

# An Integrative Data Science Pipeline to Identify Novel Drug Interactions that Prolong the QT Interval

Tal Lorberbaum<sup>1,2,3</sup> · Kevin J. Sampson<sup>4</sup> · Raymond L. Woosley<sup>5</sup> · Robert S. Kass<sup>4</sup> · Nicholas P. Tatonetti<sup>2,3,6</sup>

Published online: 10 February 2016

© The Author(s) 2016. This article is published with open access at Springerlink.com

## Abstract

**Introduction** Drug-induced prolongation of the QT interval on the electrocardiogram (long QT syndrome, LQTS) can lead to a potentially fatal ventricular arrhythmia known as *torsades de pointes* (TdP). Over 40 drugs with both cardiac and non-cardiac indications are associated with increased risk of TdP, but drug–drug interactions contributing to LQTS (QT-DDIs) remain poorly

characterized. Traditional methods for mining observational healthcare data are poorly equipped to detect QT-DDI signals due to low reporting numbers and lack of direct evidence for LQTS.

**Objective** We hypothesized that LQTS could be identified latently using an adverse event (AE) fingerprint of more commonly reported AEs. We aimed to generate an integrated data science pipeline that addresses current limitations by identifying latent signals for QT-DDIs in the US FDA’s Adverse Event Reporting System (FAERS) and retrospectively validating these predictions using electrocardiogram data in electronic health records (EHRs).

**Methods** We trained a model to identify an AE fingerprint for risk of TdP for single drugs and applied this model to drug pair data to predict novel DDIs. In the EHR at Columbia University Medical Center, we compared the QTc intervals of patients prescribed the flagged drug pairs with patients prescribed either drug individually.

**Results** We created an AE fingerprint consisting of 13 latently detected side effects. This model significantly outperformed a direct evidence control model in the detection of established interactions ( $p = 1.62E-3$ ) and significantly enriched for validated QT-DDIs in the EHR ( $p = 0.01$ ). Of 889 pairs flagged in FAERS, eight novel QT-DDIs were significantly associated with prolonged QTc intervals in the EHR and were not due to co-prescribed medications.

**Conclusions** Latent signal detection in FAERS validated using the EHR presents an automated and data-driven approach for systematically identifying novel QT-DDIs. The high-confidence hypotheses flagged using this method warrant further investigation.

**Electronic supplementary material** The online version of this article (doi:10.1007/s40264-016-0393-1) contains supplementary material, which is available to authorized users.

✉ Nicholas P. Tatonetti  
nick.tatonetti@columbia.edu

<sup>1</sup> Department of Physiology and Cellular Biophysics, Columbia University, New York, NY, USA

<sup>2</sup> Department of Biomedical Informatics, Columbia University, 622 West 168th St. PH20, New York, NY 10032, USA

<sup>3</sup> Departments of Systems Biology and Medicine, Columbia University, New York, NY, USA

<sup>4</sup> Department of Pharmacology, Columbia University, New York, NY, USA

<sup>5</sup> AZCERT, Inc., Oro Valley, AZ, USA

<sup>6</sup> Observational Health Data Science and Informatics, New York, NY, USA

## Key Points

Drug–drug interactions that prolong the QT interval (QT-DDIs) can lead to potentially fatal arrhythmias but remain poorly characterized.

We developed an integrative data science pipeline that combines mining for latent QT-DDI signals in the US FDA Adverse Event Reporting System (FAERS), and retrospective analysis of electrocardiogram laboratory results in electronic health records, at Columbia University Medical Center.

Using latent evidence of long QT syndrome to detect QT-DDIs in FAERS significantly outperformed use of solely direct evidence of this adverse event in the detection of established interactions. The pipeline significantly enriched for novel QT-DDIs and identified eight novel interactions that warrant experimental validation.

## 1 Introduction

Long QT syndrome (LQTS) is a genetic or acquired change in the electrical activity of the heart that can increase the risk of *torsades de pointes* (TdP), a dangerous ventricular tachycardia that can lead to sudden cardiac death [1]. Diagnosed using an electrocardiogram (ECG), LQTS is characterized by a prolonged QT interval and represents an abnormally increased cardiac action potential duration. While the link between QT prolongation and TdP is complex and involves the interplay of multiple factors, a QT interval >500 ms (vs. a normal range of 350–440 ms) is nonetheless considered a significant risk for arrhythmogenesis [2].

Since the first reports of TdP in the 1960s [3], mutations in 13 genes coding for cardiac ion channels and their associated proteins have been found to play roles in LQTS [1, 4–6]. Congenital LQTS can result from mutations that disrupt the  $I_{Ks}$ ,  $I_{Kr}$ , or  $I_{Na}$  ion currents; however, the acquired form of LQTS (which is often drug-induced) is almost exclusively due to block of the human ether-à-go-go-related gene (hERG) channel (*KCNH2*), which plays a role in the  $I_{Kr}$  delayed rectifier potassium current responsible for ventricular repolarization [3]. Drug-induced inhibition of  $I_{Kr}$  was first discovered for the antiarrhythmic quinidine [7], and since then over 40 drugs with both cardiac and non-cardiac indications have been found to possess either a known, possible, conditional, or congenital

link to dangerously prolonging the QT interval [8]. Terfenadine (an allergy medication) and cisapride (used to treat acid reflux) were withdrawn from the market in 1997 and 2000, respectively, for prolonging the QT interval [9], and risk of TdP is now the second leading cause for approved drug withdrawal [2].

Drug–drug interactions (DDIs) such as those between methadone (an analgesic) and quetiapine (an antipsychotic) have also been reported to increase the risk for TdP [10]. Despite the increasingly comprehensive resources available to clinicians for linking single drugs to TdP, little remains known about DDIs (QT-DDIs). We define a QT-DDI as a measurable change in effect (QT interval duration) for a drug pair compared with the effect observed for either drug alone. This includes both pharmacokinetic interactions (such as the increased plasma concentrations of methadone in patients also taking quetiapine [10]), as well as pharmacodynamic interactions. While the FDA has required clinical studies to assess the effects of drug interactions, it is intractable to prospectively evaluate every possible drug combination. With DDIs thought to play a role in upwards of 17 % of adverse events (AEs), and an increasingly aging population taking multiple drugs concurrently [11, 12], there is a pressing need for methods to identify potential interactions.

Molecular mechanism-based approaches such as biological network analysis have been previously used to prioritize drugs with molecular links to LQTS genes, but they remain limited to known drug targets and often only apply to individual drugs [6]. More recent work using machine learning on network data can overcome the requirement for known targets [13]; however, this approach has only been validated for individual drugs.

Observational healthcare datasets such as the US FDA Adverse Event Reporting System (FAERS) and electronic health records (EHRs) provide invaluable resources for adverse event prediction, but their use is tempered by multiple limitations. Spontaneous reporting systems such as FAERS are known to suffer from both reporting bias and sampling variance [14], and methods for mining FAERS traditionally rely on direct evidence between a drug exposure and AE (i.e. the number of reports with the drug and AE co-mentioned). While methods have been developed to limit high false positives by correcting for unsubstantiated drug–AE signals [15], this leads to a tradeoff between reducing false positive rates and the ability to actually detect AEs. Direct detection of AEs falters in the prediction of DDIs, where reporting numbers are often lower than for single drugs and unanticipated or unexpected events with no understood molecular explanation can go unreported. A number of advances have been made in the field, including the observation that additive baseline models tend to outperform multiplicative ones [16] and that

case reports can be combined with mechanistic information such as shared cytochrome P450 (CYP) metabolism to develop more sophisticated triage algorithms [17]. Nonetheless, most DDI signal detection algorithms have had limited success [18–20]. Additionally, AE detection in EHRs can be challenging as such data are often complex, inaccurate, and missing [21]. While use of either dataset alone can thus be problematic for QT-DDI detection, integration of these two sources using data science offers an opportunity for improved performance.

In previous work, we demonstrated that a novel signal detection algorithm could be used for detecting latent signals of previously unknown DDIs for eight severe AE classes [22, 23]. Importantly, each individual drug in the drug pair had no previously known connection to the AE class of interest. In this study, we introduce an updated pipeline called DIPULSE (Drug Interaction Prediction Using Latent Signals and EHRs) that uses latent signal detection in FAERS to generate an AE fingerprint for LQTS. This AE fingerprint—trained on individual drugs with a known link to prolonging the QT interval—represents a profile of more commonly reported side effects that together are highly predictive of underlying QT interval prolongation. We apply this fingerprint model to an independent test data set of drug pairs to predict new QT-DDIs where neither drug alone has a known association to this phenotype. We validate these predictions using ECG laboratory results in EHRs.

## 2 Methods

A graphical overview of DIPULSE can be found in Fig. 1. The individual steps of the pipeline corresponding to each panel of the figure are described in detail below. Briefly, we used AE reporting frequencies for individual drugs to identify an AE fingerprint for increased risk of TdP. We then apply this model to a test data set of AE reporting frequencies for drug pairs. We filtered for high-confidence predictions and proceeded to validate these putative QT-DDIs in the EHR by comparing the QTc (heart rate-corrected QT) intervals of patients prescribed the flagged drug pair with patients prescribed either drug alone. Finally, we perform a confounder analysis to remove any associations that can be explained by co-prescribed medications, and generated a final candidate list of novel QT-DDIs.

In developing the pipeline, our rationale was to prioritize high precision over high recall to obtain a final list of high-confidence interactions; therefore, the choices we made in designing the filtering steps described below reflect this conservative approach. We implemented the method using Python 2.7.9 and R 3.1.0.

### 2.1 Primary Data Sources

We downloaded a snapshot of the FAERS database containing 1,851,171 reports (corresponding to the first quarter of 2004 to the first quarter of 2009). Each report in FAERS contains the drugs prescribed to the patient, the drug indications, and the observed AEs. We included suspected, interacting, and concomitant drugs on the reports.

As positive controls, we downloaded a list of 180 drugs with known ( $n = 47$ ), possible ( $n = 75$ ), conditional ( $n = 31$ ), or congenital ( $n = 27$ ) risk of TdP from CredibleMeds, an online compendium of drugs associated with LQTS [8]. We also obtained a list of 2856 critical and significant DDIs from the Veteran Affairs Hospital [24].

To validate our DDI predictions, we used EHR data from Columbia University Medical Center (CUMC). In addition to patient demographics, drugs prescribed, and diagnosis codes, we also used QTc (heart rate-corrected QT interval) values obtained from ECG laboratory results. The study was approved by the CUMC Institutional Review Board.

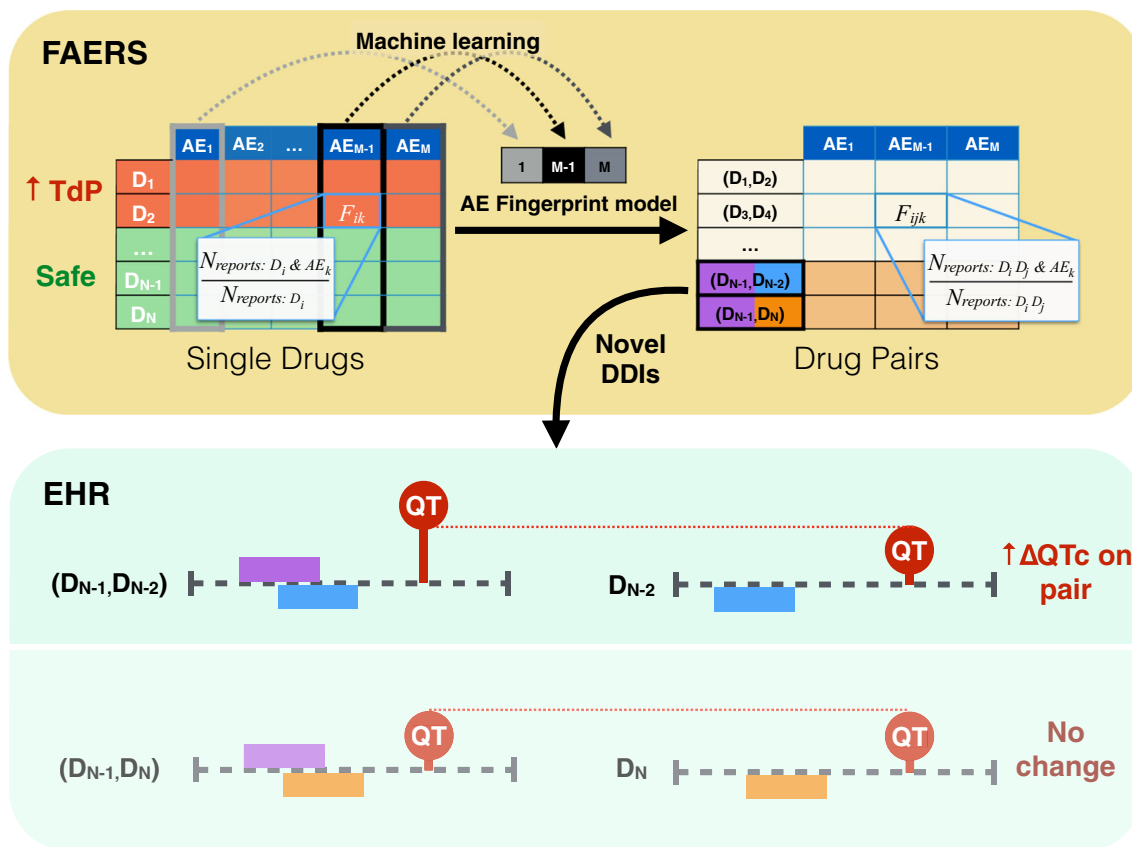
### 2.2 Generating Adverse Event (AE) Reporting Frequency Tables

We pre-processed the reports from FAERS to generate the intermediate AE reporting frequency tables in the OFFSIDES (single drug) and TWOSIDES (drug pair) databases [25]. OFFSIDES and TWOSIDES were created by training propensity score matching models to match patients exposed to a single drug or drug pair to unexposed controls on the basis of co-prescribed medications and drug indications; an advantage of this approach is that only patients for whom controls could be matched are used for drug safety prediction [25].

An intermediate step in this process is the assembly of AE frequency reporting tables for both single drugs and drug pairs, as seen in Fig. 1, with each row representing a drug and each column representing one of the AEs in FAERS. For single drugs, the value at a given row and column represents the frequency of reporting  $F_{ik}$ , defined as the fraction of reports for drug  $i$  containing the AE  $k$ . Similarly, for drug pairs, the reporting frequency  $F_{ijk}$  corresponds to the fraction of reports for drug pair  $(i, j)$  containing the AE  $k$ . We used the former matrix to train the fingerprint model, and the latter for DDI prediction.

### 2.3 Training AE Fingerprint Model

We used the AE reporting frequencies ( $F_{ik}$ ) in the frequency table for single drugs as features to train a logistic regression classifier. The binary classifier models the log odds ratio of a drug prolonging the QT interval as a linear



**Fig. 1** Overview of DIPULSE pipeline, which combines mining of FAERS and EHRs to flag novel QT-prolonging DDIs. FAERS: We generate an AE reporting frequency table (dimensions,  $N$  drugs by  $M$  AEs) for single drugs in FAERS. The value at a row and column represents the fraction of reports for drug  $i$  containing AE  $k$  ( $F_{ik}$ ). We label a drug as a positive example (shown in red) if it has a known risk of TdP (obtained from <http://www.CredibleMeds.org>). All drugs not found in CredibleMeds were labeled as negative examples (shown in green). We use machine learning to generate an AE fingerprint model that identified the most predictive subset of features (AE reporting frequencies,  $F_{ik}$ ) as latent evidence for predicting whether a drug does or does not prolong the QT interval (gray boxes). We then apply this fingerprint model to an independent test data set consisting of a matrix (with AE reporting frequencies  $F_{ijk}$ ) for drug pairs. We send pairs receiving high classifier probabilities (but where neither individual drug is known to prolong the QT interval) for EHR

validation (in this case pairs  $(D_{N-1}, D_{N-2})$  [purple-blue] and  $(D_{N-1}, D_N)$  [purple-orange]). EHR: We validate putative interactions using electrocardiogram laboratory results in the EHRs by determining whether patients prescribed a predicted interacting drug pair had increased QTc intervals compared with patients taking either drug alone. In this example, patients prescribed the drug pair  $(D_{N-1}, D_{N-2})$  have a significantly increased QT interval compared with patients on either drug alone. This is not observed for drug pair  $(D_{N-1}, D_N)$  so it is filtered out. Finally, we performed a confounder analysis to confirm that the significant increase observed in QTc interval is not due to other co-prescribed medications. DIPULSE Drug Interaction Prediction Using Latent Signals and EHRs, EHRs electronic health records, FAERS FDA Adverse Event Reporting System, DDIs drug–drug interactions, AE adverse event, TdP torsades de pointes, QTc heart rate-corrected QT interval

combination of each AE reporting frequency in the model multiplied by a weight (known as a  $\beta$  coefficient); depending on the probability threshold set, a drug above the threshold is classified as increasing the risk of TdP, and a drug below the threshold is classified as safe. Training the model requires both positive and negative examples. As positive examples, we used the subset of the 47 drugs with a known risk of TdP in CredibleMeds that were also in FAERS ( $n = 23$ ). As negative controls, we selected all drugs in FAERS that did not appear in CredibleMeds (i.e. have no known, possible, conditional, or congenital risk of TdP;  $n = 530$ ).

Because the number of features (11,305 AEs) is much greater than the number of examples (553 drugs), overfitting of the model to the training data is a concern. To ensure the model generalized to our test data set (drug pairs), we reduced the number of features by using L1 (lasso) regularization [26]. Unlike L2 (ridge) regularization (which penalizes the squares of the feature weights), L1 regularization penalizes their absolute values and is therefore preferred because it results in sparse models (i.e. most of the feature weights will be driven to zero). We generated five models, each of which contained between 5 and 20 features obtained by varying the regularization

strength for the given model. We evaluated these models using 10-fold cross-validation, and then re-fit the classifier using only the selected features. The features for each of these models constitute an *AE fingerprint* that represents latent evidence for QT interval prolongation.

As a control, we generated a logistic regression model built solely using direct evidence of QT interval prolongation (standardized Medical Dictionary for Regulatory Activities [MedDRA] query for ‘Torsade de Pointes/QT prolongation’). There were only six AEs corresponding to QT interval prolongation or TdP (electronic supplementary Table 1), and therefore feature selection was not necessary.

#### 2.4 Predicting Novel Drug–Drug Interactions (DDIs) Using the Fingerprint Model

We next applied the QT fingerprint model to an independent test data set consisting of the AE reporting frequencies ( $F_{ijk}$ ) in the frequency table for drug pairs. The model outputs a probability for a given drug pair to prolong the QT interval. We assessed model performance using two references. In the first, we labeled each drug pair containing a drug known to increase the risk of TdP as a positive example. While these may not be bonafide DDIs, they demonstrate the ability of the fingerprint model to ‘re-discover’ drugs known to prolong the QT interval within the drug pair data. We used this validation to select the optimal fingerprint model. We also performed an additional validation using a list of critical and significant DDIs from the Veteran Affairs Hospital. For both of these evaluations, we compared the performance of the ‘latent’ AE fingerprint model with the ‘direct evidence’ control model using DeLong’s test [27].

To obtain a candidate list of novel DDIs predicted by the fingerprint model, we first removed all drug pairs containing a drug in the CredibleMeds list. We then filtered for all novel predictions found at a classifier probability below a 4 % false positive rate according to the CredibleMeds evaluation. We chose this false positive rate threshold by modeling the expected increase in false discovery rate as a function of false positive rate (see electronic supplementary Fig. 1 and accompanying legend for a description of the analysis). Finally, we removed drug pairs that would receive high classifier scores regardless of the features used in the model by generating 100 logistic regression models using randomly chosen features and estimating empirical  $p$  values for each drug pair. We removed any drug pairs receiving an empirical  $p$  value  $\geq 0.01$ .

#### 2.5 Validating Novel DDIs Using Electronic Health Records

While the novel DDIs predicted using our signal detection algorithm each contain latent evidence for prolonging the

QT interval, ECG values in EHRs allow us to retrospectively evaluate the effect of these drug pairs (our cases) on QT interval duration compared with either drug alone (our controls). Because QT interval durations differ between males and females [28], we evaluated the effects of a given drug pair on each sex separately.

To obtain cases, we selected patients at New York-Presbyterian Hospital/Columbia University Medical Center who were prescribed each drug in a given drug pair within a 7-day period. Patients were also required to have an ECG lab—and corresponding QTc (heart rate-corrected QT interval)—within 36 days of the second drug prescription. We chose this limit to minimize the potential for new confounding drug prescriptions or interventions; additionally, because follow-up visits are often scheduled in units of weeks, we allowed for 5 weeks plus 1 day for laboratory tests to be performed [22]. For patients with multiple QTc values within this time period, we used the maximum value.

To obtain controls, we selected patients taking whichever individual drug in the pair yielded the greatest median QTc within a 36-day period from drug prescription; we call this drug the ‘control’ drug. We then compared QTc values between cases and controls and assessed significance using a Mann–Whitney  $U$  test, correcting for multiple hypothesis testing using Bonferroni’s method.

In order to demonstrate that the predictions being sent for EHR validation were enriched for drug interactions that actually prolonged the QT interval, we ran the above EHR case-control analysis on a set of drug pairs equal in number to that generated by the latent signal detection but randomly chosen from the frequency table for drug pairs. To generate a more representative comparison, we required that each pair be comprised of a randomly chosen drug paired with a ‘control’ drug (i.e. the drug with the greatest QTc interval alone from the latent evidence pairs). Additionally, to ensure equivalent statistical power we matched the number of patients in the case groups of the randomly chosen pairs to the case group sizes of the pairs prioritized by the latent signal detection. We counted the number of random pairs that had significant increases in QT interval, and repeated this sampling procedure 1000 times to build an empirical distribution of how many significant results would be expected after EHR analysis by chance alone.

Finally, we adjusted for confounders by confirming that the elevated QTc interval on the drug pair was not due to other co-prescribed medications. For each of our sets of cases (patients on a given drug pair) and controls (patients on an individual drug in the pair), we identified possible confounder drugs by counting the number of exposures to each drug prescribed up to 36 days prior. We evaluated each potential confounder by confirming that it was correlated both with the exposure condition and with QTc



values. For the former, we determined whether the covariate was more likely to be prescribed with the drug pair compared with the single drug using a Fisher's exact test; for the latter, we compared the QTc values for patients exposed to the covariate versus those unexposed using a Mann–Whitney  $U$  test. Both of these evaluations were performed using a Bonferroni correction for multiple hypothesis testing. We collected all drug covariates that passed these two requirements and assessed their significance (for males and females separately) using an analysis of covariance (ANCOVA). To obtain the final list of validated novel DDIs, we only kept those results (drug pairs for a given sex) receiving significant ANCOVA  $p$  values ( $p < 0.05$ ) for the DDI.

### 3 Results

#### 3.1 QT Fingerprint Model Significantly Outperforms Model Built Using Only Direct Evidence

Of the five fingerprint models evaluated, we found that the model containing 13 features achieved the best performance for drug pair data (area under the curve [AUC] = 0.69 using pairs containing a known CredibleMeds drug) (electronic supplementary Fig. 2); see Table 1 for the list of features that constitute the QT AE fingerprint. Importantly, the QT fingerprint model significantly outperformed the model built using direct evidence, as evaluated by both the CredibleMeds ( $p = 1.62\text{E}-3$ ) and Veteran Affairs ( $p = 5.22\text{E}-10$ ) drug pair standards (Fig. 2). We also compared these models to a previously published additive baseline model for predicting DDIs [19] and found that the latent evidence model outperformed this

**Table 1** Features in QT fingerprint model

Adverse event	Beta
Drug interaction	0.52
Atrial fibrillation	0.50
Arrhythmia	0.29
Electrocardiogram QT prolonged	0.28
Tachycardia ventricular	0.28
Asystole	0.27
Torsades de pointes	0.24
Completed suicide	0.21
Rhabdomyolysis	0.17
Agitation	0.07
Drug ineffective	-0.36
Hemorrhage	-0.25
Myocardial infarction	-0.18

method (electronic supplementary Fig. 3; CredibleMeds:  $p < 2.2\text{E}-16$ ; Veteran Affairs:  $p = 2.18\text{E}-11$ ). After filtering using both empirical  $p$ -values and the 4 % false positive rate cutoff, we obtained 889 putative novel DDIs to be validated in the EHR.

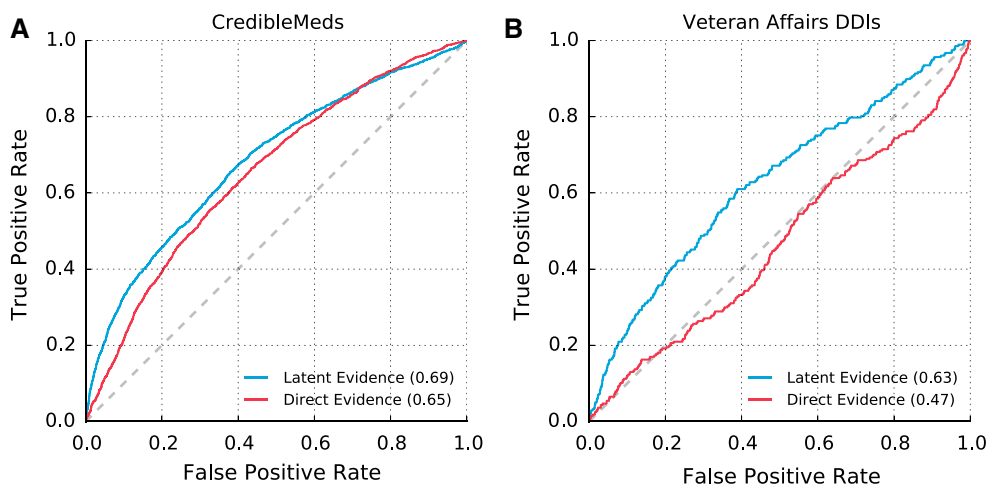
#### 3.2 EHR Validation and Confounder Analysis Confirms Novel Drug Interactions Prolonging the QT Interval

Our EHR evaluation yielded 49 results (drug pairs for males and/or females) that had significantly increased QTc intervals on the drug pair compared with either drug alone (electronic supplementary Fig. 4). This number of results was significantly greater than for randomly generated input to the EHR validation ( $p = 0.01$ ) (electronic supplementary Fig. 5). After confounder analysis, we obtained ten results (corresponding to eight distinct drug pairs) which represented validated novel DDIs that increase the risk of acquired LQTS (Table 2).

The greatest increase in median QTc (30 ms) was for octreotide (a somatostatin analog used to lower growth hormone levels) and lactulose (administered to treat constipation) compared with octreotide alone ( $p = 2.48\text{E}-4$ ) in males, and males prescribed this pair were 2 times as likely to have a QTc interval  $\geq 500$  ms (electronic supplementary Table 2). For females, co-prescription of mupirocin and vancomycin was associated with a 20 ms increase in median QTc compared with vancomycin alone ( $p = 1.3\text{E}-4$ ); females prescribed the pair were 1.7 times as likely to have a QTc interval  $\geq 500$  ms. A complete list of retrospectively validated interactions and the number of patients in the case and control groups can be found in Table 2.

### 4 Discussion

Drug-induced LQTS and its potential for fatal arrhythmia (TdP) make this disorder of critical importance both to drug discovery and pharmacovigilance. Indeed, an important step in the drug development process is confirming that the lead compound does not significantly block the hERG channel that contributes to TdP [2]. However, the inability to prospectively identify this risk is highlighted by the increasing number of drugs found to increase the risk for TdP [8]. Even more difficult to detect are DDIs that contribute to LQTS, as experimental evaluation of all possible QT-DDIs is not feasible and traditional methods for mining observational data are poorly equipped to handle low reporting numbers and high false positive rates. Because analyses of spontaneous reporting systems (such as FAERS) and EHRs alone have many limitations, in this study we developed an integrative pipeline that incorporates multiple dimensions of observational data to



**Fig. 2** Receiver operating characteristic curves for adverse event fingerprint model and direct evidence control. **a** Model validation was performed by labeling drug pairs containing a drug with known increased risk of TdP as positive examples. We compared the performance of a model built using latent evidence (AE fingerprint model) to a control model using only direct evidence of QT

prolongation. **b** A second evaluation performed using a list of critical and significant DDIs from the Veteran Affairs Hospital in Arizona. For both validations, the AE fingerprint model significantly outperformed the model built solely with direct evidence. Area under the curve (AUC) is indicated in parentheses. *DDIs* drug–drug interactions, *TdP* torsades de pointes, *AE* adverse event

**Table 2** List of novel DDIs generated by DIPULSE and validated in the EHR

Drug 1	Drug 2	Control	Sex	Estimate	<i>p</i> value	Median QTc cases	Median QTc controls	<b><math>\Delta</math>QTc (ms)</b>	No. of cases	No. of controls
Octreotide	Lactulose	Octreotide	M	74.8	2.48E−04	485	455	<b>30</b>	333	603
Mupirocin	Vancomycin	Vancomycin	F	54.5	1.30E−04	476	456	<b>20</b>	810	10,165
Metoprolol	Fosphenytoin	Metoprolol	M	40.9	2.19E−07	462	444	<b>18</b>	549	24,717
<i>N</i> -Acetylcysteine	Vancomycin	Vancomycin	M	17.4	3.74E−04	469	453	<b>16</b>	2633	9789
Cefazolin	Meperidine	Cefazolin	F	27.6	1.29E−05	455	441	<b>14</b>	1025	9172
Cefazolin	Meperidine	Cefazolin	M	18.2	8.97E−08	452	440	<b>12</b>	2110	10,013
Ceftriaxone	Lansoprazole	Ceftriaxone	M	39.1	4.21E−09	458	446	<b>12</b>	934	5734
<i>N</i> -Acetylcysteine	Morphine	<i>N</i> -Acetylcysteine	M	12.1	3.19E−02	460	451	<b>9</b>	2525	6046
Meperidine	Vancomycin	Vancomycin	F	34.6	4.77E−03	464	457	<b>7</b>	1105	9894
<i>N</i> -Acetylcysteine	Morphine	<i>N</i> -Acetylcysteine	F	22.3	7.93E−04	459	455	<b>4</b>	1900	4803

The bolded column highlights the  $\Delta$ QTc for a given drug pair

*DDIs* drug–drug interactions, *DIPULSE* Drug Interaction Prediction Using Latent Signals and EHRs, *EHRs* electronic health records, *M* male, *F* female, *QTc* corrected QT interval

allow for identification of true QT-DDI signals. We demonstrated the applicability of this data science approach by identifying latent signals of LQTS in FAERS and retrospectively validating these novel QT-DDI predictions using EHRs. Comparing our AE fingerprint model for QT prolongation with a direct evidence control demonstrated that latent evidence of drug-induced LQTS in FAERS can outperform direct evidence in the detection of established interactions.

While most drugs prolong the QT interval by interacting with the hERG channel, the clinical data used in this analysis do not permit a mechanistic explanation for the synergistic effects of the identified DDIs. Electrophysiology experiments to directly assay the effect of individual

drugs and drug pairs on hERG channel activity can provide further evidence for, and molecular mechanisms of, these effects [2]. Importantly, QTc correction formulas still used today were developed in 1920 and are known to be inaccurate when heart rate changes occur outside the baseline range used to define the formula [2]. As such, drugs that do not directly affect ventricular repolarization but instead alter the patient’s heart rate may be incorrectly attributed to increasing the QTc. It is possible that some of the interactions we identified were confounded by this complexity. This limitation highlights the need for experimental validation of our QT-DDI predictions to directly assess hERG channel block or effects on other ion channels.

In considering the features selected for the QT fingerprint model (Table 1), many of the features are expected, including ECG QT prolonged, TdP, arrhythmia, and even rhabdomyolysis, as this condition can be induced by hypokalemia which also predisposes patients to LQTS [3, 29]. However, other features are more unexpected, including completed suicide and agitation. One explanation for the selection of these features is that a number of the positive control drugs (including chlorpromazine, citalopram, and haloperidol) from CredibleMeds are indicated for conditions characterized by agitation and suicidality. We purposefully did not manually exclude any features on the basis of wanting to develop a purely data-driven model that is not limited to current clinical knowledge of (non-cardiac) side effects that are highly predictive of underlying QT prolongation; however, because of the relatively small number of positive controls (predominantly with psychological, antibacterial, and anti-arrhythmic indications), we acknowledge the possibility that inclusion of these features may be driven by the indications of the positive controls rather than their effects on QT prolongation.

Our EHR control analysis (while limited to comparing the number of significant findings prior to confounder adjustment) demonstrated that our method significantly enriched for QT-prolonging drug pairs compared with random selection. Approximately 4 % of pairs investigated 'passed' the EHR validation prior to confounder analysis. Of the 889 pairs flagged by latent signal detection in FAERS, 251 of these pairs (28 %) had no patients prescribed the pair in our EHR and therefore could not be evaluated. The other pairs that did not pass validation were either prescribed at low numbers (and could therefore be false negatives due to insufficient statistical power) or may be false positives from FAERS. While we believe the 7-day window between drug prescriptions represents a fairly stringent cutoff for confirming that patients were taking both drugs in a pair concurrently, challenges in estimating the duration of treatment in EHRs also has implications for accurately selecting all of the desired patients in the case group. Follow-up analyses could repeat the EHR analysis at additional institutions to both replicate these results and investigate drug pairs that could not be validated in our EHR.

Because our EHR analysis filtered for interactions (pairs with significantly greater QT interval prolongation compared with either drug alone), a final potential explanation for pairs identified in FAERS that could not be validated in the EHR is that the highlighted pair represented a novel single drug that prolongs the QT interval. While we limited the scope of this study to identifying QT-DDIs, resources such as CredibleMeds continue to use signals in FAERS as part of their evidence portfolio for the inclusion and

removal of new individual drugs to/from the database [30]. An important challenge to overcome in the evaluation of potential QT-prolonging single drugs in the EHR would be the identification of proper controls; propensity score matching offers one opportunity for addressing this [25].

We note that the AE reporting frequencies for drug pairs ( $F_{ijk}$ ) cannot intrinsically distinguish between interactions and single-drug effects from either drug  $i$  or drug  $j$  alone. To distinguish between these two explanations for a drug pair receiving a high classifier score, it is therefore necessary to remove all single-drug effects (attributable to not only a known but also possible, conditional, or congenital link to TdP). CredibleMeds uses a number of signals (including FAERS, laboratory and clinical research reports, and clinical trial data) to populate their database [30]. Thus, while it is possible that CredibleMeds does not contain complete coverage of all QT-prolonging drugs, we believe it represents the most reliable resource for justifying removal of drug pairs that receive high scores due to the effects of single drugs. Application of our method to other AEs would therefore necessitate a similarly reliable resource of single-drug effects to minimize the possibility of falsely labeled interactions. While our confounder analysis investigated the effects of co-prescribed medications in addition to the drug pair of interest, follow-up work could also incorporate the dose of each drug in the pair as a potential confounder.

While cases of drug-induced LQTS have predominantly been found to be due to blocking of  $I_{Kr}$ , we do not discount the possibility for other potential mechanisms of these QT-DDIs. Biological network analysis [6, 13] may be useful for identifying other proteins, in addition to or instead of hERG, that are affected by these drugs.

## 5 Conclusions

In this study we have developed and validated DIPULSE, an automated integrated pipeline for flagging novel DDIs that can prolong the QT interval using data from both spontaneous reporting systems (FAERS) and EHRs. By identifying latent signals of QT interval prolongation, this method is able to overcome some of the limitations in mining for DDIs. The method significantly outperforms DDI detection solely using direct evidence for QT prolongation in the detection of established interactions. This study highlights the utility of integrative data science approaches in mining for new and potentially fatal DDIs.

**Acknowledgments** The authors would like to thank Sam Roe for thoughtful discussions about the manuscript.

**Author contributions** Tal Lorberbaum and Nicholas P. Tatonetti designed and performed the research, and analyzed the data. Kevin J.



Sampson, Robert S. Kass, and Raymond L. Woosley contributed new reagents/analytical tools. Tal Lorberbaum and Nicholas P. Tatonetti wrote the article.

### Compliance with Ethical Standards

This study was approved by the CUMC Institutional Review Board.

**Funding** Tal Lorberbaum is supported by a training grant from the National Institute of General Medical Sciences (T32GM082797); Tal Lorberbaum and Nicholas P. Tatonetti are supported by the National Institute of General Medical Sciences (R01GM107145); and Kevin J. Sampson and Robert S. Kass are supported by the National Institute of Health (5R01GM109762-02).

**Conflict of interest** Tal Lorberbaum, Kevin J. Sampson, and Robert S. Kass declare that they have no conflicts of interest. Nicholas P. Tatonetti is a paid advisor to Advera Health, Inc.; he declares no conflicts of interest. Raymond L. Woosley is an uncompensated officer of the non-profit organization AZCERT.org, which is supported by FDA HHSF223201400189C and which maintains the website <http://www.CredibleMeds.org> utilized in this study; he declares no conflicts of interest.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits any noncommercial use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

### References

- Roden DM. Clinical practice. Long-QT syndrome. *N Engl J Med*. 2008;358(2):169–76.
- Fermini B, Fossa AA. The impact of drug-induced QT interval prolongation on drug discovery and development. *Nat Rev Drug Discov*. 2003;2(6):439–47.
- Kannankeril P, Roden DM, Darbar D. Drug-induced long QT syndrome. *Pharmacol Rev*. 2010;62(4):760–81.
- Marx SO, Kurokawa J, Reiken S, Motoike H, D'Armiento J, Marks AR, et al. Requirement of a macromolecular signaling complex for beta adrenergic receptor modulation of the KCNQ1-KCNE1 potassium channel. *Science*. 2002;295(5554):496–9.
- Moss AJ, Kass RS. Long QT syndrome: from channels to cardiac arrhythmias. *J Clin Invest*. 2005;115(8):2018–24.
- Berger SI, Ma'ayan A, Iyengar R. Systems pharmacology of arrhythmias. *Sci Signal*. 2010;3(118):ra30.
- Roden DM, Woosley RL, Primm RK. Incidence and clinical features of the quinidine-associated long QT syndrome: implications for patient care. *Am Heart J*. 1986;111(6):1088–93.
- Woosley RL, Romero K. Assessing cardiovascular drug safety for clinical decision-making. *Nat Rev Cardiol*. 2013;10(6):330–7.
- Woosley RL, Chen Y, Freiman JP, Gillis RA. Mechanism of the cardiotoxic actions of terfenadine. *JAMA*. 1993;269(12):1532–6.
- Uehlinger C, Crettol SV, Chassot P, Brocard M, Koeb L, Brauwand-Amey M, et al. Increased (R)-methadone plasma concentrations by quetiapine in cytochrome P450s and ABCB1 genotyped patients. *J Clin Psychopharmacol*. 2007;27(3):273–8.
- Pirmohamed M, James S, Meakin S, Green C, Scott AK, Walley TJ, et al. Adverse drug reactions as cause of admission to hospital: prospective analysis of 18,820 patients. *BMJ*. 2004;329(7456):15–9.
- Hajjar ER, Cafiero AC, Hanlon JT. Polypharmacy in elderly patients. *Am J Geriatr Pharmacother*. 2007;5(4):345–51.
- Lorberbaum T, Nasir M, Keiser MJ, Vilar S, Hripcsak G, Tatonetti NP. Systems pharmacology augments drug safety surveillance. *Clin Pharmacol Ther*. 2015;97(2):151–8.
- Bate A, Evans SJW. Quantitative signal detection using spontaneous ADR reporting. *Pharmacoepidemiol Drug Saf*. 2009;18(6):427–36.
- Szarfman A, Machado SG, O'Neill RT. Use of screening algorithms and computer systems to efficiently signal higher-than-expected combinations of drugs and events in the US FDA's spontaneous reports database. *Drug Saf*. 2002;25(6):381–92.
- Juhlin K, Soeria-Atmadja D, Thakrar B, Norén GN. Evaluation of statistical measures for adverse drug interaction surveillance. *Pharmacoepidemiol Drug Saf*. 2014;23(S1):294–5.
- Strandell J, Caster O, Hopstadius J, Edwards IR, Norén GN. The development and evaluation of triage algorithms for early discovery of adverse drug interactions. *Drug Saf*. 2013;36(5):371–88.
- DuMouchel W, Pregibon D. Empirical Bayes screening for multi-item associations. In: Proceedings of the 7th ACM SIGKDD international conference on knowledge discovery and data mining. New York: Association for Computing Machinery (ACM); 2001. p. 67–76.
- Norén GN, Sundberg R, Bate A, Edwards IR. A statistical methodology for drug–drug interaction surveillance. *Stat Med*. 2008;27(16):3057–70.
- Harpaz R, Chase HS, Friedman C. Mining multi-item drug adverse effect associations in spontaneous reporting systems. *BMC Bioinform*. 2010;11(Suppl 9):S7.
- Hripcsak G, Albers DJ. Next-generation phenotyping of electronic health records. *J Am Med Inform Assoc*. 2013;20(1):117–21.
- Tatonetti NP, Denny JC, Murphy SN, Fernald GH, Krishnan G, Castro V, et al. Detecting drug interactions from adverse-event reports: interaction between paroxetine and pravastatin increases blood glucose levels. *Clin Pharmacol Ther*. 2011;90(1):133–42.
- Tatonetti NP, Fernald GH, Altman RB. A novel signal detection algorithm for identifying hidden drug–drug interactions in adverse event reports. *J Am Med Inform Assoc*. 2012;19(1):79–85.
- Olvey EL, Clauschee S, Malone DC. Comparison of critical drug–drug interaction listings: the Department of Veterans Affairs medical system and standard reference compendia. *Clin Pharmacol Ther*. 2010;87(1):48–51.
- Tatonetti NP, Ye PP, Daneshjou R, Altman RB. Data-driven prediction of drug effects and interactions. *Sci Transl Med*. 2012;4(125):125ra31.
- Tibshirani R. Regression shrinkage and selection via the Lasso. *J R Stat Soc Ser B Stat Methodol*. 1996;58(1):267–88.
- Robin X, Turck N, Hainard A, Tiberti N, Lisacek F, Sanchez J-C, et al. pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinform*. 2011;12(1):77.
- Rautaharju PM, Zhou SH, Wong S, Calhoun HP, Berenson GS, Prineas R, et al. Sex differences in the evolution of the electrocardiographic QT interval with age. *Can J Cardiol*. 1992;8(7):690–5.
- Vanholder R, Sever MS, Ereik E, Lameire N. Rhabdomyolysis. *J Am Soc Nephrol*. 2000;11(8):1553–61.
- Woosley RL, Romero KA. QTdrugs List. Available at: <http://www.CredibleMeds.org>. Accessed 22 Dec 2015.