The Tribune-Columbia collaboration resulted in two published scientific papers, including this one in the Journal of the American College of Cardiology, a leading cardiovascular journal.

JOURNAL OF THE AMERICAN COLLEGE OF CARDIOLOGY
© 2016 BY THE AMERICAN COLLEGE OF CARDIOLOGY FOUNDATION
PUBLISHED BY ELSEVIER

VOL. 68, NO. 16, 2016
ISSN 0735-1097/$36.00
http://dx.doi.org/10.1016/j.jacc.2016.07.761

Coupling Data Mining and Laboratory Experiments to Discover Drug Interactions Causing QT Prolongation

Tal Lorberbaum, MA,[a,b] Kevin J. Sampson, PhD,[c] Jeremy B. Chang, PhD,[b] Vivek Iyer, MD, MSE,[d] Raymond L. Woosley, MD, PhD,[e] Robert S. Kass, PhD,[c] Nicholas P. Tatonetti, PhD[b]

ABSTRACT

BACKGROUND QT interval-prolonging drug-drug interactions (QT-DDIs) may increase the risk of life-threatening arrhythmia. Despite guidelines for testing from regulatory agencies, these interactions are usually discovered after drugs are marketed and may go undiscovered for years.

OBJECTIVES Using a combination of adverse event reports, electronic health records (EHR), and laboratory experiments, the goal of this study was to develop a data-driven pipeline for discovering QT-DDIs.

METHODS 1.8 million adverse event reports were mined for signals indicating a QT-DDI. Using 1.6 million electrocardiogram results from 380,000 patients in our institutional EHR, these putative interactions were either refuted or corroborated. In the laboratory, we used patch-clamp electrophysiology to measure the human ether-à-go-go-related gene (hERG) channel block (the primary mechanism by which drugs prolong the QT interval) to evaluate our top candidate.

RESULTS Both direct and indirect signals in the adverse event reports provided evidence that the combination of ceftriaxone (a cephalosporin antibiotic) and lansoprazole (a proton-pump inhibitor) will prolong the QT interval. In the EHR, we found that patients taking both ceftriaxone and lansoprazole had significantly longer QTc intervals (up to 12 ms in white men) and were 1.4 times more likely to have a QTc interval above 500 ms. In the laboratory, we found that, in combination and at clinically relevant concentrations, these drugs blocked the hERG channel. As a negative control, we evaluated the combination of lansoprazole and cefuroxime (another cephalosporin), which lacked evidence of an interaction in the adverse event reports. We found no significant effect of this pair in either the EHR or in the electrophysiology experiments. Class effect analyses suggested this interaction was specific to lansoprazole combined with ceftriaxone but not with other cephalosporins.

CONCLUSIONS Coupling data mining and laboratory experiments is an efficient method for identifying QT-DDIs. Combination therapy of ceftriaxone and lansoprazole is associated with increased risk of acquired long QT syndrome. (J Am Coll Cardiol 2016;68:1756–64) © 2016 by the American College of Cardiology Foundation.

Torsades de pointes is a ventricular tachycardia that can result in sudden death (1) and occurs as an adverse effect of more than 40 medications that prolong the QT interval, referred to as acquired long QT syndrome (LQTS) (2). The U.S. Food and Drug Administration (FDA) has established strict guidelines for evaluating the risk of acquired LQTS for new compounds when

Listen to this manuscript's audio summary by *JACC* Editor-in-Chief Dr. Valentin Fuster.

From the [a]Department of Physiology and Cellular Biophysics, Columbia University, New York, New York; [b]Department of Biomedical Informatics, Columbia University, New York, New York; [c]Department of Pharmacology, Columbia University, New York, New York; [d]Department of Cardiology, Columbia University, New York, New York; and [e]Azcert, Inc., Oro Valley, Arizona. Mr. Lorberbaum is supported by National Institute of General Medical Sciences (NIGMS) training grant T32GM082797. Mr. Lorberbaum and Dr. Tatonetti are supported by NIGMS grant R01GM107145. Drs. Sampson and Kass are supported by National Institutes of Health (NIH) grant 5R01GM109762-02. Dr. Iyer is supported by NIH grant K08HL116790. Dr. Tatonetti is a compensated advisor to Advera Health, Inc.; he declares no conflict of interest. Dr. Woolsey is an uncompensated officer of the nonprofit organization Azcert.org, which is supported by U.S. Food and Drug Administration Safe Use Initiative contract HHSF223201400189C and sponsors the CredibleMeds.org website used in this study. All other authors have reported that they have no relationships relevant to the contents of this paper to disclose.

Manuscript received October 6, 2015; revised manuscript received July 11, 2016, accepted July 12, 2016.

administered individually. Nonantiarrhythmic compounds that increase the QT/QTc interval by 20 ms or more are unlikely to be approved, and a compound associated with an increase of 10 ms or more would face many challenges (3). Even a 5 ms increase would prompt an evaluation of the risks and benefits of the new compound (3). Studies of both cardiac and noncardiac compounds found that a QTc interval above 500 ms is associated with significant risk of Torsades de pointes (4,5).

Acquired LQTS is of particular concern when it is not anticipated and occurs as the result of a QT interval-prolonging drug-drug interaction (QT-DDI) (2,6). QT-DDIs are not routinely evaluated preclinically and can go undiscovered for years. For example, quetiapine (an antipsychotic agent) was on the market for nearly 10 years before reports of a QT-DDI with methadone (an analgesic agent) prompted investigation into a possible mechanism (7). It took 3 more years before a label change was made to caution against the use of quetiapine in combination with other drugs known to prolong the QT interval.

SEE PAGE 1765

Large clinical databases, such as electronic health records (EHR), represent an opportunity to rapidly detect QT-DDIs and save lives (8,9). Drug safety algorithms could be applied to health record data in near real time, flagging potentially dangerous drug interactions before they become widespread. Furthermore, these analyses are in situ and therefore focus on the most important drug combinations: those that are actually used in clinical practice. Unfortunately, analysis of medical records is complex, due to issues of missing data, noise, and bias (10). This leads to high false positive rates and algorithms that often will mislead health care providers. Laboratory experiments, especially if they are high-throughput, can be used to screen data-mined hypotheses for plausibility. Following observational analysis with confirmatory prospective experiments can remove the spurious signals, enabling clinically useful discoveries (11).

We developed a data science pipeline to mine potential QT-DDIs from clinical databases. In this pipeline, we combine evidence of QT-DDIs from the FDA Adverse Event Reporting System (FAERS) and the EHR at New York-Presbyterian/Columbia University Medical Center (CUMC-EHR). We identified a putative interaction between lansoprazole (a proton-pump inhibitor [PPI]) and ceftriaxone (a cephalosporin antibiotic). Importantly, this is an interaction that would not have been suspected using current surveillance methods. We used patch-clamp electrophysiology of cells stably expressing human ether-à-go-go-related gene (hERG) channels to establish a physiological mechanism. We further confirmed the specificity of our pipeline by also investigating the combination of cefuroxime (another cephalosporin) and lansoprazole, a drug pair that did not have evidence of an interaction in FAERS. In the clinic, patients on the combination of ceftriaxone and lansoprazole had 12 ms (95% confidence interval [CI]: 7 to 15 ms) longer QTc intervals than patients exposed to either drug alone and were 1.4 times as likely to have a QTc interval above 500 ms. The negative control showed no significant effect. A QT-DDI between ceftriaxone and lansoprazole has the potential for significant morbidity and mortality.

ABBREVIATIONS AND ACRONYMS

APD70 = action potential duration at 70% of repolarization

DDI = drug-drug interaction

ECG = electrocardiogram

EHR = electronic health records

FAERS = Food and Drug Administration adverse event reporting system

hERG = human ether-à-go-go-related gene

LQTS = long QT syndrome

PPI = proton-pump inhibitor

QT-DDI = QT interval-prolonging drug-drug interaction

METHODS

DATA SOURCES. We used 2 independent databases to investigate possible QT-DDIs. The first database (Twosides) was a derivative of 1.8 million adverse event reports from FAERS mined for evidence of adverse drug-drug interactions that could not be explained by the individual effects of the drugs (12). The second database consisted of 1.6 million electrocardiograms (ECGs) from 382,221 patients treated at New York-Presbyterian/CUMC between 1996 and 2014. To obtain the heart rate-corrected QT (QTc) intervals, we wrote a parser to automatically extract the patient identifier, laboratory date, and QTc value from the ECG reports. QTc values were calculated using Bazett's formula. We manually checked 50 abnormal ECGs (defined as QTc >500 ms) to confirm we were extracting the correct values and found that the parser obtained 100% precision and recall. We implemented the pipeline using Python 2.7.9 (Python Foundation, Wilmington, Delaware) and R version 3.2.2. (R Foundation for Statistical Computing, Vienna, Austria).

IDENTIFICATION OF CANDIDATE QT-DDIs. We used the side effect reporting frequencies in Twosides to find drug pairs significantly over-reported with the 6 adverse events in the standardized MedDRA (Medical Dictionary for Regulatory Activities) query for "Torsade de Pointes/QT prolongation"; we call this the direct evidence model (12). However, most drug pairs are not directly reported with QT prolongation. In addition, we performed latent signal detection, a method we have previously validated (13,14), to

identify candidate QT-DDIs that lacked prior direct evidence. To perform latent signal detection, we used machine learning to define and validate a side effect profile of 13 side effects associated with known QT-prolonging compounds. Some of these latently identified side effects (such as arrhythmia and rhabdomyolysis) are positively correlated with QT interval prolongation, whereas others (such as hemorrhage and myocardial infarction) are negatively correlated (**Figure 1B**). We previously validated the method using drug pairs containing a known QT-prolonging drug (2) and demonstrated high specificity and sensitivity (Online Figure 1). We then scanned for novel drug interactions in the TWOSIDES database that matched the side effect profile; we refer to this as indirect evidence. We scored each drug pair for the amount of both direct and indirect evidence.

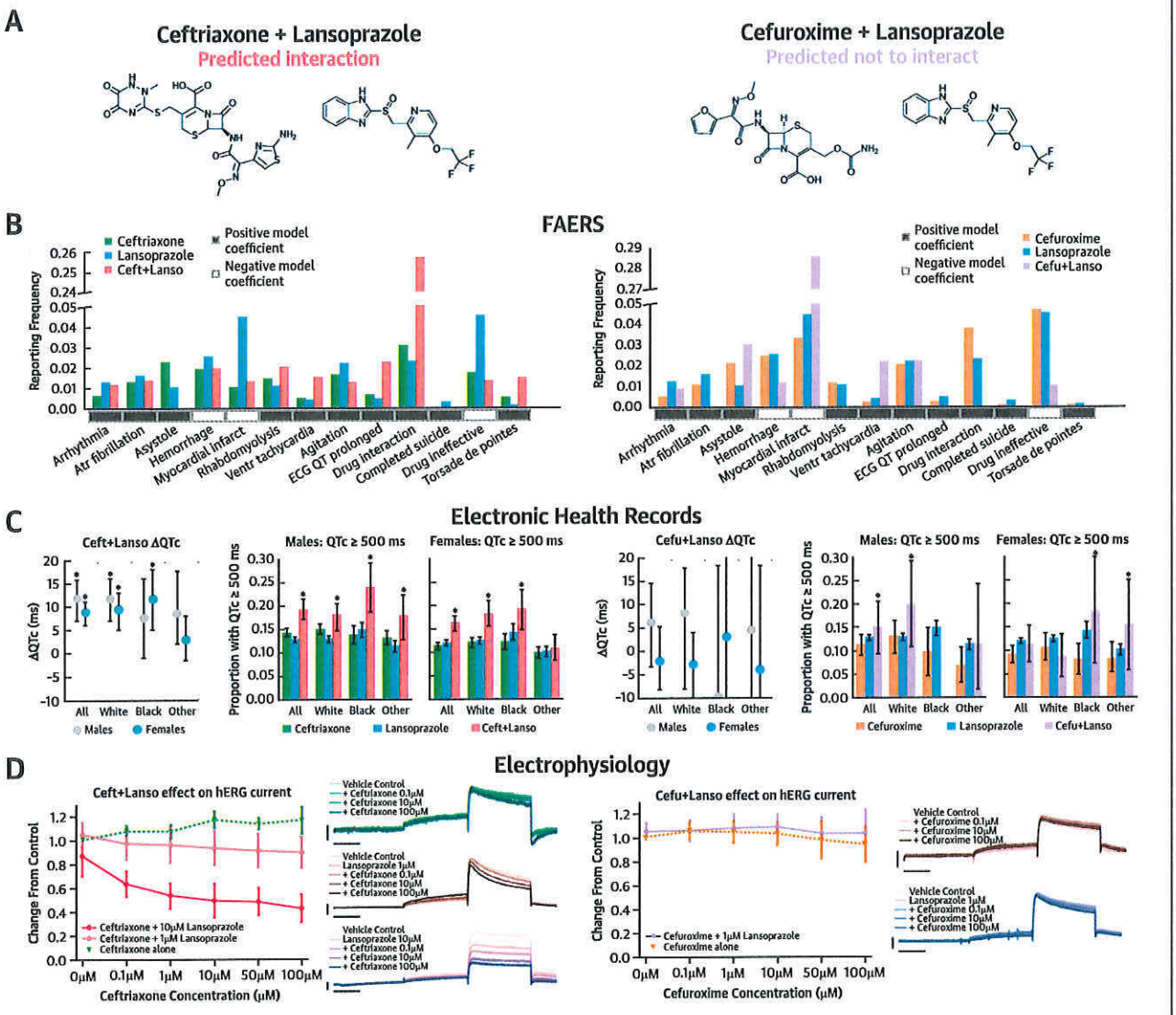
EVALUATION OF CANDIDATE QT-DDIs USING THE EHR. We attempted to corroborate (or refute) each of the candidate QT-DDI hypotheses using the heart rate-corrected QTc values from ECGs stored in the CUMC-EHR. For each candidate drug-drug interaction, we defined an exposed cohort and 2 control cohorts. Those patients included in the exposed cohort were administered both of the drugs within a 7-day window. Those in the control cohorts had evidence of exposure to only 1 of the 2 drugs ever in their records. Only patients who had at least 1 ECG in the following 36 days after drug exposure (either combination or single) were included. Corroboration required that we found significantly longer heart rate-corrected QTc intervals in patients on combination treatment compared with patients on either drug alone. The CUMC-EHR uses Bazett's formula by default; we also evaluated the change in QT interval using the Fridericia, Framingham, and Hodges correction formulae (15). Because the distributions of QTc intervals were non-normal, we assessed significance using a Mann-Whitney *U* test with a Bonferroni correction for multiple hypothesis testing. We further verified that this effect could not be explained by concomitant medications (analysis of covariance with concomitant medications modeled as categorical variables) (14). This analysis was stratified by sex because QT interval durations are known to differ between men and women (16). We evaluated the effects of each drug pair both on individual races and on all races combined (Mann-Whitney *U* test). We also performed a post hoc power analysis to estimate our ability to detect a change in QTc interval for the sample and effect sizes present in our EHR (17). Only those QT-DDIs corroborated by the EHR data (in either men, women, or both) were considered for laboratory analysis.

PATCH-CLAMP ELECTROPHYSIOLOGY. QT-prolonging drugs have in common the ability to block the hERG channel (which conducts I_{Kr}) in the heart. We evaluated the combination of ceftriaxone and lansoprazole by performing patch-clamp electrophysiology of cells stably expressing I_{Kr}. Using an automated patch-clamp system (PatchLiner, Nanion, Germany) in voltage clamp mode, we examined the concentration-dependent block of the I_{Kr} current by each drug individually, as well as in combination, using dimethyl sulfoxide as vehicle control (**Figure 1D**). We applied a voltage protocol with a step to +40 mV, followed by a return to −40 mV, to elicit the inward-rectifying tail current. This protocol was repeated every 20 s for the length of the experiment, and after 10 consecutive sweeps in each concentration, the concentration was increased. We then averaged the current at the end of each drug application and normalized it to the control to measure the block by each compound. We assessed significance by using a test of repeated measures on the log-normalized block percentages.

We performed patch-clamp electrophysiology experiments as described for ceftriaxone alone, lansoprazole alone, and ceftriaxone and lansoprazole combined, and similarly for the negative controls of cefuroxime alone and cefuroxime and lansoprazole combined. We evaluated the ability of ceftriaxone or cefuroxime to block the hERG channel at concentrations of 0.1, 1, 10, 50, and 100 μM. For lansoprazole, we evaluated at 0.1, 1, and 10 μM. We performed 3 combination experiments. For the combination of ceftriaxone and lansoprazole, we held lansoprazole constant at either 1 μM or 10 μM and increased the dose of ceftriaxone stepwise from 0.1 to 100 μM. To evaluate our negative control of cefuroxime and lansoprazole, we held lansoprazole constant at 1 μM and increased the dose of cefuroxime stepwise from 0.1 to 100 μM. The concentrations tested were chosen to include the range of plasma concentrations usually reached during routine clinical use of the drugs (1.9 to 3.9 μM for lansoprazole, 24 to 228 μM for ceftriaxone, and 35 to 428 μM for cefuroxime) (18–21).

COMPUTATIONAL MECHANISTIC MODEL. We used a computational model of the human ventricular myocyte (22) to simulate the action potential for the hERG block we observed for ceftriaxone, lansoprazole, and the combination from our laboratory experiments. We ran the model for a ventricular action potential paced at 1 Hz with baseline conditions and 10% or 55%

FIGURE 1 Data Science and Experimental Pipeline for Identifying and Validating QT-DDIs



(A) Chemical structures for ceftriaxone (cephalosporin) and lansoprazole (proton pump inhibitor), which we predicted would have a QT-DDI. We predicted cefuroxime (cephalosporin) and lansoprazole not to interact. **(B)** QT-DDI discovery in FAERS: data-driven side effect profile containing latent evidence of a QT-DDI (**solid boxes** = positive correlation with QT prolongation; **open boxes** = negative correlation). Each bar represents the reporting frequency of a given side effect in FAERS for ceftriaxone **(green)**, lansoprazole **(blue)**, cefuroxime **(orange)**, ceftriaxone + lansoprazole **(red)**, and cefuroxime + lansoprazole **(purple)**. **(C)** Retrospective corroboration in electronic health records. **(Left)** Differences in QTc interval (mean ± 95% CI) between cases (patients prescribed the drug pair) and controls (patients on only 1 drug). We stratified the analysis by sex (men = **gray**; women = **teal**) and evaluated all races combined, as well as whites, blacks, and "other, including Hispanic" separately. The **asterisk** indicates the change in QTc intervals is statistically significant (Mann-Whitney *U* test with Bonferroni correction). We obtained 95% CIs by bootstrapping case and control QTc distributions and calculating the change in median QTc for each iteration. **(Right)** Percentage of patients with a QTc interval ≥500 ms (mean ± 95% CI), stratified by sex and race. The **asterisk** indicates the combination had a significantly greater proportion of patients with a QTc interval ≥500 ms than either drug alone (independent samples Student *t*-test with Bonferroni correction, comparing means of single drug and combination therapy percentage$_{\geq 500}$ distributions generated using bootstrapping). **(D)** Experimental validation using patch-clamp electrophysiology. **(Left)** Change in hERG current from control (mean ± SD) for increasing concentrations of cephalosporin alone **(dashed line)**, and increasing concentrations of cephalosporin in the presence of a single concentration of lansoprazole **(solid lines)**. **(Right)** Representative traces from each patch-clamp electrophysiology experiment. **(Top to bottom)** hERG channel current in the presence of vehicle only (control), and then cephalosporin at 3 concentrations (0.1, 10, and 100 μM); hERG channel current in the presence of lansoprazole alone and then in combination with progressively increasing concentrations of cephalosporin. CI = confidence interval; other abbreviations as in **Central Illustration**.

JACC VOL. 68, NO. 16, 2016
OCTOBER 18, 2016:1756–64

TABLE 1 Demographic and Clinical Characteristics of Cohort

	Men	Women
Combination of ceftriaxone + lansoprazole		
n	934	1,414
Demographic		
Age, yrs	61.3 ± 16.9	66.5 ± 18.5
% Race distribution		
White	57.3	52.9
African American	18.9	20.2
Other/unknown	23.8	26.9
QTc, ms	458 (398-588)	457 (401-571.7)
% Patients with QTc ≥500 ms	19.27	16.34
Combination of cefuroxime + lansoprazole		
n	107	228
Demographic		
Age, yrs	66.1 ± 15.7	67.6 ± 17.9
% Race		
White	56.1	60.1
African American	13.1	14.9
Other/unknown	30.8	25.0
QTc, ms	450 (393.6-579.4)	443.5 (398.7-579.2)
% Patients with QTc ≥500 ms	14.95	11.40
Ceftriaxone only		
n	5,734	6,850
Demographic		
Age, yrs	59.5 ± 17.9	63.7 ± 19.8
% Race		
White	46.6	45.1
African American	19.0	18.4
Other/unknown	34.4	36.5
QTc, ms	446 (394-566)	448 (398-560)
% Patients with QTc ≥500 ms	14.21	11.43
Cefuroxime only		
n	636	957
Demographic		
Age, yrs	61.5 ± 17.6	66.0 ± 19.3
% Race		
White	54.1	50.3
African American	20.6	19.3
Other/unknown	25.3	30.4
QTc, ms	435 (391.9-552.1)	439 (397-551.1)
% Patients with QTc ≥500 ms	11.16	9.09
Lansoprazole only		
n	12,271	13,074
Demographic		
Age, yrs	60.0 ± 15.8	63.1 ± 17.7
% Race		
White	60.8	54.6
African American	13.9	16.7
Other/unknown	25.3	28.7
QTc, ms	443 (395-572)	445 (399-569)
% Patients with QTc ≥500 ms	12.84	12.07

Values are n, mean ± SD, or median (95% confidence interval).

block of hERG current (chosen using the current block observed in the electrophysiology experiments). We evaluated the action potential duration at 70% of repolarization (APD70).

RESULTS

CANDIDATE QT-DDI DISCOVERY VIA DATA SCIENCE. We detected 889 putative signals in FAERS, of which 34 (1.42× more than expected by chance, $p = 0.003$) were corroborated by the CUMC-EHR, after multiplicity correction. Twenty-six signals were eliminated by confounder analysis for concomitant medications. The remaining 8 combinations could not be explained by concomitant medications and were not previously associated with acquired LQTS (14). We prioritized the combination of ceftriaxone and lansoprazole for experimental validation, as lansoprazole is available over the counter and is one of the top 200 most-prescribed drugs (totaling over 2.6 million prescriptions in 2010) (23). An interaction with a PPI could therefore have a profound impact on patient safety. As a negative control, we chose to evaluate the combination of cefuroxime and lansoprazole as, according to our algorithm, it did not match the side-effect profile for QT prolongation in FAERS (**Figures 1A and 1B**).

CO-MEDICATION OF CEFTRIAXONE AND LANSOPRAZOLE IS ASSOCIATED WITH PROLONGED QT IN THE EHR. Overall, the QTc intervals (Bazett's correction) for male patients taking this combination were 12 ms (95% CI: 7 to 15 ms; $n = 934$) longer than those of patients taking either drug alone ($p < 0.001$); for female patients, QTc intervals for patients taking the combination were 9 ms (95% CI: 5.2 to 11.3 ms; $n = 1,414$) longer than those of patients taking either drug alone ($p < 0.001$) (**Figure 1C**). We evaluated QT interval prolongation post hoc using the Fridericia, Framingham, and Hodges correction formulae. In men, all 3 formulae were significant, with $p < 0.01$ (Online Table 1), and in women, Fridericia and Hodges formulae were significant, with $p < 0.01$. When stratifying by race in addition to sex, we observed the largest effects were in white men (12 ms increase; 95% CI: 6.5 to 17 ms; $p < 0.001$) and in black women (12 ms increase; 95% CI: 3.7 to 18.5 ms; $p < 0.001$). We performed a regression analysis which confirmed the increased sensitivity to the drug pair in white patients ($p = 0.049$) (Online Table 2). In 19% of men taking the combination, the QTc was ≥500 ms, an accepted threshold for clinical concern (3), compared with 14% ($p < 0.001$) of patients taking only 1 drug (**Table 1**).

Applying the same case-control analysis to cefuroxime and lansoprazole showed no significant differences in QTc intervals for either men (7 ms increase; 95% CI: −4.5 to 17 ms, $n = 107$; $p = 0.167$) or women (1.5 ms decrease; 95% CI: −9.3 to 4.3 ms;

n = 228; p = 0.155). We observed no significant changes in QTc interval when further stratifying by race. See Figure 1C for complete results.

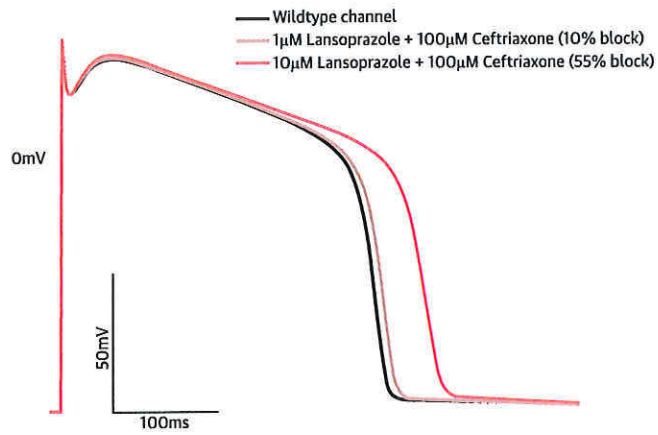
We performed sample-size and effect-size analyses, which demonstrated that, with 100 patients prescribed either combination, we would be able to detect a 10 ms QT interval prolongation with 80% power; with 1,000 patients, the same effect size could be detected with 100% power (Online Figure 2).

A total of 603 patients taking ceftriaxone and lansoprazole had ECGs both before and after they started combination treatment. To control for baseline confounders, we performed a paired analysis comparing each of these patient's highest QTc interval from ECGs performed up to 36 days before and after exposure to ceftriaxone and lansoprazole. We stratified the analysis by both sex and race. We observed a statistically significant increase in QTc interval for both white men (14.0 ± 4.0 ms increase; $p = 6.56 \times 10^{-4}$) and white women (12.9 ± 3.3 ms increase; $p = 1.03 \times 10^{-4}$). We observed no significant change in QTc interval for patients prescribed our negative control. See Online Table 3 for complete results.

IN COMBINATION, CEFTRIAXONE AND LANSOPRAZOLE BLOCK THE hERG CHANNEL. Using a test of repeated measures, we found no significant effect from ceftriaxone on the hERG channel (p = 0.096). We found a significant effect from lansoprazole alone ($p = 1.63 \times 10^{-4}$), causing a drop in current to 86.6 ± 16.7% at 10 μM (no effect at 1 or 0.1 μM). In the presence of 1 μM lansoprazole, ceftriaxone caused a dose-dependent drop in current (96.8 ± 13.2% of control at 0.1 μM; and 89.3 ± 13.2% at 100 μM; $p = 1.07 \times 10^{-4}$). In the presence of 10 μM lansoprazole, ceftriaxone caused a dose-dependent drop in current (63.1 ± 10.9% of control at 0.1 μM; and 42.4 ± 11.6% at 100 μM; $p < 3.45 \times 10^{-5}$) (Figure 1D, left). For our negative control, we saw a small block in cefuroxime alone (94.0 ± 14.8% of control at 100 μM cefuroxime; $p = 5.62 \times 10^{-5}$) but no dose-dependent response of cefuroxime combined with 1 μM lansoprazole (p = 0.083) (Figure 1D, right).

COMPUTATIONAL MODEL RECAPITULATES CLINICAL OBSERVATIONS. Using the hERG current blocks observed in the electrophysiology experiments as input to the computational model, the APD prolongation (measured as APD70) was 9 ms for the combination of 1 μM lansoprazole and 100 μM ceftriaxone and 50 ms for 10 μM lansoprazole and 100 μM ceftriaxone (Figure 2). For the combination of 1 μM lansoprazole and 100 μM cefuroxime, the APD70 was shortened by 2 ms.

FIGURE 2 Results of the Computational Model of Ventricular Epicardial Myocytes



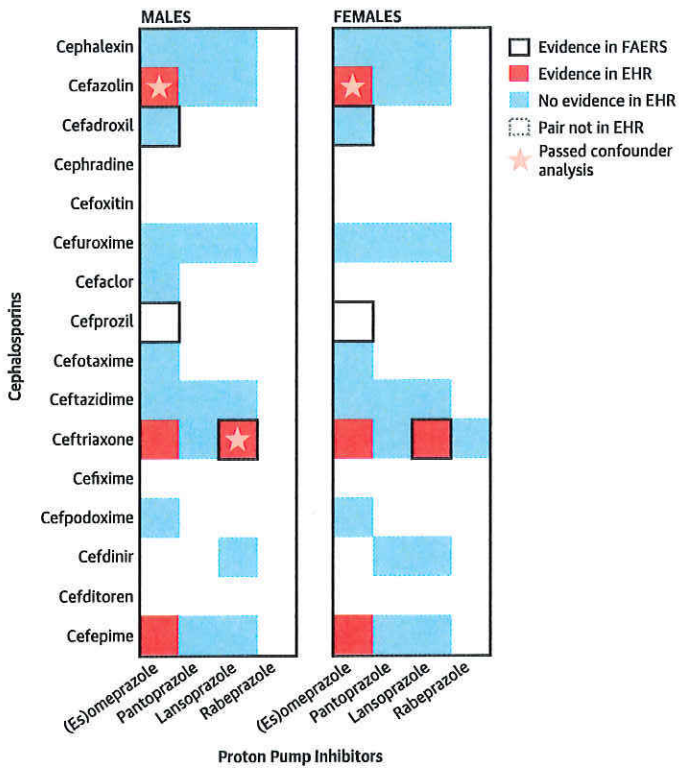
The APD prolongation (measured as APD70) for each case are 9 ms and 50 ms, simulating 1 μM lansoprazole + 100 μM ceftriaxone and 10 μM lansoprazole + 100 μM ceftriaxone, respectively. Briefly, the model was run for a ventricular action potential paced at 1 Hz with baseline conditions **(black)** and 10% or 55% block of peak hERG current **(brown and red** respectively). APD70 = action potential duration at 70% of repolarization; other abbreviations as in **Central Illustration**.

NO EVIDENCE OF CLASS EFFECTS BETWEEN CEPHALOSPORINS AND PPIs. Given our identification of a putative drug interaction between a cephalosporin antibiotic and a PPI, we systematically evaluated all combinations of cephalosporins and PPIs for evidence of a drug interaction in FAERS, EHR, or both (Figure 3). The combination of ceftriaxone and lansoprazole in men was the only drug pair that had evidence in both FAERS and the EHR that also passed our confounder analysis for concomitant medications.

DISCUSSION

NEW DATA SOURCES PRESENT NEW AVENUES FOR DISCOVERY. Data science and large clinical databases present new opportunities to discover adverse drug effects and drug-drug interactions. This is especially true in situations where traditional methods are impractical or unfeasible, as is often the case for DDIs. There are many advantages to taking a retrospective approach for detecting DDIs. The analyses are relatively rapid and inexpensive to perform, and because they are in situ, they focus on drug combinations that are actually used together in clinical practice. In particular, our use of latent signal detection to mine for DDIs using side-effect profile models allowed us to circumvent many of the limitations inherent in

FIGURE 3 Analysis of Class Effects Between Cephalosporins and PPIs



We analyzed each cephalosporin and PPI pair for evidence of an interaction in FAERS (**solid box** = drug pair matches side effect profile), EHR (**red** = patients on combination have significantly prolonged QT intervals compared with those on either drug alone; **blue** = no significant change between cases and controls; **open** = no patients on the drug pair in the EHR), and that the change seen in the EHR was not due to concomitant medications (**red star**). We stratified the analysis between men and women. Only ceftriaxone and lansoprazole in men passed each of these criteria. Abbreviations as in Central Illustration.

conventional data mining approaches that rely solely on direct evidence between drug pairs and side effects (13,14). However, there are many disadvantages as well. Retrospective analysis, and data mining in particular, are notorious for their potential biases and high false discovery rates. There are simply too many potentially confounding variables to make strong statements about causal relationships.

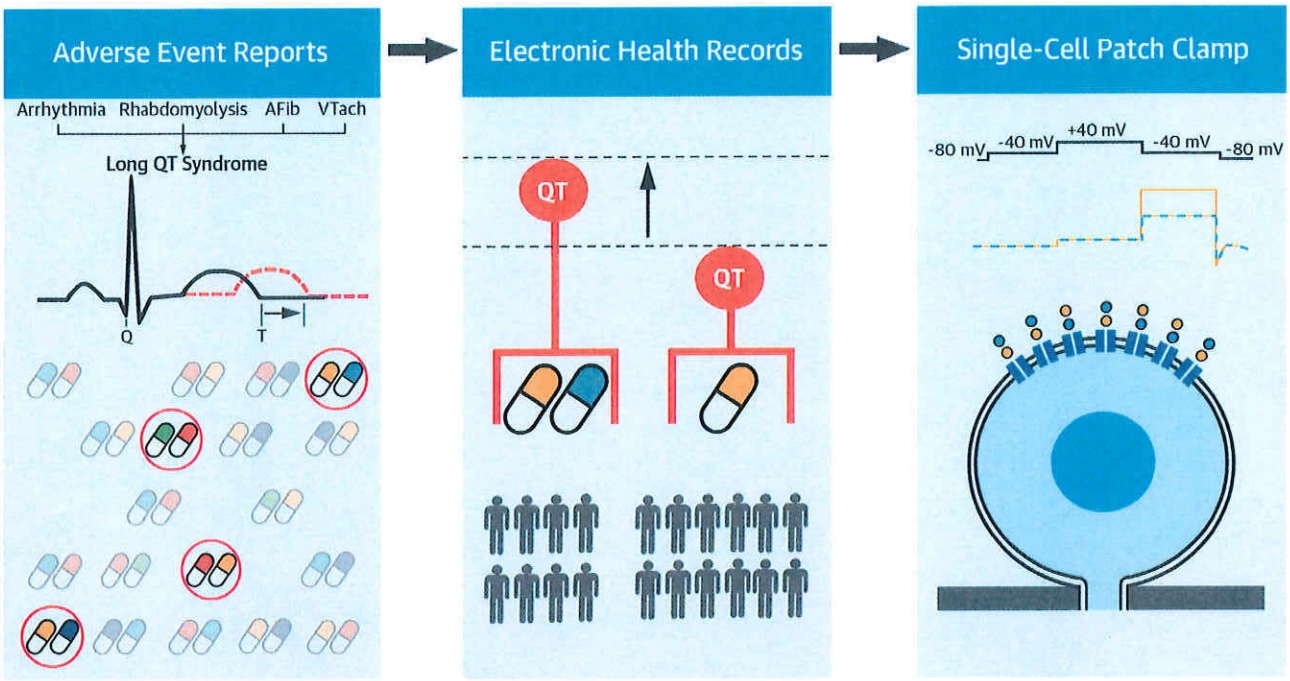
Here, we present a novel strategy that couples observational data mining with laboratory experiments to identify QT-DDIs (Central Illustration). Our observational analysis establishes the presence of a clinically significant association between co-medication and a prolonged QT interval. There are many hypotheses that may explain such an association. For example, a patient prescribed the putative interacting drugs may also be prescribed a known QT-prolonging agent. In fact, this is what we observed. Of 34 drug combinations that were associated with increased QT intervals, 26 could be dismissed as likely confounded by a known agent. Alternatively, it may be that there is a real drug interaction, in the pharmacological sense. The most common physiological explanation would be hERG block; therefore, we tested this hypothesis for our top prediction (ceftriaxone/lansoprazole) by using patch-clamp electrophysiology. This atypical path, going from the clinic into the laboratory, has great potential to increase the efficiency of DDI discovery.

CRITICAL EVALUATION OF DATA MINING USING LABORATORY EXPERIMENTS. We combined data from FAERS with our local EHR to find evidence of QT-prolonging drug interactions. Either data source alone provides only weak evidence of a potential DDI producing thousands of equivalent hypotheses. By integrating these data, we increased power and focused the analysis on only the strongest candidates. Most importantly, we followed up on these DDI hypotheses by using laboratory experiments to identify a possible mechanism.

AN INTERACTION BETWEEN CEFTRIAXONE AND LANSOPRAZOLE IS UNEXPECTED. Our top candidate, ceftriaxone and lansoprazole, would not have been suspected using current surveillance methods. In the clinical records, we found that co-medication of these 2 common drugs is associated with significantly prolonged QTc intervals. This increase was highest for white men and black women, in whom we observed an average increase of 12 ms. It is important to note that, if this effect size was observed for a single drug, it would be well above the threshold for regulatory concern during the approval stage (3). In the laboratory, we found that, in combination, lansoprazole and ceftriaxone block the hERG channel up to 57.6%, corresponding to an APD70 increase of 50 ms. At these higher lansoprazole concentrations, it is likely that, if treated as a single entity, the combination would not have received regulatory approval.

STUDY LIMITATIONS. We discovered that ceftriaxone and lansoprazole were significantly associated with prolonged QT intervals using clinical data. Our laboratory analysis suggests that this effect may be mediated through the hERG potassium channel, the most common mechanism by which drugs prolong the QT interval. However, the molecular explanation is not clear. Possibilities include a chemical interaction between the 2 compounds, cooperative binding to the channel, or an indirect mechanism through proteins that function with hERG. Furthermore, we

CENTRAL ILLUSTRATION **Ceftriaxone and Lansoprazole Are Associated With Acquired LQTS**



Lorberbaum, T. et al. J Am Coll Cardiol. 2016;68(16):1756–64.

We combined mining of adverse event reports, corroboration in electronic health records, and experimental validation using single-cell patch clamp to discover and validate a QT-DDI between ceftriaxone and lansoprazole. We used a data-driven profile of side effects that are predictive of LQTS to prioritize drug pairs in FAERS. We corroborated these findings in the electronic health records by comparing the QTc intervals of patients administered the prioritized drug pair to patients exposed to either drug alone. We then validated our top prediction (ceftriaxone/lansoprazole) by measuring the dose-dependent changes in hERG channel current using patch-clamp electrophysiology. AFib = atrial fibrillation; FAERS = FDA Adverse Event Reporting System; hERG = human Ether-à-go-go-Related Gene; LQTS = long QT syndrome; QT-DDI = QT-prolonging drug-drug interaction; VTach = ventricular tachycardia.

found significantly different effects when our analysis was stratified by race and ethnicity. White men and women appear to be sensitive to the interaction, whereas black men experience only an intermediate change, and women identifying as "other, including Hispanic" experience no detectable effect. This is consistent with the large amount of ethnic heterogeneity in cardiac potassium channels (24,25) and may guide a structural analysis of the interaction.

PRIOR EVIDENCE OF RELATED ADVERSE EVENTS. Lansoprazole is a commonly used PPI that is available over-the-counter. In retrospective analyses, PPIs were associated with a slightly increased risk of myocardial infarction (26). Additionally, there have been a large number of deaths reported to the FDA for patients taking this class of drugs, although this association is not statistically significant. Our discovery of a drug interaction with a PPI may explain these observations, although this requires follow-up study. Notably, evaluation of cefuroxime and lansoprazole, a pair predicted not to interact from the FAERS reporting frequencies, suggests that our pipeline is capable of distinguishing between safe and unsafe pairs, even within the same drug class.

CONCLUSIONS

We present evidence of a novel QT-DDI between lansoprazole and ceftriaxone. This interaction was discovered by using a combination of data mining and laboratory experiments. Our clinical data suggest that patients taking this pair of interacting drugs are more likely to have acquired LQTS, and the experimental study suggests that this effect may be mediated by blocking the hERG channel, the most common mechanism of acquired LQTS. This interaction appears to be

specific to ceftriaxone and does not extend to other cephalosporin antibiotics in combination with lansoprazole. Follow-up studies are required to confirm our findings and should include evaluation of the mechanism of the interaction at the hERG channel, the effect of ceftriaxone and lansoprazole on other ion channels, and investigation of these drugs in combination with other hERG blockers.

ACKNOWLEDGMENT The authors thank Sam Roe for insight and perspective throughout the design and execution of this study.

REPRINT REQUESTS AND CORRESPONDENCE: Dr. Nicholas P. Tatonetti, Department of Biomedical Informatics, Columbia University, 622 West 168th St., PH20, New York, New York 10032. E-mail: nick.tatonetti@columbia.edu.

PERSPECTIVES

COMPETENCY IN SYSTEMS-BASED PRACTICE: Data science methodologies may accelerate evaluation of the safety of drug combinations. Data from large numbers of patients and ECGs in electronic health records can corroborate clinical event reports of adverse drug interactions and, for those involving QT-interval prolongation, laboratory methods can then be used to elucidate the electrophysiological mechanisms involved.

TRANSLATIONAL OUTLOOK: Further work is needed to systematically extend this paradigm to evaluate the safety of commonly used drug combinations in other cardiovascular domains.

REFERENCES

1. Roden DM. Repolarization reserve: a moving target. Circulation 2008;118:981–2.

2. Woosley RL, Romero K. Assessing cardiovascular drug safety for clinical decision-making. Nat Rev Cardiol 2013;10:330–7.

3. Fermini B, Fossa AA, Fermini B, et al. The impact of drug-induced QT interval prolongation on drug discovery and development. Nat Rev Drug Discov 2003;2:439–47.

4. Woosley RL, Chen Y, Freiman JP, et al. Mechanism of the cardiotoxic actions of terfenadine. JAMA 1993;269:1532–6.

5. Shah RR. Drug-induced prolongation of the QT interval: regulatory dilemmas and implications for approval and labelling of a new chemical entity. Fundam Clin Pharmacol 2002;16:147–56.

6. Itzhaki I, Maizels L, Huber I, et al. Modelling the long QT syndrome with induced pluripotent stem cells. Nature 2011;471:225–9.

7. Uehlinger C, Crettol S, Chassot P, et al. Increased (R)-methadone plasma concentrations by quetiapine in cytochrome P450s and ABCB1 genotyped patients. J Clin Psychopharmacol 2007;27:273–8.

8. Hripcsak G, Duke JD, Shah NH, et al. Observational Health Data Sciences and Informatics (OHDSI): opportunities for observational researchers. Stud Health Technol Inform 2015;216:574–8.

9. Psaty BM, Breckenridge AM. Mini-sentinel and regulatory science—big data rendered fit and functional. N Engl J Med 2014;370:2165–7.

10. Hripcsak G, Albers DJ. Next-generation phenotyping of electronic health records. J Am Med Inform Assoc 2012;20:117–21.

11. Tatonetti NP, Denny JC, Murphy SN, et al. Detecting drug interactions from adverse-event reports: interaction between paroxetine and pravastatin increases blood glucose levels. Clin Pharmacol Ther 2011;90:133–42.

12. Tatonetti NP, Ye PP, Daneshjou R, et al. Data-driven prediction of drug effects and interactions. Sci Transl Med 2012;4:125ra31.

13. 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;19:79–85.

14. Lorberbaum T, Sampson KJ, Woosley RL, et al. An integrative data science pipeline to identify novel drug interactions that prolong the QT interval. Drug Saf 2016;39:433–41.

15. Indik J, Pearson EC, Fried K, et al. Bazett and Fridericia QT correction formulas interfere with measurement of drug-induced changes in QT interval. Heart Rhythm 2006;3:1003–7.

16. Rautaharju PM, Zhou SH, Wong S, et al. Sex differences in the evolution of the electrocardiographic QT interval with age. Can J Cardiol 1992;8:690–5.

17. Collings BJ, Hamilton MA. Estimating the power of the two-sample Wilcoxon test for location shift. Biometrics 1988;44:847–60.

18. Sakurai Y, Hirayama M, Hashimoto M, et al. Population pharmacokinetics and proton pump inhibitory effects of intravenous lansoprazole in healthy Japanese males. Biol Pharm Bull 2007;30:2238–43.

19. Tolman KG, Sanders SW, Buchi KN, et al. The effects of oral doses of lansoprazole and omeprazole on gastric pH. J Clin Gastroenterol 1997;24:65–70.

20. Rocephin [package insert]. South San Francisco, CA: Genentech USA, Inc, 2015.

21. Foord RD. Cefuroxime: human pharmacokinetics. Antimicrob Agents Chemother 1976;9:741–7.

22. Iyer V, Mazhari R, Winslow RL. A computational model of the human left-ventricular epicardial myocyte. Biophys J 2004;87:1507–25.

23. Pharmaceutical Sales 2010. Verispan, VONA. Available at: http://www.drugs.com/top200_units.html. Accessed August 9, 2016.

24. Modell SM, Lehmann MH. The long QT syndrome family of cardiac ion channelopathies: a HuGE review. Genet Med 2006;8:143–55.

25. Ackerman MJ, Tester DJ, Jones GS, et al. Ethnic differences in cardiac potassium channel variants: implications for genetic susceptibility to sudden cardiac death and genetic testing for congenital long QT syndrome. Mayo Clin Proc 2003;78:1479–87.

26. Shah NH, LePendu P, Bauer-Mehren A, et al. Proton pump inhibitor usage and the risk of myocardial infarction in the general population. PLoS One 2015;10:e0124653.

KEY WORDS data mining, data science, drug-drug interaction, long QT syndrome

APPENDIX For supplemental tables and figures, please see the online version of this article.

The Tribune-Columbia collaboration also resulted in a scientific paper published in Drug Safety.

Drug Saf (2016) 39:433–441
DOI 10.1007/s40264-016-0393-1

ORIGINAL RESEARCH ARTICLE

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

Tal Lorberbaum[1,2,3] · Kevin J. Sampson[4] · Raymond L. Woosley[5] · Robert S. Kass[4] · Nicholas P. Tatonetti[2,3,6]

Published online: 10 February 2016

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

1 Department of Physiology and Cellular Biophysics, Columbia University, New York, NY, USA

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

3 Departments of Systems Biology and Medicine, Columbia University, New York, NY, USA

4 Department of Pharmacology, Columbia University, New York, NY, USA

5 AZCERT, Inc., Oro Valley, AZ, USA

6 Observational Health Data Science and Informatics, New York, NY, USA

Key Points

Drug–drug interactions that prolong the QT interval (QT-DDIs) can 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

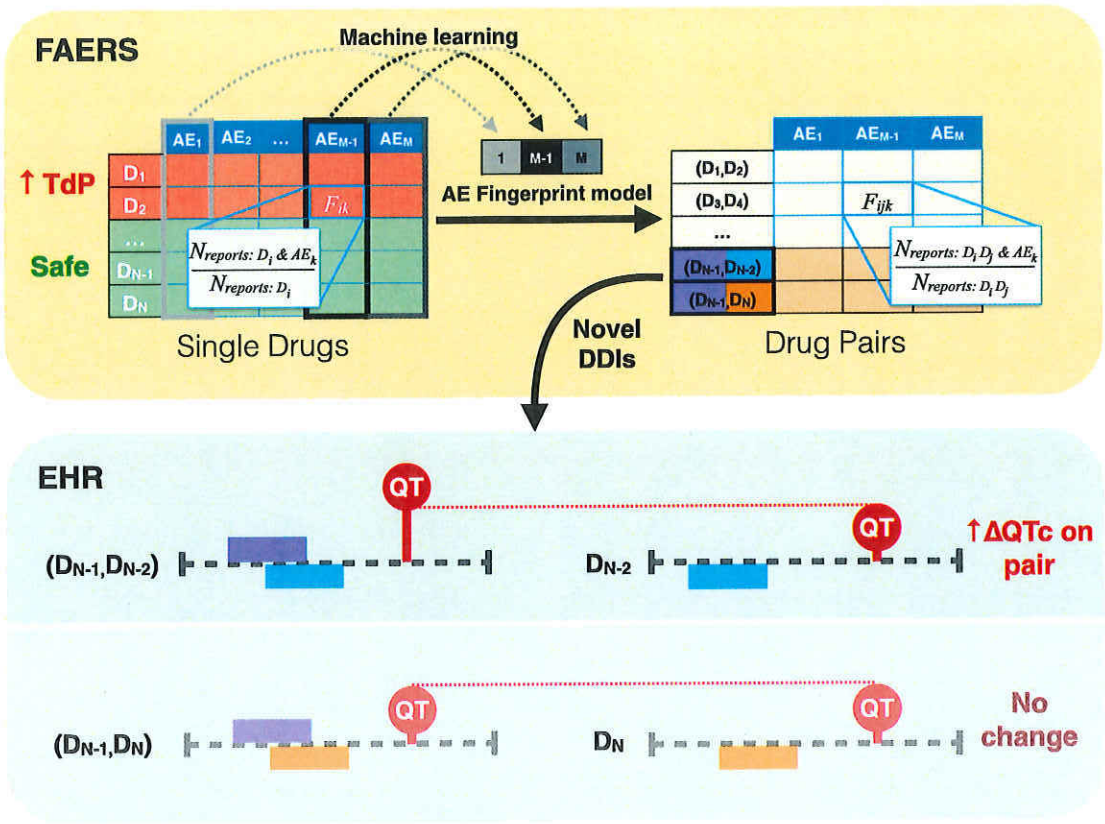


Fig. 1 Overview of DIPULSE pipeline, which combines mining of FAERS and EHRs to flag novel QT-prolonging DDIs. FAERS: We generate an AE reporting frequency table (dimensions, *N* drugs by *M* AEs) for single drugs in FAERS. The value at a *row* and *column* represents the fraction of reports for drug *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 β 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 ≥ 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 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.

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

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

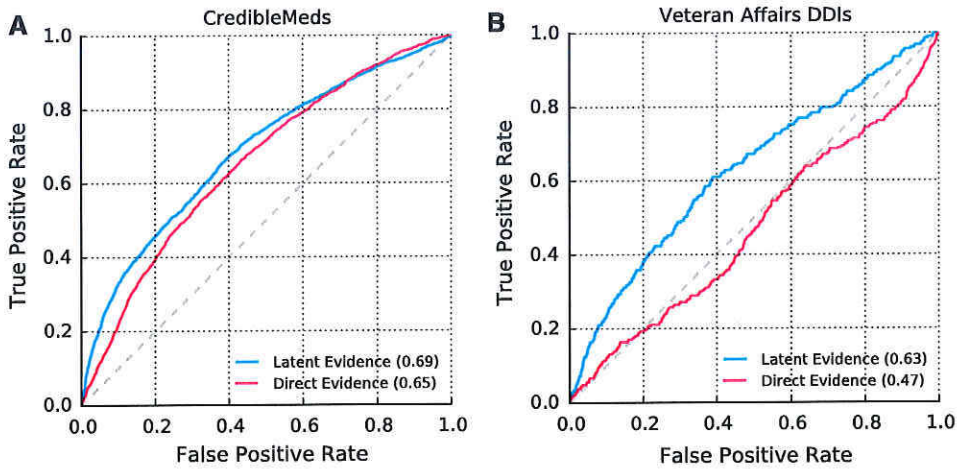


Fig. 2 Receiver operating characteristic curves for adverse event fingerprint model and direct evidence control. **a** Model validation was performed by labeling drug pairs containing a drug with known increased risk of TdP as positive examples. We compared the performance of a model built using latent 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	*p* value	Median QTc cases	Median QTc controls	ΔQTc (ms)	No. of cases	No. of controls
Octreotide	Lactulose	Octreotide	M	74.8	2.48E−04	485	455	**30**	333	603
Mupirocin	Vancomycin	Vancomycin	F	54.5	1.30E−04	476	456	**20**	810	10,165
Metoprolol	Fosphenytoin	Metoprolol	M	40.9	2.19E−07	462	444	**18**	549	24,717
N-Acetylcysteine	Vancomycin	Vancomycin	M	17.4	3.74E−04	469	453	**16**	2633	9789
Cefazolin	Meperidine	Cefazolin	F	27.6	1.29E−05	455	441	**14**	1025	9172
Cefazolin	Meperidine	Cefazolin	M	18.2	8.97E−08	452	440	**12**	2110	10,013
Ceftriaxone	Lansoprazole	Ceftriaxone	M	39.1	4.21E−09	458	446	**12**	934	5734
N-Acetylcysteine	Morphine	*N*-Acetylcysteine	M	12.1	3.19E−02	460	451	**9**	2525	6046
Meperidine	Vancomycin	Vancomycin	F	34.6	4.77E−03	464	457	**7**	1105	9894
N-Acetylcysteine	Morphine	*N*-Acetylcysteine	F	22.3	7.93E−04	459	455	**4**	1900	4803

The bolded column highlights the ΔQTc 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.

References

1. Roden DM. Clinical practice. Long-QT syndrome. N Engl J Med. 2008;358(2):169–76.
2. 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.
3. Kannankeril P, Roden DM, Darbar D. Drug-induced long QT syndrome. Pharmacol Rev. 2010;62(4):760–81.
4. 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.
5. Moss AJ, Kass RS. Long QT syndrome: from channels to cardiac arrhythmias. J Clin Invest. 2005;115(8):2018–24.
6. Berger SI, Ma'ayan A, Iyengar R. Systems pharmacology of arrhythmias. Sci Signal. 2010;3(118):ra30.
7. 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.
8. Woosley RL, Romero K. Assessing cardiovascular drug safety for clinical decision-making. Nat Rev Cardiol. 2013;10(6):330–7.
9. Woosley RL, Chen Y, Freiman JP, Gillis RA. Mechanism of the cardiotoxic actions of terfenadine. JAMA. 1993;269(12):1532–6.
10. Uehlinger C, Crettol SV, Chassot P, Brocard M, Koeb L, Brawand-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.
11. 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.
12. Hajjar ER, Cafiero AC, Hanlon JT. Polypharmacy in elderly patients. Am J Geriatr Pharmacother. 2007;5(4):345–51.
13. 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.
14. Bate A, Evans SJW. Quantitative signal detection using spontaneous ADR reporting. Pharmacoepidemiol Drug Saf. 2009;18(6):427–36.
15. 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.
16. 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.
17. 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.
18. 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.
19. Norén GN, Sundberg R, Bate A, Edwards IR. A statistical methodology for drug–drug interaction surveillance. Stat Med. 2008;27(16):3057–70.
20. Harpaz R, Chase HS, Friedman C. Mining multi-item drug adverse effect associations in spontaneous reporting systems. BMC Bioinform. 2010;11(Suppl 9):S7.
21. Hripcsak G, Albers DJ. Next-generation phenotyping of electronic health records. J Am Med Inform Assoc. 2013;20(1): 117–21.
22. 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.
23. 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.
24. 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.
25. Tatonetti NP, Ye PP, Daneshjou R, Altman RB. Data-driven prediction of drug effects and interactions. Sci Transl Med. 2012;4(125):125ra31.
26. Tibshirani R. Regression shrinkage and selection via the Lasso. J R Stat Soc Ser B Stat Methodol. 1996;58(1):267–88.
27. 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.
28. 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.
29. Vanholder R, Sever MS, Erek E, Lameire N. Rhabdomyolysis. J Am Soc Nephrol. 2000;11(8):1553–61.
30. Woosley RL, Romero KA. QTdrugs List. Available at: http://www.Crediblemeds.org. Accessed 22 Dec 2015.