Circulation

ORIGINAL RESEARCH ARTICLE

Artificial Intelligence–Enabled Assessment of the Heart Rate Corrected QT Interval Using a Mobile Electrocardiogram Device

Editorial, see p 1299

John R. Giudicessi, MD, PhD*
Matthew Schram, PhD*
J. Martijn Bos, MD, PhD
Conner D. Galloway, BS, MSc
Jacqueline B. Shreibati, MD, MS
Patrick W. Johnson, BS
Rickey E. Carter, PhD
Levi W. Disrud, CCT, CRAT
Robert Kleiman, MD
Zachi I. Attia, PhD
Peter A. Noseworthy, MD
Paul A. Friedman, MD
David E. Albert, MD
Michael J. Ackerman, MD, PhD

BACKGROUND: Heart rate–corrected QT interval (QTc) prolongation, whether secondary to drugs, genetics including congenital long QT syndrome, and/or systemic diseases including SARS-CoV-2–mediated coronavirus disease 2019 (COVID-19), can predispose to ventricular arrhythmias and sudden cardiac death. Currently, QTc assessment and monitoring relies largely on 12-lead electrocardiography. As such, we sought to train and validate an artificial intelligence (AI)–enabled 12-lead ECG algorithm to determine the QTc, and then prospectively test this algorithm on tracings acquired from a mobile ECG (mECG) device in a population enriched for repolarization abnormalities.

METHODS: Using >1.6 million 12-lead ECGs from 538 200 patients, a deep neural network (DNN) was derived (patients for training, n = 250 767; patients for testing, n = 107 920) and validated (n = 179 513 patients) to predict the QTc using cardiologist-overread QTc values as the "gold standard". The ability of this DNN to detect clinically-relevant QTc prolongation (eg, QTc ≥500 ms) was then tested prospectively on 686 patients with genetic heart disease (50% with long QT syndrome) with QTc values obtained from both a 12-lead ECG and a prototype mECG device equivalent to the commercially-available AliveCor KardiaMobile 6L.

RESULTS: In the validation sample, strong agreement was observed between human over-read and DNN-predicted QTc values (–1.76±23.14 ms). Similarly, within the prospective, genetic heart disease–enriched dataset, the difference between DNN-predicted QTc values derived from mECG tracings and those annotated from 12-lead ECGs by a QT expert (–0.45±24.73 ms) and a commercial core ECG laboratory [10.52±25.64 ms] was nominal. When applied to mECG tracings, the DNN's ability to detect a QTc value ≥500 ms yielded an area under the curve, sensitivity, and specificity of 0.97, 80.0%, and 94.4%, respectively.

CONCLUSIONS: Using smartphone-enabled electrodes, an AI DNN can predict accurately the QTc of a standard 12-lead ECG. QTc estimation from an AI-enabled mECG device may provide a cost-effective means of screening for both acquired and congenital long QT syndrome in a variety of clinical settings where standard 12-lead electrocardiography is not accessible or cost-effective.

*J. Giudicessi and M. Schram contributed equally.

Key Words: artificial intelligence ■ electrocardiography ■ long QT syndrome ■ machine learning

Sources of Funding, see page 1285

https://www.ahajournals.org/journal/circ

Clinical Perspective

What Is New?

- A deep neural network can learn, without any instruction or predetermined mathematical algorithm, to calculate accurately QTc values within a large collection of human overread 12-lead ECGs.
- The ability of a deep neural network to recognize clinically meaningful QTc prolongation (eg, >500 ms) on mobile ECG device tracings was similar to 12-lead ECG–based QTc assessments determined by both a QT expert and a commercial core laboratory.

What Are the Clinical Implications?

- The ability to equip mobile ECG devices with accurate artificial intelligence–powered approaches capable of calculating accurately the QTc represents a potential paradigm shift regarding how and where the QT interval can be assessed.
- Artificial intelligence–enabled mobile ECG devices may provide an effective, ubiquitous, and cost-effective means to screen for congenital long QT syndrome, monitor for acquired/drug-induced QT prolongation, and prevent QT-mediated torsades de pointes/sudden cardiac death in a variety of clinical settings not amenable to conventional 12-lead electrocardiography.

Prolongation of the heart rate–corrected QT interval (QTc), defined minimally by current expert consensus guidelines as a QTc value >450 ms in males and >460 ms in females,[1] and practically by sex-specific 99.5th percentile QTc values (>460 ms for prepubertal males and females, >470 ms for postpubertal males, and >480 ms for postpubertal females),[2] can predispose to torsades de pointes (TdP), ventricular fibrillation, and potentially sudden cardiac death (SCD).[3] Whereas an acute, exaggerated QTc increase (ie, ΔQTc >60 ms) serves as a marker of TdP/SCD risk in drug-induced long QT syndrome (LQTS),[4] persistent, otherwise unexplained prolongation of the QTc ≥500 ms at baseline is highly suspicious for congenital LQTS.[5,6] Furthermore, even modest QTc prolongation, perhaps secondary to elevated sympathetic tone, subclinical atherosclerosis, and/or underlying electrolyte/metabolic abnormalities, in middle-aged and older adults is associated with an increased risk of cardiovascular events, stroke, and all-cause mortality independent of traditional clinical risk factors.[7–9]

Despite a growing body of evidence that suggests QTc prolongation, whether secondary to drugs (www.crediblemeds.org), electrolyte abnormalities (hypokalemia), genetics (congenital LQTS), or systemic diseases that now includes SARS-CoV-2–mediated coronavirus disease 2019 (COVID-19), is associated with adverse outcomes in a variety of clinical situations, routine assessment and monitoring of cardiac repolarization still relies largely on traditional 12-lead electrocardiography.[3] As a result, accessibility and cost issues serve as barriers to widespread implementation of universal ECG screening for SCD-predisposing disorders, such as congenital LQTS,[10] or expanded QTc monitoring to identify/prevent drug-induced TdP/SCD in patients receiving ≥1 QTc-prolonging drug(s).[11]

Therefore, with the emergence of smartphone-enabled, single- and multilead ECG devices assisted by artificial intelligence (AI)–aided interpretation algorithms, we sought to (1) train, test, and validate a deep neural network (DNN) to accurately predict the 12-lead ECG–derived QTc value from a subset of those 12 ECG leads that would be obtained from a mobile ECG (mECG device), and (2) prospectively test this DNN-aided QTc detection algorithm using lead I and II ECG data collected from both a standard 12-lead ECG and a U.S. Food and Drug Administration (FDA)–cleared smartphone-enabled mECG device in 686 patients (50% with congenital LQTS) evaluated in the context of the Mayo Clinic Windland Smith Rice Genetic Heart Rhythm Clinic.

METHODS

Data Sources and Definitions

The data, analytic methods, and study materials will not be made available to other researchers for purposes of reproducing the results or replicating the procedure to protect patient privacy. In this Mayo Clinic Institutional Review Board–approved study, all patients with at least 1 digital standard 12-lead ECG acquired in the Mayo Clinic Heart Rhythm and Physiological Monitoring Laboratory between December 1993 and July 2017 were eligible for inclusion in the development and validation datasets used in the AI-based DNN analysis. All standard 12-lead ECGs were acquired in the supine position at a sampling rate of 500 Hz using GE Marquette systems where the GE Marquette system–derived QTc value was then overread by physician-supervised ECG technicians to identify and correct rare, but inherent measurement errors associated with the proprietary approach to QT/QTc assessment utilized by the Marquette 12SL program.[12] Raw electrocardiographic data were stored using GE's MUSE data management system (GE Healthcare, Chicago, IL) for later retrieval.

To generate the training and validation datasets, fully anonymized clinical data, including digitally-stored ECGs and basic demographic/health data, on 674 063 patients was extracted from the Mayo Clinic's proprietary Digital Data Vault. The overall dataset was then partitioned by randomizing patients with two-thirds allocated for model development (development dataset, n=449 378 total patients) and one-third allocated for validation (validation dataset, n=224 685 total patients) as outlined in Figure 1. After exclusion of ECGs that lacked an annotated QT interval, ability to generate an average beat, or patients that were previously part of a hyperkalemia AI ECG study (to avoid cross-contamination),[13] the development dataset was further partitioned randomly by

Table 1. Baseline Characteristics of the Training, Validation, and Testing Study Populations/Datasets

	Training set (n=250 767)	Testing set (n=107 920)	Validation set (n=179 513)
Demographics			
Average age at ECG, y	61±17	61±17	61±17
Female, n (%)	119 443 (48%)	51 482 (48%)	85 829 (48%)
Race			
American Indian/Alaskan Native	1013 (0.5%)	444 (0.5%)	716 (0.5%)
Asian	2638 (1.2%)	1134 (1.2%)	1970 (1.2%)
Black	3971 (1.8%)	1790 (1.9%)	2855 (1.8%)
Native Hawaiian/Pacific Islander	132 (0.1%)	61 (0.1%)	91 (0.1%)
White	209 801 (95%)	90 195 (95%)	150 051 (95%)
Other/not reported	33 211 (13%)	14 296 (13%)	23 830 (13%)
ECG parameters			
Median heart rate (IQR), bpm	71 (62–83)	72 (62–83)	71 (62–83)
Median QT interval (IQR), ms	396 (368–424)	396 (368–422)	396 (368–424)
Median QTcB (IQR), ms	429 (412–451)	429 (412–451)	429 (412–451)
Median QTcF (IQR), ms	417 (402–436)	416 (401–435)	417 (402–435)
Comorbid conditions			
Cancer, n (%)	37 973 (15%)	16 462 (15%)	26 953 (15%)
Coronary artery disease, n (%)	17 911 (7.1%)	7772 (7.2%)	12 721 (7.1%)
Congestive heart failure, n (%)	32 424 (13%)	13 823 (13%)	22 888 (13%)
Diabetes with complications, n (%)	12 551 (5.0%)	5318 (4.9%)	8879 (4.9%)
Diabetes without complications, n (%)	29 983 (12%)	12 859 (12%)	21 546 (12%)
Renal disease, n (%)	21 225 (8.5%)	9015 (8.4%)	15 039 (8.4%)

bpm indicates beats per minute; IQR, interquartile range; QTcB, Bazett's heart rate-corrected QT interval; and QTcF, Fridericia's heart rate-corrected QT interval.

and human-overread QTc values for all patients/ECGs included in a subset of the validation dataset are depicted in Figure 3.

Prospective Genetic Heart Disease–Enriched Study Population

As the 2-lead DNN QTc accurately predicted 12-lead–derived QTc values in the development phase, we next sought to evaluate prospectively the 2-lead DNN QTc in consecutive patients who were referred to the Mayo Clinic Windland Smith Rice Genetic Heart Rhythm Clinic. Of the 705 patients enrolled in the study between April 2018 and February 2020, 686 (97%) had both an interpretable 12-lead ECG and mECG tracing and were thus included in the final analyses. Of note, only 8 of 705 (1.1%) eligible patients were excluded because of issues related to ambulatory signal acquisition (poor quality signal attributable to tremor/movement disorder, electronic pump, or physical limitation in 4 patients, unobtainable signal in 3 patients, and data loss in 1 patient).

Overall, this genetic heart disease–enriched study population was 57% female with an average age of 28.7±18.5 years old and median computer-/ECG technician–annotated QTc value of 446 (IQR, 425–470) (Table 2). The most common diagnoses were congenital LQTS (343 of 686 [50%]) (Table 2), followed by unaffected relatives and patients who were dismissed as normal after a comprehensive cardiovascular evaluation and were thereby collectively considered otherwise heathy controls (148 of 686 [22%]). A complete list of the diagnoses rendered within the prospective study population are summarized in Table 2.

Detection of QTc Prolongation with the Use of an AI-Enabled mECG Device

Within the genetic heart disease–enriched prospective study population, the median QTc as determined by standard-of-care (ie, ECG technician overread), QT expert overread (M.J.A.), commercial core laboratory overread, and 2-lead DNN QTc prediction was 446 ms (IQR, 425–470), 446 ms (IQR, 425–468), 434 ms (IQR, 411–460), and 444 ms (IQR, 425–467), respectively (Table 2). Similar to the validation dataset, clinically nominal mean differences were observed between the 2-lead DNN QTc and ECG technician–annotated QTc (–5.56±17.85 ms; Figure 4A), QTc expert–overread (–5.14±17.31 ms; Figure 4B), and commercial core laboratory–overread (5.82±17.38 ms; Figure 4C) QTc values derived from

Table 2. Basic Characteristics of Consecutive Patients Referred to the Mayo Clinic Windland Smith Rice Genetic Heart Rhythm Clinic With Analyzable 12-Lead ECG and mECG Device Tracings

Characteristic	Mayo Clinic Windland Smith Rice Genetic Heart Rhythm Clinic Prospective Cohort (n=686)
Average at ECG, y	28.7±18.5
Female, n (%)	390 (57)
Cardiovascular diagnoses	
None (control), n (%)	148 (22)
ABiMVPS, n (%)	11 (1.6)
ACM, n (%)	23 (3.4)
BrS/ERS, n (%)	16 (2.3)
CPVT, n (%)	32 (4.7)
Congenital LQTS, n (%)	343 (50)
LQT1, n (% LQTS cases)	147 (43)
LQT2, n (% LQTS cases)	106 (31)
LQT3, n (% LQTS cases)	55 (16)
Minor, n (% LQTS cases)	17 (5.0)
Compound, n (% LQTS cases)	10 (2.9)
Genotype negative, n (% LQTS cases)	6 (1.7)
Multisystem, n (% LQTS cases)	2 (0.6)
DCM, n (%)	5 (0.7)
Drug-induced LQTS, n (%)	4 (0.6)
HCM, n (%)	36 (5.2)
IVF, n (%)	14 (2.0)
Multiple disorders, n (%)	10 (1.4)
Other CVD, n (%)	25 (3.6)
Other GHD, n (%)	20 (2.9)
Electrocardiographic data*	
Heart rate (IQR), bpm	67 (57-75)
QT (IQR), ms	426 (398-460)
12-lead QTc (IQR), ms	446 (425-470)
12-lead expert-overread QTc (IQR), ms	446 (425-468)
12-lead core laboratory–overread QTc (IQR), ms	434 (411-460)
2-lead DNN-predicted (12-lead) QTc (IQR), ms	439 (418–463)
2-lead DNN-predicted (mECG) QTc (IQR), ms	444 (425–467)
12-lead expert-overread QTc distribution	
QTc<460 ms	456 (66)
QTc≥ 460 ms	230 (34)
QTc≥470 ms	163 (24)
QTc≥500 ms	40 (5.8)

ABiMVPS indicates arrhythmogenic bileaflet mitral valve prolapse syndrome; ACM, arrhythmogenic cardiomyopathy; BrS, Brugada syndrome; ERS, early repolarization syndrome; CPVT, catecholaminergic polymorphic ventricular tachycardia; CVD, cardiovascular disease; DCM, dilated cardiomyopathy; DNN, deep neural network; GHD, genetic heart disease; HCM, hypertrophic cardiomyopathy; IVF, idiopathic ventricular fibrillation; IQR, interquartile range; LQT1, long QT syndrome 1; LQT2, long QT syndrome 2; LQT3, long QT syndrome 3; LQTS, long QT syndrome; mECG, mobile ECG; and QTc, heart rate–corrected QT interval.

*Median values are reported for all electrocardiographic variables.

an independent risk factor for SCD and a predictor of both all-cause and cardiovascular mortality in a variety of clinical settings.[7–9] Importantly, studies from our institution and others have demonstrated that ≈1% of all individuals who receive an inpatient or outpatient 12-lead ECG have a QTc ≥500 ms.[8,23,24] Of note, this QTc threshold, when met or exceeded, carries an ≈2- to 4-fold increased risk of death[8,23,24] and serves, at least in the short term (ie, 30 days), as a powerful predictor of all-cause mortality that outperforms conventional comorbidity indices.[24]

Furthermore, in many circumstances, the development of a QTc ≥500 ms is driven, at least in part, by the presence of a potentially lethal, but highly treatable genetic condition (ie, congenital LQTS) and/or modifiable risk factors (eg, electrolyte abnormalities, use of ≥1 QTc prolonging medication, or underlying QT-agitating diseases which now includes SARS-CoV-2–mediated COVID-19).[8] Therefore, in many circumstances, the identification of substantial QTc prolongation provides an important opportunity to (1) identify vulnerable, at-risk hosts and (2) make potentially life-saving change(s) (ie, initiation of β-blockers, discontinuation of QTc-prolonging medications, or correction of hypokalemia and hypomagnesemia) needed to mitigate the risk of TdP and SCD.

However, because of the reliance on bulky 12-lead ECG systems and trained ECG technicians, the clinical settings in which the QTc can be derived and monitored are limited largely to snapshot assessments in hospitals and outpatient clinics. Of note, recent attempts to utilize unproven QTc-prolonging medication(s) in the midst of the COVID-19 pandemic, specifically hydroxychloroquine ± azithromycin, have illuminated issues surrounding how and where the QTc is monitored as the use of conventional 12-lead electrocardiography to routinely monitor the QTc in COVID-19 patients carries both a risk of additional exposure(s) (ie, ECG technicians) and the potential to tax the already limited supply of personal protective equipment.[11]

To this end, the marriage of AI/deep learning to mECG devices that can be used by the patient themselves (in hospital or at home) represents the sort of disruptive technology capable of transforming how the QTc value is obtained and utilized clinically. In the current study, an AI DNN–derived QTc was trained and validated using >1.6 million ECGs and then applied prospectively to tracings from AliveCor's FDA-approved, smart phone–enabled, mECG device. This DNN QTc was capable of detecting a QTc ≥500 ms with an AUC of 0.95 to 0.97 compared with overreads by either a QT expert genetic cardiologist (M.J.A.) or a commercial core ECG laboratory with expertise in thorough QTc studies. Of note, when the QTc threshold was lowered to ≥460 ms, the 99.9th percentile QTc value for otherwise

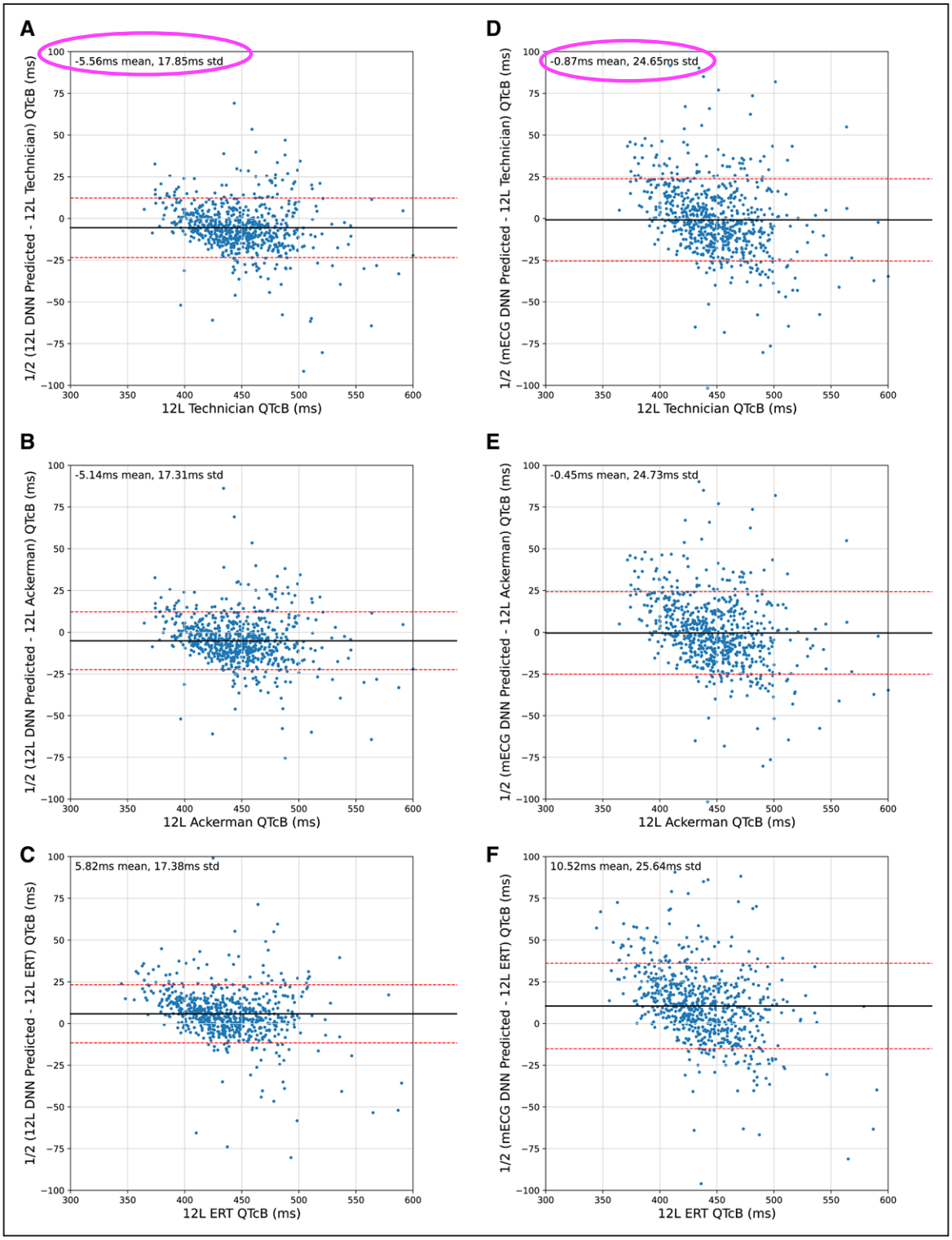

Figure 4. Bland–Altman plots of 2-lead DNN–predicted and ECG technician–, QT expert–, and commercial ECG core laboratory–overread 12-lead QTcB values in the prospective, genetic heart disease–enriched dataset.

A, Two-lead DNN– vs ECG technician–overread QTcB values from 12-lead ECG data. **B**, Two-lead DNN– vs QT expert–overread QTcB values from 12-lead ECG data. **C**, Two-lead DNN– vs eResearchTechnology–overread QTcB values from 12-lead ECG data. **D**, Two-lead DNN from mECG tracings vs ECG technician–overread 12-lead QTcB values. **E**, Two-lead DNN from mECG tracings vs QT expert–overread 12-lead QTcB values. **F**, Two-lead DNN from mECG tracings vs ERT-overread 12-lead QTcB values. Dr Ackerman was the QT expert; ERT was the commercial ECG core laboratory. DNN indicates deep neural network; ERT, eResearchTechnology; and QTcB, Bazett's heart rate–corrected QT interval.

healthy prepubertal females and males, the performance of the DNN on mECG device tracings was similar (AUC, 0.92).

Importantly, the performance characteristics of an AI-enabled mECG device to detect clinically significant QTc prolongation (ie, ≥460 ms, ≥470 ms, and ≥500

equipment and personnel costs associated with using a conventional 12-lead ECG.[33]

As QTc values calculated from an AI-enabled mECG device closely approximate those obtained from a standard 12-lead ECG, as was previously demonstrated by Karacan et al,[34] our study provides the first large-scale evidence that a viable alternative now exists. However, a prospective population-based study, potentially in combination with an AI-enhanced mECG-based screening for other SCD-predisposing disorders such as hypertrophic cardiomyopathy,[35] may be needed before universal AI-enabled mECG screening for SCD-predisposing disorders garners a formal recommendation from various medical and cardiac societies.

Importantly, the potential applications of an AI-enabled mECG device–based approach to QTc assessment and monitoring extend well beyond universal screening for the early detection of congenital LQTS. At present, QTc prolonging drugs account for ≈3% of prescriptions worldwide and QTc-related safety issues continue to represent a major reason why potentially efficacious drugs are withdrawn, restricted, or fail to reach the market.[3] With an increasingly medically-complex patient population, the number of patients receiving ≥1 QTc-prolonging medication(s) with underlying modifiable (eg, electrolyte abnormalities) and nonmodifiable (eg, diabetes, renal failure) QTc aggravating risk factors is likely to grow. Further, knowing that ≈1% of individuals of European descent and ≈8% of individuals of African descent possess a potentially proarrhythmic common variant (eg, p.Asp85Asn-KCNE1 and p.Ser1103Tyr-SCN5A, respectively) associated with a roughly 8-fold increase in the risk of drug-induced LQTS/TdP,[36,37] one can envision how widespread AI-enabled mECG device–based QTc screening and monitoring could be used to prevent QTc-related adverse drug events.

For example, the use of an AI-enabled mECG to perform an in-pharmacy QTc spot check before dispensing a known QTc-prolonging prescription(s) or facilitate the outpatient loading of known QTc-prolonging medications, such as dofetilide in relatively low-risk patients, has the ability to reduce not only drug-related adverse events/QTc polypharmacy in multimorbid and genetically-predisposed individuals, but also medical and health care costs.

Furthermore, the development of AI-enabled mECG-based QTc monitoring protocols in cooperation with pharmaceutical companies may provide a potential avenue for otherwise promising QTc-prolonging medications, particularly those in infectious disease and hematology/oncology, to progress through the drug pipeline and ultimately transition safely into clinical practice with the recommendation to include routine in-home QTc monitoring rather than be relegated to the sideline or abandoned altogether secondary to the drug's potential QT-prolonging signal.

Last, with the increasing availability of consumer mECG devices, including those attached to smartwatch/activity tracker bands, and the ability to equip these devices with AI-enabled QTc prediction algorithms, as demonstrated in the current study, a means to assess and monitor the QTc in a ubiquitous manner analogous to traditional vital signs such as heart rate, respiratory rate, temperature, and blood pressure finally exists. Although the technological advances detailed in this study further the merits of the QTc as the next vital sign,[3] this labeling is still likely to be the subject of debate. Designation as a vital sign or not, uncoupling QTc assessment from the standard 12-lead ECG with the use of AI-enabled mECG devices should provide health care professionals and patients/consumers with an important tool to aid in the identification and monitoring of individuals at risk for both congenital and acquired LQTS-mediated TdP and SCD.

Study Limitations

Although this study demonstrates the potential capability of mECG devices to revolutionize how and where the QTc is assessed and monitored in a large-scale prospective fashion, it is not without some inherent limitations. First, the 12-lead ECGs and mECG tracings, used for comparative studies within the prospective arm of the study, were not obtained exactly simultaneously. Cardiac repolarization is known to be both dynamic and positional in nature, and we anticipated that both the time differences between recordings and positional differences (supine for 12-lead versus seated, upright for mECG) might introduce some variability between the true 12-lead– and mECG-derived QTc values. However, despite this expectation, the DNN QTc from the mECG far exceeded our expectations with its tight correlation. Because of these differences, the reported performance measures likely underestimate the actual ability of a QTc DNN–equipped mECG device to approximate the "gold standard" 12-lead ECG–derived QTc values of both a QT expert genetic cardiologist and commercial core ECG laboratory.

Second, the DNN QTc was intentionally limited to 2 leads to allow for potential application to a wide range of mECG devices (eg, single-lead and multilead). In light of the recent study by Cheung et al suggesting that the acquisition of accurate and reproducible QTc values from a single-lead mECG device are possible only after obtaining multiple leads (I, II, and precordial),[16] it is possible that future iterations of the DNN QTc later trained, validated, and tested on a number of limb and precordial leads could increase further the performance of a DNN QTc–equipped mECG device. Nevertheless, given the degree of similarity between the 2-lead DNN QTc– and 12-lead ECG–derived QTc values and QTc prolongation detection performance metrics observed in