ-PATH, XGBoost model shows better specificity.

	% Long QTc	% Predicted High-risk	Sensitivity	Specificity
Original Tisdale Cohort	30.7		74	77
XGBoost	5.7 [5.2, 6.2]	21.6 [16.4, 26.7]	**74.3 [65.7, 82.8]**	**81.6 [76.5, 86.7]**
Original Optimized RISQ-PATH Cohort	5.9		87.4 [86.2, 88.5]	46.2 [45.8, 46.6]
XGBoost	5.7 [5.2, 6.2]	52.8 [41.8, 63.7]	**92.4 [88.2, 96.6]**	**49.7 [38.3, 61.0]**

Geisinger

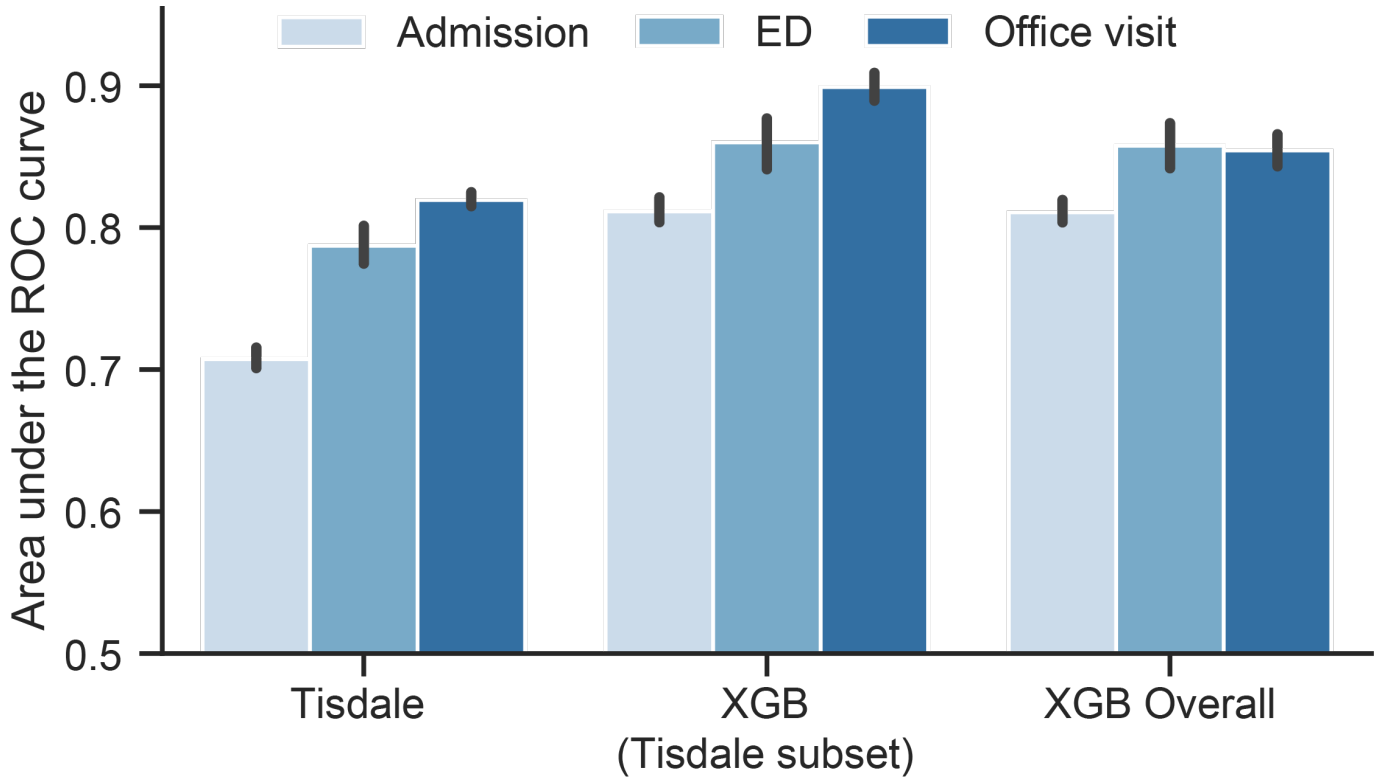
Results are reported as mean [95% confidence interval] across 5 test folds

TEMPUS

Patient Characteristics

Variable	Median[IQR] or %	Variable	Median[IQR] or %
N patients	345,371	**Heart failure**	13%
Long QTc	5.7%	**Diabetes**	21%
Age, yrs	62.2 [48.4, 74.4]	**Hypertension**	49%
Male sex	45.1%	**Myocardial Infarction**	11%
White race	95.7%	**QTc, ms**	440 [421, 462]
BMI, kg/cm^2	29.5 [25.1, 35]	**Max QTc ever**	456 [434, 484]
Ever smoker	51.7%	**QTc med count**	2 [1, 3]
Potassium	4.2 [3.9, 4.5]	**Sr. Creatinine**	0.9 [0.8, 1.1]

Machine Learning Shows Superior Performance in All Clinical Settings



Setting	Data size	% Long QTc
Tisdale subset		
Office Visit	32,271	3.3
Admission	62,528	13
ED	15,759	3.6
Overall dataset		
Office Visit	262,294	4.1
Admission	66,137	12.9
ED	16,940	3.7