Fragment-Based Prediction of the Clinical Occurrence of Long QT Syndrome and Torsade de Pointes

Matthew Clark* and Jeffrey S. Wiseman

Pharmatrope Ltd., 324 Croton Road, Wayne, Pennsylvania 19087

Received March 30, 2009

We present a study directly linking clinically observed adverse events to molecular structure. The method is applied to predict the long QT syndrome (LQS) and the resulting condition torsade de pointes (TdP) that can lead to sudden death. The predictive models are created by correlating biochemically significant chemical substructures, derived from a database of marketed drugs, to reports of adverse events in the FDA adverse event reporting system, which contains all events reported to the FDA since 1997. We compute the reporting ratio for each drug/event combination and perform a χ^2 test to determine whether there is a statistically significant association of each drug to reports of LQS and TdP. Linear models are then used to identify chemical substructures that are most consistently associated with the adverse event. The results for LQS and TdP are compared to models for LQS based on human ether-a-go-go-related gene binding and tested for statistical significance by comparing to models created with a randomized dependent variable. The ability to identify compounds associated with LQS and TdP is approximately five times improved in comparison to models based on randomized data, suggesting that there is a significant relationship between specific chemical structures and these adverse events.

INTRODUCTION

Clinical adverse events are among the most significant barriers to successful drug development. The FDA 2005 Report to the Nation lists 32 high-profile drugs recently withdrawn from the market due to postmarketing safety concerns.[1] Seven of these involved fatal cardiac arrhythmias. In each case, the drugs passed the required regimen of animal and clinical trials before marketing. Methods to anticipate and avoid these costly safety issues could be enormously beneficial to both drug development and public safety. Linking the large database of adverse event reports from the FDA to the chemical structure may provide a means to anticipate adverse events and to understand the influence of chemical substructures on important adverse events. While statistical models cannot replace preclinical and clinical studies, the models may suggest additional safety studies that could prevent severe adverse events in the general public.

One of the most significant drug safety concerns relates to the role of QT interval prolongation in cardiac arrhythmias. The electrocardiogram QT interval represents the duration of ventricular depolarization and subsequent repolarization and is measured from the beginning of the QRS complex to the end of the T wave. A delay in cardiac repolarization creates an electrophysiological environment that favors the development of cardiac arrhythmias, including torsade de pointes (TdP).[2] TdP is a polymorphic ventricular tachyarrhythmia that appears on the ECG as continuous twisting of the vector of the QRS complex around the isoelectric baseline. TdP can degenerate into ventricular fibrillation, leading to sudden death.[3,4] The discovery of the link between sudden death due to TdP and drug-induced QT prolongation has had significant impact on the drug development process.[5]

While long QT syndrome, LQS, is an imperfect biomarker for proarrhythmic risk, in general, there is a qualitative relationship between the QT prolongation and the risk of TdP, especially for drugs that cause substantial prolongation of the QT interval. Several drugs, including terfenadine, astemizole, grepafloxicin, terodiline, droperidol, lidoflazine, sertindole, levomethadyl, and cisapride, have been withdrawn from the market, or their use severely restricted due to their association with drug-induced LQS.

In turn, the biological mechanism of QT prolongation has been linked to the human ether-a-go-go-related gene (hERG), which encodes the $K_v11.1$ potassium ion channel protein responsible for cardiac repolarization.[6,7] This protein is commonly called hERG, after the encoding gene. The discovery of a biomolecular mechanism for QT elongation has led to a body of work relating drug structure to hERG binding in attempts to understand the relationship between the molecular structure, the QT elongation, and the final outcome, TdP. This previous work centered on modeling and QSAR methods looking for relationships between hERG-binding IC_{50} and structure. Table 1 lists recent publications of in silico models for hERG-binding affinity.

While these models are valuable, hERG binding is only a surrogate for QT elongation, which in turn is a precursor for torsade de pointes. The ability to directly anticipate clinical adverse events, such as torsade de pointes, based on drug structure could make drug development dramatically more efficient, since toxicity and other adverse events are responsible for the majority of drug failures during clinical trials.

Despite the significant body of work relating chemical structure to properties, such as hERG binding (Table 1), and to a variety of cell-based and animal toxicities,[4] there has been little work relating chemical structure directly to

* mclark@pharmatrope.com.

10.1021/ci900116q CCC: $40.75
Published on Web 11/06/2009

Table 1. Recent Publications of In Silico Modeling of hERG Blockers

references	methods employed
	homology modeling
Mitcheson et al., 2000[8]	alanine-scanning mutagenesis, homology modeling using a closed K^+ channel structure (Kcsa) and docking analysis
Pearlstein et al. 2003[9]	homology model that based on open K^+ channel structure (MthK) and CoMSIA model
	pharmacophore modeling
Cavalli et al., 2002[10]	conformational search and clustering using MacroModel and 3D-QSAR (CoMFA) models
Ekins et al., 2002[11]	pharmacophore model using Catalyst
Zolotoy, et al., 2003[12]	conformational analysis and ab initio calculation to understand the physicochemical determinants for hERG blockers
	2D QSAR (classification and regression)
Qikprop[13]	QikProp descriptors and multiple linear regression
Roche et al., 2002[14]	a variety of QSAR descriptors and statistical classification methods
Keseru et al., 2003[15]	hologram QSAR descriptors and partial least-squares
Aronov et al., 2004[16]	classification model based on 2D topological similarity filter and 3D pharmacophore ensemble
Bains et al., 2004[17]	fragment-based and experimental descriptors and evolutionary algorithm
Song, Clark,[18] 2006	fragment based method using a variety of statistical analyses.
Sierstad, Agrafiotis 2006[19]	structure keys and physicochemical descriptors

clinically observed effects. This has been due, in large part, to the lack of sufficient data in a form convenient for analysis.[20]

The main public source of data regarding clinical outcomes is now the FDA adverse event reporting system (AERS).[21] This resource was created by the FDA to monitor the ongoing safety of marketed drugs. AERS provides an efficient approach for the receipt, processing, and analysis of safety reports through advanced automated functions that cover the full life cycle of post-marketing surveillance activities, including information capture, storage, analysis, and dissemination. These functions, combined with administrative and business capabilities, make AERS an effective tool for national and international pharmacovigilance. The AERS system was brought on line in 1997 to replace the older spontaneous reporting system (SRS) database that was compiled from 1969 to 1997.

The AERS consists of a large set of individual case reports of adverse events. A case report includes all drugs administered to the patient, the drug(s) suspected of causing the event, drug manufacturer, dosage, frequency, administration dates, and administration route. Patient information includes the age, sex, and weight. The adverse events are reported using the medical dictionary for regulatory activities (MedDRA)[22] classification, and the outcome of the patient is reported. Starting from the first report in 1996 to the second quarter of 2008, the database includes 2 061 597 case reports covering 1 871 unique drugs or drug combinations (after consolidation of various trade names for the same drug substance) and 7 933 712 individual adverse events.

Analyzing the AERS to draw conclusions on the relationship of drugs and adverse events still presents some challenges, however, since the data is not uniformly collected and the reporting is voluntary. Thus, some physicians are more inclined than others to report events; the threshold of event that triggers a report may differ among physicians, and the inclination to report may vary based on perception and opinion of the drugs used and on public awareness.

An adverse event report does not always imply causation between the drug and the effect, since reports are based only on the physician's impression. In most cases, the patient has been administered multiple drugs, and the physician determines which one is most responsible for the adverse event by coding it as "primary suspect" or "secondary suspect". Drawing conclusions across all drugs may also be confused based on the nature of the patients; drugs administered to the elderly or terminally ill may have higher rates of events related to these demographics. In particular, in many cases, the detection of specific drug adverse events is biased due to the condition being treated by the drug. For example, the adverse event "t-cell depletion" may only be detected if the patient has a disorder warranting the appropriate tests to detect this condition. The total number of reports for a drug will depend on time in the market for the drug and on the number of prescriptions written. And last, physicians and manufacturers may have submitted some reports in duplicate.

Since data reporting is uneven, many previous efforts in this field have focused on normalizing the data within the database to allow detection of statistically significant signals. Self-normalizing procedures involve computing the average number of times each adverse event is reported per drug, and signal detection for significance beyond this average has been studied by using a Bayesian data-mining approach,[23] disproportionality mining,[24,25] by several groups.[26−29] Other methods to detect significant signals include: multi-item gamma Poisson shrinker, a Bayesian confidence propagation neural network,[30] proportional reporting ratios (PRR and aPRR),[31] odds ratios, and incidence rate ratios.[32] In general, this previous work has been performed on the older SRS database, which is less curated and, therefore, requires more sophistication to draw conclusions.

Given methods to detect significant drug−event relationships, Matthews, et al., were able to use SRS data to successfully predict toxicity in general from chemical structure.[33] The Tropsha group have used the k-nearest neightbore (kNN) method to study human liver adverse events.[34] The current work is an attempt to increase the granularity of this kind of analysis by relating chemical structure to specific clinical adverse events. In order to take the first steps, we study LQS and TdP and their relationship to a specific

mechanism, hERG binding. Linear QSAR models were first constructed to relate each activity individually to the chemical substructures of drugs showing a significant association with the activity. The relationships between the three activities were then assessed by identifying chemical fragments that contributed significantly to more than one model. The robustness of the models was verified by scrambling the dependent variable and analyzing the statistics of the randomized models. The models were also assessed by predicting the LQS and TdP propensities for a test set of drugs that were withdrawn from the market due to QT-related events.

METHODS

Data Sets. The AERS was loaded into a PostgreSQL database system for analysis.[35] The database consists of several tables linked by the "ISR", a unique number that identifies each case report. Each case report contains the patient demographic data, the drugs that were administered, the adverse events, and the case outcome. However, many case reports include only a subset of this information. The most important tables for this study are the drug table, which lists the drugs administered to the patient when the adverse event took place, and the event table, which lists the adverse events that occurred. Several drugs and adverse events may be reported for each case report.

The AERS database uses drug names as reported by physicians, which means that a variety of trade names and formulation names are reported for each unique drug entity. The drug names were rationalized by assigning the FDA active ingredient name to each reported drug using the drugs@FDA data file to map the reported trade names to the FDA active ingredient names.[36] This database comprises 2 190 unique active ingredients. Of these, 1 214 appear as primary suspect agents in AERS case reports. In addition, for each case, the drugs administered are identified by role codes that specify the perceived relation of the drug to the adverse event. These are primary suspect, secondary suspect, concomitant, and interacting. In this study, only drugs identified as primary or secondary suspects for an event were used to correlate to the event. FDA researchers have created similar databases that combine data from the SRS database and the AERS and from several commercial databases.[37]

Drug structures for each of the 1 222 active ingredients were retrieved from PubChem using a name search and downloaded as structure files. Counter ions were removed to isolate the chemical entity defining the active drug structure.

The AERS database is not a uniform sampling of adverse events. The number of events reported for a drug may depend on several factors, including the number of times the drug is prescribed, the duration on the market, and the inclination of the reporting physician. Therefore, several methods have been used to make the database "self-normalizing." This normalization uses the large base of reports to compute a reporting baseline frequency of each adverse event over all drugs and then analyzes occurrences of drug/adverse event combinations that occur more frequently than expected from the baseline.

We use a relative reporting ratio for normalization, computed using a 2 × 2 table as shown in Figure 1. The baseline frequency for a drug/event combination is computed using eq 1:

	Reports with drug j	Reports w/o drug j	Total
Reports with event i	a	b	a+b
Reports w/o event i	c	d	c+d
Total	a+c	b+d	a+b+c+d

Figure 1. Statistics table used to compute relative reporting ratio.

	Reports with drug i	Reports w/o drug i	Total
Reports with event j	17	102	119
Reports w/o event j	15,104	7,131,022	7,146,126
Total	15,121	7,131,124	7,146,245

Figure 2. Statistics table for drug aspirin and event Reye's syndrome.

$$\text{baseline count } e_{ij} = (a + b)(a + c)/(a + b + c + d) \tag{1}$$

$$\text{relative ratio } \mathrm{rr}_{ij} = n_{ij}/e_{ij}, \quad \text{where } n_{ij} = a \tag{2}$$

The relative reporting ratio, the ratio of the number of reports of a specific event to the baseline is given by eq 2, where n_{ij} is the observed count of event type i reported for drug j. The relative ratio rr_{ij} provides a metric of how much more often an event is reported than would be expected given the baseline of all other drugs.

This can be further illuminated with an example. Reye's syndrome is now a well-known adverse event of aspirin. The number of reported cases has dropped precipitously since the connection has been publicized.[38] Figure 2 shows the cells of Figure 1 filled in with the reports for aspirin.

There were 17 reported cases of Reye's syndrome associated with aspirin, and 102 cases reported without aspirin. There are 15 104 events reported for aspirin other than Reye's syndrome, and 7 131 022 events reported without either aspirin or Reye's syndrome. The expected baseline number of events is, therefore, 0.25 = 119(15 121/7 146 245), as compared to the 17 observed events, leading to a relative ratio of 67.5.

The statistical significance of the relative ratio is computed using the χ^2 test, as computed in eq 3. In this equation, the baseline expected events e_{ij} are computed using eq 1, n_{ij} is the count of observed events for event i, drug j, and all events is the total number of events reported:

$$X_{ij}^2 = \left(\frac{n_{ij} - e_{ij}}{\sqrt{(e_{ij}(1 - e_{ij}/\text{all events})}} \right)^2 \tag{3}$$

For this and the analyses below, we require a confidence level of 95% for the drug/event relations with a corresponding χ^2 value of 3.84. The threshold requires that the observed chi-squared value must exceed the value for the confidence level and was used to eliminate spurious correlations that arise due to small sample sizes. In addition, in keeping with

accepted practice, drug/event combinations with less than five reported events for either a or b in Figure 1 were removed, regardless of the χ^2 value, to further reduce outliers due to small sample sizes. For most drugs only a small subset of adverse events survives this statistical test. The relative ratio of 67.5 for the aspirin−Reye's syndrome association corresponds to a χ^2 value of 33.5. Thus, the association is statistically significant at the 95% level.

Fragment Fingerprints. A fragment-based characterization of chemical structures was carried out as in previous works.[18,39] The structural fragments were selected based on two criteria. The first was to include all of the basic chemical functional groups, those containing carbon, oxygen, nitrogen, sulfur, and phosphorus. Carbon-containing groups include ethane, ethylene, and aromatic rings. Oxygen-containing groups include aldehydes, ketones, ethers, carboxylic acids, and amides. Similar functional groups were chosen for nitrogen, phosphorus, and sulfur. A second set of fragments was based on small heterocyclic ring systems found in drug molecules. These two sets provided 321 total fragments. While other QSAR researchers have used much larger descriptor sets in hopes of finding correlations,[6,40] the goal of this study is to focus on chemically meaningful substructures that can be interpreted by chemists to gain understanding and guide modification to reduce adverse events. A complete list of fragments is given in the Supporting Information.

The methods of encoding chemical structures as bitstrings for characterization and comparison are well-known for the computing of molecular similarity and the creation of QSAR models.[41,42] Bitstring fingerprints, which encode the presence of a fragment by setting a bit to "1" if present or "0" if absent, were created by searching each drug structure for the presence of each fragment described above. These searches were carried out with custom in-house substructure searching software, although any system capable of substructure searches could be used. The resulting bitstrings were used as indicator variables for correlating the relative ratio of adverse events to the presence of the fragments in the drug molecules, using the values of 1 and 0. This general method has shown to be effective for creating models for solubility, log*P*, and melting point[9] as well as hERG affinity.[4] In these previous models, the count of times the substructure occurred was used as the indicator variable; in this case, we use only the bit map.

Correlations. Correlations were performed with the R statistical package.[43] Since there are more drug structures than fragments in this study, straightforward multiple linear regression models could be used, but since there are significant correlations among the descriptors, the partial least-squares method (PLS) was preferable and was used instead. Correlation among the descriptors can occur when one fragment is the substructure of another, such as benzene and fluorobenzene, or when two nonrelated fragments consistently co-occur. Correlation, therefore, depends on both the fragment and training sets of drug molecules. PLS is effective both at determining meaningful coefficients when the descriptors are correlated and at separating signal from random noise to make meaningful correlations.[44] For each model created, PLS was carried out by first using cross-validation to determine the optimal number of components and then by using that number of components to create the final model.

Direct use of the reporting ratio for an adverse event as the dependent variable led to spurious correlations when one or two drugs had a reporting ratio an order of magnitude larger than the other drugs. In order to ensure even scaling and compensate for the noisiness of the data, the dependent variable was reduced to an indicator variable. The value was 1 if the χ^2 test for the reporting ratio met the significance criteria and 0 if it did not. Experiments with other variations, such as using the logarithm of the reporting ratio were attempted; however, the models were not effective. This binary approach is similar to one reported for creating models for cardiovascular events from FDA data.[45]

The resulting linear model computes the likelihood of significant numbers of adverse events for a drug or a drug candidate. Each individual coefficient of the resulting model is associated with a molecular fragment and represents the contribution of that fragment to the likelihood of adverse events for the drug. The goal of the model is to classify the predicted compound as having an association with the adverse event or not. The predicted value, computed from the regression coefficients, varies roughly between 0 and 1 and is compared to a threshold to decide whether the drug is predicted to be associated with the event. Since it is the sum of the coefficients associated with each fragment, the indicator variable may be less than 0 or greater than 1, depending on the fragment makeup of the molecule and the coefficients created by the training-set and PLS model. We call the predicted value the "indicator variable" since it indicates an association between the drug and the event. For example, if the threshold is 0.5 and the model computes an indicator variable of 0.5 or greater for a drug, the drug is classified as linked to the event. The impact of the choice of threshold is studied in the results section.

RESULTS

Long QT Syndrome (LQS) Modeling. A total of 77 drugs are associated with QT events at the 95% confidence level, as measured by the chi-squared test described above. The analysis of the predictive model generated for these drugs starts with a truth-table that measures: true positives, which are compounds both predicted and observed to be related to LQS; false positives, which are compounds predicted to be related to LQS, but not observed to be; true negatives, which are neither predicted nor observed to be related to LQS; and false negatives, which are observed to be related to LQS in clinical data but are not predicted to be related. In the best case, the model would classify the compounds into both true positives and negatives, with no false predictions. Figure 3 shows the truth table for QT elongation with an indicator variable threshold of 0.4, i.e., if the fragment-based model computes a value of 0.4 or greater for the indicator variable for a compound, that compound is predicted to be associated with LQS.

The utility of this model is best evaluated in the context of a specific purpose. One purpose, for example, might be to identify compounds early in a drug discovery program that have a very high probability of causing an adverse event, so that these compounds might be barred from further development or earmarked for more extensive in vivo

		COUNTS Predicted TRUE	FALSE	total
Observed	TRUE	35	42	77
	FALSE	12	1220	1232
	total	47	1262	
Recall =		35 / 77 =	45%	
Precision =		35 / 47 =	74%	
Specificity =		1220 / 1232 =	99%	

Figure 3. Truth table and quality metrics for prediction of long QT syndrome (LQS) at an indicator threshold of 0.4. The LQS model, based on AERS data, was used to compute an indicator variable for each compound in the AERS collection. The indicator value varies approximately between 0 and 1, and higher values of the indicator threshold represent more stringent selection criteria. If the value of the indicator variable meets or exceeds a threshold, 0.4 in this case, that compound is scored as related to the event (positive). Recall, precision, and specificity factors are defined in eqs 4−6 and are metrics for the quality of the model.

evaluation. In this case, we want to identify as many true positives as possible, but at the same time, we want to minimize the number of false positives so that otherwise desirable compounds are not dropped inappropriately. The number of true positives selected compared to the total number of true positives in the compound collection is commonly termed recall, and the number of true positives compared to the total number of compounds selected is termed precision. These metrics are defined in eqs 4−5.

If the goal is to create a compound collection that is depleted of compounds causing adverse events, recall is still used to give the number of active compounds selected, in this case for removal from the collection, compared to the total number of actives. The metric for comparison, however, is specificity, which is the number of true negatives retained compared to the total number of inactive compounds. This is shown in eq 6:

$$\text{recall} = \text{true positives}/(\text{true positives} + \text{false negatives}) \quad (4)$$

$$\text{precision} = \text{true positives}/(\text{sum of true and false positives}) \quad (5)$$

$$\text{specificity} = \text{true negatives}/(\text{true negatives} + \text{false positives}) \quad (6)$$

The truth table recall, precision, and specificity are presented in Figure 3 for the LQS model for a threshold value of 0.4.

Acceptable values for false positive and negative rates will depend on the requirements of a particular study. It will seldom be necessary to require near 100% recall or specificity, for example, and lower values of these factors can be traded for higher precision. The threshold factor can, therefore, be tuned to the desired compromise between recall, specificity, and precision. Figure 4A shows how these metrics vary as a function of the indicator threshold for the LQS model. The graph shows that approximately 60% of the active compounds would be identified at a threshold value of 0.3, for example, as measured by the recall factor.

The metrics in Figure 4A apply only to the training set and do not necessarily predict performance outside the training set, however. One reason for this is that random structural patterns unrelated to biological activity can be observed, especially in unbalanced data such as this. In order to better predict performance of the model for compounds outside the training set, the performance of the LQS model was compared to the performance of models built using randomized data. We constructed randomized data sets for this purpose by selecting 77 compounds at random from the AERS database to match the number of significant drug−LQS pairs. A new model was then created to identify the fragment fingerprints of the randomly selected compounds, and the average of results for 50 such calculations were used for comparison to the LQS model. The resulting model for random data represents the "null hypothesis", and metrics for the model are shown in Figure 4B, where it is qualitatively clear that the LQS model is more predictive.

The difference between the LQS model and its random comparator is shown more clearly in Figure 5A in which plots of the recall for the LQS and the random models are overlaid. If one draws a vertical line at any threshold value on the *x*-axis, the distance between the lines for true and randomized data is the increase in signal over the random value. In addition, the distance between the lines can be used as a correction factor to estimate the quality of the model for predictions outside of the training set. The correction is defined in eqs 7−10. In those eqs the value "true positives$_{\text{random}}$" refers to the number of randomly selected positives that are predicted true by the model made using the randomly selected compounds as positives. When these corrections are applied to the LQS model, the results are shown in Figure 5B.

$$\text{corrected true positives}(a') = \text{true positives}_{\text{model}} - \text{true positives}_{\text{random}} \quad (7)$$

$$\text{corrected false negatives}(b') = \text{false negatives}_{\text{model}} + \text{true positives}_{\text{random}} \quad (8)$$

$$\text{corrected false positives}(c') = \text{false positives}_{\text{model}} + \text{true positives}_{\text{random}} \quad (9)$$

$$\text{corrected true negatives}(d') = \text{true negatives}_{\text{model}} - \text{true positives}_{\text{random}} \quad (10)$$

Unlike the results in Figure 4A, the graph in Figure 5B based on corrected truth-table values shows that both recall and precision reach maximum values less than 100%. The corrected recall factor reaches a maximum of approximately 30%. This value indicates that the model that will reliably identify at most 30% of the active compounds in a compound library. The maximum value for the corrected precision indicates that at best we can expect 80% of compounds selected to be true positives. The compromise between identifying true positives (recall) and minimizing false positives (precision) is optimum in the range 0.4−0.5. At threshold values below 0.4 strong patterns are identified in random data, and the categorization of active molecules at these threshold values is suspect. Corrected curves such as the one in Figure 5B are more representative of the performance of the models for data outside the training set than are the curves in Figure 4A.

The fragments found to be most significant in the LQS model are shown in Table 2. High positive coefficients mean a fragment is strongly associated with LQS, and large negative coefficients mean a fragment is associated with the absence of LQS.

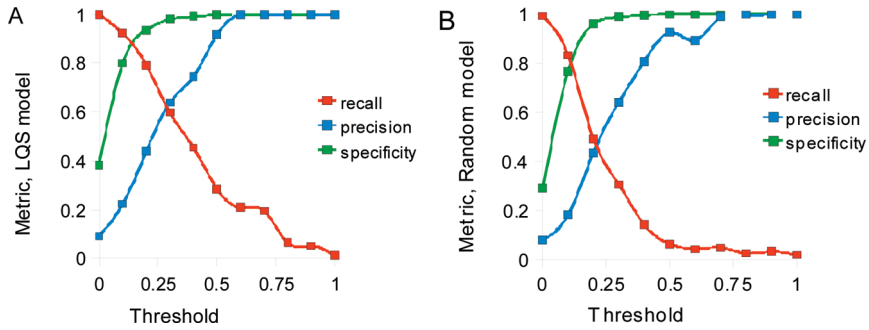

Figure 4. Metrics for the quality of the model predicting LQS adverse events. Recall, precision, and specificity are defined as in Figure 3 and eqs 4−6. Panel A: Identification of compounds with LQS liability using an LQS model. Panel B: Results for a randomized data set for comparison. From the AERS compound collection, 77 compounds were selected at random and arbitrarily defined as active. The metrics measure the efficiency of identifying these 77 random compounds using a new model constructed for the randomized data. The data shown are averages over 50 such calculations.

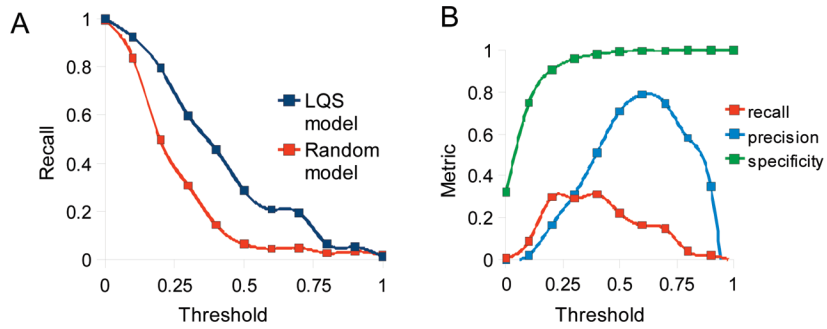

Figure 5. Comparison of models based on real and randomized data. Panel A: Comparison of the recall for the LQS model and the random model from Figure 4. Panel B: Metrics for the LQS predictive model from Figure 4A replotted using corrected truth-table values computed, as shown in eqs 7−10.

Table 2. Molecular Fragments with a Significant Association with QT Elongation

fragment	smiles	coefficient
aminosulfonamidene	C=NS(=O)(=O)N	1.16
N amino methanimidamide	N=CNN	–0.98
pyridazine	n1ncccc1	0.89
triazine-4-one	C1=C2C(=O)N=CNN2C=N1	0.73
benzobenzothiazepine	C1=CC=C2C(=C1)C=NC3=CC=CC=C3S2	0.73
1,2 benzoxazole	C1=CC=C2C(=C1)C=NO2	0.49
1,4 cyclohexadiene	C1=CCC=CC1	0.49
benzofuran	C1(OC=C2)=C2C=CC=C1	0.45
cyclohexane nitrile	C1CCCCC1(C#N)	–0.39
1,4 thiazene	C1C=CSC=C1	–0.38
3-azabicyclo[3.1.0]hexane	C12C(C1)CNC2	–0.34
4-fluorobenzene	c1cc[3](F)ccc1	0.33
azahydroxyquinoline	c1c2c(ccc1)cnNc2=O	–0.32
4-aminoquinoline	c1c2c(ccc1)nccc2N	0.31
phenanthrene	c1c2c(ccc1)ccc1c2cccc1	–0.3
secondary imine	C(C)(C)=N	–0.3
2,5-dihydro-1H-azepine	C1C=CCNC=C1	0.28
1,3-oxazinan-2-one	C1C=CCOCC1	–0.27
1,3 oxazinane	C1CCNCO1	–0.26

COUNTS

		Predicted TRUE	Predicted FALSE	*total*
Observed	**TRUE**	23	48	71
	FALSE	6	1231	1237
	total	29	1279	

Recall =	23 / 71 =	32%
Precision =	23 / 29 =	79%
Specificity =	1231 / 1237 =	100%

Figure 6. Truth table and quality metrics for prediction of TdP at an indicator threshold of 0.4. The TdP model based on AERS data is used to compute an indicator variable for each compound in the AERS collection. The indicator value varies approximately between 0 and 1, and higher values of the indicator threshold represent more stringent selection criteria. If the value of the indicator variable exceeds a threshold, 0.4 in this case, then that compound is scored as active (positive). Recall, precision, and specificity factors are defined in eqs 4−6 and are metrics for the quality of the model.

Torsade de pointes (TdP). In addition to modeling LQS as an indicator of TdP, the incidence of TdP was modeled directly. There are 1 038 reports of TdP in the AERS database and 71 compounds associated with TdP at the 95% confidence level. Only 38 of these compounds have a significant relation to the LQS adverse event. The truth table for TdP at a threshold value of 0.4 is shown in Figure 6.

The decision metrics for the TdP model as a function of threshold are shown in Figure 7A, and the corrected metrics are shown in Figure 7B. These results can be compared to results for the LQS syndrome in Figure 5A and B. The corrected metrics for the TdP model are similar to those of the LQS model and indicate that the optimum threshold for decision-making is 0.3−0.5.

The most significant fragments associated with TdP are shown in Table 3.

Analysis of a Test Set: Drugs Withdrawn From Market. The predictive ability of the model was tested on the series of drugs withdrawn from the market due to QT-related events (Table 4).[1,46] Of these compounds, astemizole and terodiline are used as standards in animal studies to test QT elongation,[47,48] and sertindole[49] and terfenadine[50] are strongly associated with LQS. We created cross-validation models based specifically on these compounds rather than randomly selected compounds to test performance against a particularly relevant test set. Five of the selected compounds, lidoflazine, terfenadine, sertindole, astemizole, and grepafloxacin were withdrawn before the start of data collection

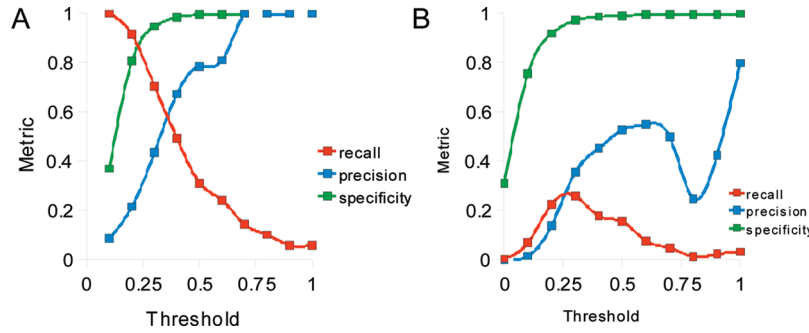

Figure 7. Metrics for the quality of the model predicting TdP adverse events. Panel A: Recall, precision, and specificity factors are defined, as in Figure 4 and eqs 4−6. Panel B: Metrics from panel A corrected as defined in eqs 7−10, based on comparison of predictions for real and randomized data sets.

Table 3. Most Significant Fragments Related to Compounds Associated with TdP

Fragment	Smiles	Coefficient
aminosulfonamidene	C=NS(=O)(=O)N	1.06
phenanthrene	c1c2c(ccc1)ccc1c2cccc1	0.85
N amino methanimidamide	N=CNN	–0.55
pyridazine	n1ncccc1	0.45
4-aminoquinoline	c1c2c(ccc1)nccc2N	0.44
3,6 diazabycyclo[3.1.0]hexane	C1C2C(N2)CN1	–0.43
1,3-oxazinan-2-one	C1=CCOC(=O)N1	–0.4
3-azabicyclo[3.1.0]hexane	C12C(C1)CNC2	–0.38
2 fluorobenzene	c1cc(F)c([R])cc1	0.36
benzofuran	C1(OC=C2)=C2C=CC=C1	0.35
1,2 dihalobenzene	c([Hal])1c([Hal])cccc1	–0.35
triazine-4-one	C1=C2C(=O)N=CNN2CC=N1	0.34
substituted propofol	CC(C)c1cccc(C(C)C)c1[R]	0.34
azahydroxyquinoline	c1c2c(ccc1)cnNc2=O	–0.27
1,4 thiazene	C1C=CSC=C1	–0.26
1,4 *N* disubstituted benzene	c1(N)ccc(N)cc1	0.26
secondary imine	C(C)(C)=N	–0.26
benzobenzothiazepine	C1=CC=C2C(=C1)C=NC3=CC=CC=C3S2	–0.24
dihydropyridine	N1C=CCC=C1	0.23
cyclohexane nitrile	C1CCCCC1(C#N)	–0.23

Table 4. Predicted Adverse Event Indicators for a Test Set of Drugs Withdrawn from the Marked Due to Cardiac Issues[a]

drug	LQS score predicted while in model	crossvalidated LQS score	TdP score predicted while in model	crossvalidated TdP score
astemizole[b]	0.8	0.5	0.6	0.2
cisapride	0.8	0.6	0.6	0.2
dofetilide	0.3	0.2	0.2	0.1
droperidol	0.5	0.2	0.3	–0.2
grepafloxacin[b]	0.6	0.4	0.6	0.5
levomethadyl	0.3	0.2	0.3	0.1
lidoflazine[b]	0.8	0.6	0.5	0.2
sertindole[b]	0.6	0.4	0.4	0.1
terfenadine[b]	0.4	0.2	0.3	–0.1

[a] Predictions performed with and without this set of molecules in the training set. [b] These compounds were withdrawn from use before the creation of the AERS and are not represented in the database. In order to include them in the LQS and TdP models, they were manually added to the set of compounds used to make the models.

for the AERS, and there is no data for these compounds in the database. We created two models for comparison of performance, one in which the older, missing compounds were manually assigned to have an association with both LQS and TdP and one in which all compounds in Table 4 were excluded in building the model for cross-validation. These nine compounds represented 12% of the active compounds of the training set. When the compounds were excluded from the training set, the LQS model correctly identified five out of nine (55%) of the withdrawn drugs at a threshold value of 0.4, which compares favorably to an expectation of 31% based on the corrected recall value (Figure 5B). The TdP model identified only one out of nine (11%), but the expected recall value is only 18% for this model (Figure 7B).

DISCUSSION

AERS Data Mining. Adverse events in the FDA databases are reported in association with the drugs that a patient is taking at the time of the event. However, simple association between drug and event does not necessarily imply cause and effect. Moreover, because of the subjective nature of the reporting system and the relatively rare occurrence of adverse events, the absence of association does not guarantee lack of cause and effect. Much of the previous work in mining these databases, therefore, has concentrated on simply identifying statistically significant drug−event combinations. This was particularly important for the older SRS database, which lacked a controlled vocabulary. With the adoption of the more structured AERS database, this is less of an issue, and we have taken a fairly straightforward and standard approach to defining significant relationships. The method simply looks for drug−event reporting frequencies that exceed the overall frequencies of appearance of the drug and the adverse event in the database, and we require that this frequency be significant at the 95% level. We are not aware of any other work that would validate our test for statistical significance specifically for LQS or TdP adverse events, but we have also applied this method to the detection of a relationship between simvastatin and cataracts, which was studied extensively by Hauben, et al. using data from the SRS database.[51] We identified a reporting ratio (defined in eq 2) of 5.8, which agrees well with ratios of 5.8−5.9 reported by this previous effort. Ultimately, the validity of the method for identifying significant drug−event associations will be determined by the utility of predictive models generated from the data, which for our models is discussed below.

Comparison of QSAR models to models derived from randomized dependent variables is a powerful method for gaining a realistic evaluation of a model and has recently been used to measure the robustness of models for carcinogenicity.[52] In the present study, correction of recall for random effects provides a much better metric than the raw recall rates. In Figure 4A, for example, we vary the threshold

used to accept the prediction of a drug adverse event association from the LQS model. As expected when testing recall against the training set, the rate reaches 100% at a threshold of 0, where there is no barrier to accepting the prediction and decreases monotonically to 0% as the threshold is increased. However, for test sets outside the training set, recall rates near 100% are an unrealistic expectation at any threshold. In particular, when nearing a threshold of 0, predictions for the test set will become more and more dominated by random effects, and recall rates should actually decrease toward 0% instead of 100%. We, therefore, use a recall rate for the training set that is corrected by the rate for randomized data, and the corrected recall for the LQS of Figure 5B shows that a more realistic expectation with a maximum expected recall of 30%. Similarly, the maximum expected precision is 80% after correcting for a random model. Depending on the needs of the analysis, such as filtering a set of compounds to eliminate compounds likely to be troublesome, one can use these corrected metrics to better select a threshold resulting in maximum recall, at the expense of more false negatives, or in maximum precision, at the expense of more true positives. This data suggests that the corrected values are a much better guide than those of the raw uncorrected metrics for the choices of threshold to optimize any selected statistic or set of statistics.

LQS and TdP Models. Using the more robust corrected models, we predict that the models will identify approximately 25% of the drugs in a compound collection that have a LQS or TdP liability and that these drugs will be identified with relatively high precision, so that the ratio of true to false positives will be approximately 2:1. In order to test this prediction, we have identified a set of nine drugs that are known to be have LQS and TdP liability and that were pulled from the market due to cardiac-related toxicities. In a cross-validation study with these compounds as the test set, between 55% (LQS) and 11% (TdP) of the molecules were correctly classified as active. These results are indication that the corrected recall metrics discussed above are appropriately conservative. Grepafloxacin is a good example of the potential use of the model. Because this drug was withdrawn from the market due to cardiac issues before inception of the AERS,[53] the AERS data does not contain information suggesting a link between the two. Nonetheless, the models correctly predict a high association of grepafloxacin with both LQS and TdP.

The difference in performance of the LQS and TdP models against the test set reflects the fact that these two events are not fully coupled based on AERS data. Only 38 compounds are linked to both models compared to 77 and 71 total compounds with statistically significant association to the individual molecules, respectively. In general, the utility of QSAR models is dependent on the number and the diversity of compounds in the training set, and this number is still relatively small in the AERS database. We can say, therefore, that the withdrawn drugs are not as well represented in the TdP training set as in the LQS training set, but whether this is an accident of the small numbers involved or perhaps represents a meaningful mechanistic distinction will require additional analysis.

Link between LQS, TdP, and hERG Binding. The LQS and TdP syndromes are thought to be linked, but there is some controversy about the strength of the relationship.[54]

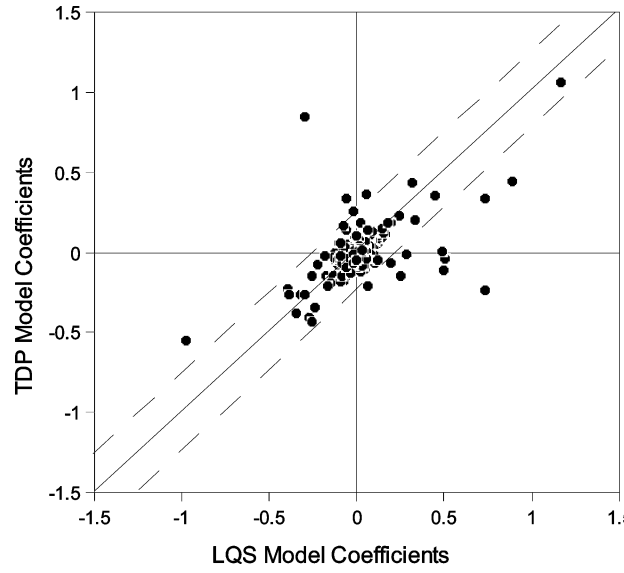

Figure 8. Comparison of fragment coefficients for the LQS and TdP models. Solid diagonal line represents the theoretical line for 1:1 correspondence of values with unit slope and 0 intercept. The standard deviation from this line is 0.12 units, and the broken lines in the graph represent ±2 standard deviations.

The syndromes are clearly distinct based solely on the drug−event data in the AERS database, and this is discussed above relative to the performance of LQS and TdP models against a test set of withdrawn drugs. This distinction cannot be taken as either proof or lack of association between the syndromes, however, since there is a subjective element to reporting drug−event combinations. The molecular fragments used to construct our models represent an opportunity to reexamine this possible relationship at a higher granularity of molecular structure. This possibility is investigated by plotting the fragment coefficients for the TdP event vs the coefficients for the LQS event (Figure 8). In the figure, the correlation of fragment coefficients is compared to the theoretical line for an exact correlation. The standard deviation from this line is 0.12 units, and the broken lines in the graph represent ±2 standard deviations. Given the relatively large lack of concordance at the molecular level, the correlation at the fragment level is unexpectedly strong. The standard deviation from the theoretical correlation does not represent any physically meaningful quantity, however, so that the observed correlation does not in itself represent proof of an association between the two adverse events. The results do, however, suggest new experiments that may be most efficient in identifying molecules for which the two events are not correlated. For example, there is a significant outlier, phenanthrene, which has a strong negative coefficient in the LQS model, −0.30, and a positive coefficient, 0.85, in the TdP model. Molecules containing this fragment are strong candidates for additional experimentation in this regard. The full list of molecules with LQS and TdP activities and the model coefficients for the fragments are provided in the Supporting Information.

Further, binding to the hERG ion channel is thought to be related to both LQS and TdP, but the strength of the relationship is also controversial. Although multiple fragment- and pharmacophore-based QSAR models for hERG binding have been reported previously (Table 1), the specific fragments and pharmacophores used differ from those used

herein. It is difficult, therefore, to compare models directly, but some consistencies are apparent. The fragments most related to hERG binding were fluorobenzene, aniline, piperidine, sulfoxyamine, and cresol in the fragment-based study,[18] with heterocyclic amines and secondary, tertiary, and quaternary aliphatic amines being generally identified across multiple studies.[12,17] The most significant fragments for LQS and TdP in this study also included heterocyclic amines; this class of fragments is, therefore, a consistent predictor of all three activities.

Similarly, studies by both Agrafiotis and Song found fluorine-containing fragments to be major structural factor for hERG binding,[18,19] and here we observe that compounds containing aromatic fluorine increase the chances of observing LQS and TdP. In the LQS model, 4-substituted fluorobenzene has a coefficient of 0.33, and in the TdP model, 2-fluorobenzene has a coefficient of 0.36. The original study by Song did not distinguish between substitution points of fluorobenzene, using only the presence of the fragment to characterize the molecules.

The links of hERG and LQS to aliphatic amines, however, are more tenuous. Although tertiary amines have been associated with hERG binding, they are present in both strong and weak hERG binders, suggesting that they are not a strong discriminant of hERG binding.[18] Consistent with results for hERG, tertiary aliphatic amines make only a weak positive contribution, and tertiary aromatic amines have a small negative coefficient in both the LQS and TdP models. Further, secondary and quaternary aliphatic amines span a range from weakly predictive to weakly counter-indicated in LQS and TdP models. Quaternary amines are identified in hERG and TdP models as a significant indicator but are not significant in the LQS model. The behavior of aliphatic amines, therefore, provide some differentiation of hERG from LQS and TdP activities, but the correlations seem to be weak and may be dominated by other factors, such as bioavailability.

CONCLUSIONS

QSAR models have traditionally been used to detect adverse events through surrogates, such as hERG binding and other in vitro toxicological analyses. This work is among the first to investigate the direct link between chemical structures and significant levels of adverse event reporting. The ability to anticipate postmarketing safety concerns could increase both the efficiency of the drug discovery and development process and the public safety by suggesting safety studies that may detect important risks before they affect the public.

The methodology and the data are not without limitations. The diversity of the compounds comprising FDA approved drugs is limited. Therefore, the ability to assess the impact of chemical structures outside these drugs may also be limited. In addition, the AERS clinical data used in this study is not collected in a uniform way, which may contain reporting biases, and suffers from other potential statistical flaws. Grepafloxacin was discontinued from use before the AERS database was initiated, for example, but is known to have a high liability for LQS. Missing data such as this reduces the generality of the model, and augmenting the AERS data with additional external data would increase the utility of future models. Nonetheless we find a detectable signal in the data that relates chemical substructures to statistically significant reports of LQS and TdP.

Normally structure−activity related events begin with known or suspected mechanisms of action, such as hERG binding, that can be measured at relatively high throughput and low cost. Use of such surrogate in vitro activities or animal models to predict clinical events introduces a source of uncertainty and limits the extensibility of such models in prediction of the clinical event. This uncertainty is reflected in the inconsistencies between QSAR models for hERG, LQS, and TdP, as referenced above. The emerging ability to relate clinical outcomes directly to drug structures reduces this uncertainty. In addition to using the human data directly, drug−event relations identified in the AERS can be used in tandem with the existing in vitro and animal models to validate the models and to differentiate between possible biological mechanisms that may underlie the toxicities of diverse drug classes.

A preliminary analysis of other adverse drug−event relationships in the AERS database (data not shown) has identified additional adverse events with equal or better statistical relationships to chemical structure than LQS or TdP, suggesting that the current work is extensible.

Supporting Information Available: The fragment set, model coefficients, database of drugs and those found from the AERS data to be related to LQS and TdP are provided. This material is available free of charge via the Internet at http://pubs.acs.org.

ACKNOWLEDGMENT

This work was carried out with open source and publicly available resources. The authors thank the developers of R, babel, postgreSQL, Linux, OpenOffice, and Java as well as the pubMed, pubChem, the FDA AERS and Drugs@FDA groups for their contributions. In addition we thank the reviewers for their insightful comments and advice in preparing the publication.

REFERENCES AND NOTES

(1) *About the Center for Drug Evaluation and Research (CDER) Report to the Nation 2005*; U.S. Food and Drug Administration: Silver Springs, MD, 2005; http://www.fda.gov/cder/reports/rtn/2005m/rtn2005.htm. Accessed June 1, 2009.
(2) Dessertenne, F. La tachycardie ventriculaire à deux foyers opposés variables. *Arch Mal Coeur Vaiss.* **1966**, *59*, 263–72.
(3) *Guidance for Industry E14 Clinical Evaluation of QT/QTc Interval Prolongation and Proarrythmic Potential for Non-Antiarrhyhmic Drugs*; U.S. Food and Drug Administration Office of Training and Communication Rockville, MD; http://www.fda.gov/cber/gdlns/iche14qtc.htm. Accessed June 1, 2009.
(4) Keating, M. T.; Sanguinetti, M. C. Molecular genetic insights into cardiovascular disease. *Science* **1996**, *272*, 681–685.
(5) Shah, R. R. The significance of QT interval in drug development. *Br. J. Clin. Pharmacol.* **2002**, *54*, 188–202.
(6) Chadwick, D. J. *The HERG Cardiac Potassium Channel: Structure, Function And Long QT Syndrome*; John Wiley and Sons: New York, 2005.
(7) Sanguinetti, M. C.; Jiang, C.; Curran, M. E.; Keating, M. T. A mechanistic link between an inherited and an acquired cardiac arrhythmia: HERG encodes the IKr potassium channel. *Cell* **1995**, *81*, 299–307.
(8) Mitcheson, J. S.; Chen, J.; Lin, M.; Culberson, C.; Sanguinetti, M. C. A structural basis for drug-induced long QT syndrome. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 12329–12333.
(9) Pearlstein, R. A.; Vaz, R. J.; Kang, J.; Chen, X.-L.; Preobrazhenskaya, M.; Shchekotikhin, A. E.; Korolev, A. M.; Lysenkova, L. N.;

Miroshnikova, O. V.; Hendrix, J.; Rampe, D. Characterization of HERG potassium channel inhibition using CoMSiA 3D QSAR and homology modeling approaches. *Bioorg. Med. Chem. Lett.* **2003**, *13*, 1829–1835.
(10) Cavalli, A.; Poluzzi, E.; De Ponti, F.; Recanatini, M. Toward a pharmacophore for drugs inducing the long QT syndrome: insights from a CoMFA study of HERG K(+) channel blockers. *J. Med. Chem.* **2002**, *45*, 3844–3853.
(11) Ekins, S.; Crumb, W. J.; Sarazan, R. D.; Wikel, J. H.; Wrighton, S. A. Three-Dimensional Quantitative Structure-Activity Relationship for Inhibition of Human Ether-a-Go-Go-Related Gene Potassium Channel. *J. Pharmacol. Exp. Ther.* **2002**, *301*, 427–434.
(12) Zolotoy, A. B.; Plouvier, B. P.; Beatch, G. B.; Hayes, E. S.; Wall, R. A.; Walker, M. J. A. Physicochemical determinants for drug induced blockade of HERG potassium channels: Effect of charge and charge shielding. *Curr. Med. Chem. Cardiovasc. Agents* **2003**, *1*, 225–241.
(13) Duffy, E. M.; Jorgensen, W. L. Prediction of Properties from Simulations: Free Energies of Solvation in Hexadecane, Octanol, and Water. *J. Am. Chem. Soc.* **2000**, *122*, 2878–2888.
(14) Roche, O.; Trube, G.; J., Z.; Pflimlin, P.; Alanine, A.; Schneider, G. A virtual screening method for prediction of the HERG potassium channel liability of compound libraries. *ChemBioChem* **2002**, *3*, 455–459.
(15) Keseru, G. M. Prediction of hERG potassium channel affinity by traditional and hologram QSAR methods. *Bioorg. Med. Chem. Lett.* **2003**, *13*, 2773–2775.
(16) Aronov, A. M.; Goldman, B. B. A model for identifying HERG K+ channel blockers. *Bioorg. Med. Chem.* **2004**, *12*, 2307–2315.
(17) Bains, W.; Basman, A.; White, C. hERG binding specificity and binding site structure: evidence from a fragment-based evolutionary computing SAR study. *Prog. Biophys. Mol. Bio.* **2004**, *86*, 205–233.
(18) Song, M.; Clark, M. Development and Evaluation of an in Silico Model for hERG Binding. *J. Chem. Inf. Model.* **2006**, *46*, 392–400.
(19) Seierstad, M.; Agrafiotis, D. K. A QSAR Model of hERG Binding Using a Large, Diverse, and Internally Consistent Training Set. *Chem. Biol. Drug Des.* **2006**, *67*, 284–296.
(20) AERS Data; National Technical Information Center (NTIS): Alexandria: VA; http://www.ntis.gov. Accessed June 1, 2009. Pharmapendium; Elsevier B.V.: New York, NY; http://www.pharmapendium.com. Accessed June 1, 2009.
(21) Adverse Event Reporting System; U.S. Food and Drug Administration: Silver Springs, MD; http://www.fda.gov/cder/aers. Accessed June 1, 2009.
(22) *MedDRA the Medical Dictionary for Regulatory Activities*; Northrop Grumman Corp.: Chantilly VA, 2008.
(23) DuMouchel, W. Bayesian data mining in large frequency tables, with an application to the FDA spontaneous reporting system. *J. Am. Stat. Assoc.* **1999**, *53*, 177–190.
(24) Almenoff, J. S.; DuMouchel, W.; Kindman, L. A.; Yang, X.; Fram, D. M. Disproportionality analysis using empirical Bayes data mining: a tool for the evaluation of drug interactions in the post-marketing setting. *Pharmacoepidemiol. Drug Saf.* **2003**, *12*, 517–521.
(25) Hammond, I. W.; Rich, D. S.; Gibbs, T. G. Effect of consumer reporting on signal detection: using disproportionality analysis. *Expert Opin. Drug Saf.* **2007**, *6*, 705–712.
(26) O'Neill, R. T.; Szarfman, A. Some FDA perspectives on data mining for pediatric safety assessment. *Curr. Ther. Res. Clin. Exp.* **2001**, *62*, 650–663.
(27) Szarfman, A.; Machado, S. G.; O'Neill, R. T. 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**, *6*, 381–392.
(28) van Puijenbroek, E. P.; Diemont, W. L.; van Grootheest, K. Application of Quantitative Signal Detection in the Dutch Spontaneous Reporting System for Adverse Drug Reactions. *Drug Saf.* **2003**, *5*, 293–301.
(29) Chen, Y.; Guo, J.; Steinbuch, M.; Lin, X.; Buncher, C. R.; Patel, N. C. Comparison of Sensitivity and Timing of Early Signal Detection of Four Frequently Used Signal Detection Methods: An Empirical Study Based on the US FDA Adverse Event Reporting System Database. *Pharm. Med.* **2008**, *6*, 359–365.
(30) Andrew, B. Bayesian confidence propagation neural network. *Drug Saf.* **2007**, *7*, 623–625.
(31) Waller, P.; van Puijenbroek, E.; Egberts, A.; Evans, S. The reporting odds ratio versus the proportional reporting ratio: 'deuce'. *Pharmacoepidemiol. Drug Saf.* **2003**, *8*, 52–526.
(32) Strombom, I.; Wernicke, J. F.; Seeger, J.; D'Souza, D. N.; Acharya, N. Hepatic Effects of Duloxetine-III: Analysis of Hepatic Events Using External Data Sources. *Curr. Drug Saf.* **2008**, *3*, 154–162.
(33) Matthews, E. J.; Kruhlak, N. L.; Weaver, J. L.; Benz, R. D.; Contrera, J. F. Assessment of the health effects of chemicals in humans: II. Construction of an adverse effects database for QSAR modeling. *Curr. Drug Discov. Tech.* **2004**, *1*, 243–254.
(34) Rodgers, A. D.; Zhu, H.; Rusyn, I.; Tropsha, A. QSAR Modleing of Human Liver Adverse Effects Database Using kNN Method. Presented at Society of Toxicology 47th Annual Meeting and ToxExpo, Seattle Washington, March 16−20, 2008; Abstract 244.
(35) PostgreSQL relational database, Version 8.2.1; University of California at Berkeley: Berkeley, CA, 2008.
(36) Drugs@FDA Data Files; U.S. Food and Drug Administration: Silver Springs, MD; http://www.fda.gov/cder/drugsatfda/datafiles/. Accessed June 1, 2009.
(37) Frid, A. A.; Matthews, E. J.; Kruhlak, N. L.; Benz, R.; Contrera, J. F. Adverse Effects of Pharmaceuticals: A. Construction of a Relational Databaes of Adverse Cardiological Effects. Using *FDA Archives, Pharmapendium, and Public Sources*. Proceedings of the Society of Toxicology 47th Annual Meeting and ToxExpo, Seattle, Washington, March 16−20, 2008; Abstract 242.
(38) Glasgow, J. F. Reye's syndrome: the case for a causal link with aspirin. *Drug Saf.* **2006**, *12*, 111–121.
(39) Clark, M. Generalized Fragment-Substructure Based Property Prediction Method. *J. Chem. Inf. Model.* **2005**, *45*, 30–38.
(40) Kier, L. B.; Hall, L. H. *Molecular Connectivity in Structure-Activity Analysis*; John Wiley & Sons: New York, 1986.
(41) Bocker, A. Toward an Improved Clustering of Large Data Sets Using Maximum Common Substructures and Topological Fingerprints. *J. Chem. Inf. Model.* **2008**, *48*, 2097–2107. (a) Xue, L.; Godden, J. W.; Bajorath, J. Evaluation of descriptors and mini-fingerprints for the identification of molecules with similar activity. *J. Chem. Inf. Comput. Sci.* **2000**, *5*, 1227–1234.
(42) Kratochwil, N. A.; Huber, W.; Muller, F.; Kansy, M.; Gerber, P. R. Predicting plasma protein binding of drugs: a new approach. *Biochem. Pharmacol.* **2002**, *64*, 1355–1374.
(43) The R Project *for Statistical Computing* 2.2.0; The R Foundation: Vienna, 2005.
(44) Clark, M.; Cramer, R. D. The Probability of Chance Correlation Using Partial Least Squares (PLS). *Quant. Struct.-Act. Relat.* **1993**, *12*, 137–145.
(45) FDA Model Applier Leadscope, version 1.0; Leadscope, Inc.: Columbus OH, 2008.
(46) Katchman, A. N.; Koerner, J.; Tosaka, T.; Woosley, R. L.; Ebert, S. N. Comparative evaluation of HERG currents and QT intervals following challenge with suspected torsadogenic and nontorsadogenic drugs. *J. Pharmacol. Exp. Ther.* **2006**, *316*, 1098–1106.
(47) Hanton, G.; Yvon, A.; Racaud, A. Temporal variability of QT interval and changes in T wave morphology in dogs as markers of the clinical risk of drug-induced proarrhythmia. *J. Pharmacol. Toxicol. Methods* **2008**, *3*, 194–201.
(48) Fossa, A. A. Assessing QT prolongation in conscious dogs: validation of a beat-to-beat method. *Pharmacol. Therapeut.* **2008**, *2*, 133–140.
(49) Lançon, C.; Toumi, M.; Sapin, C.; Hansen, K. The Sertindole Safety Survey: a retrospective analysis under a named patient use programme in Europe. *BMC Psychiatry* **2008**, *8*, 1–8.
(50) Pinney, S. P.; Koller, B. S.; Franz, M. R.; Woolsey, R. L. Terfenadine increases the QT interval in isolated guinea pig heart. *J. Cardiovasc. Pharmacol.* **1995**, *1*, 30–34.
(51) Hauben, M.; Reich, L.; Gerrits, C. M.; Younus, M. Illusions of objectivity and a recommendation for reporting data mining results. *Eur. J. Clin. Pharmacol.* **2007**, *63*, 517–521.
(52) Hao, Z.; Rusyn, I.; Richard, A.; Tropsha, A. Use of Cell Viability Assay Data Improves the Prediction Accuracy of Conventional Quantitative Structure-Activity Relationship Models of Animal Carcinogenicity. *Environ. Health Perspect.* **2008**, *116*, 506–513.
(53) Ball, P. Quinoline-induced QT interval prolongation: a not-so-unexpected class effect. *J. Antimicrob. Chemother.* **2000**, *45*, 557–559.
(54) Morganroth, J.; Gussak, I. *Cardiac Safety of Noncardiac Drugs: Practical Guidelines for Clinical Research and Drug Development*. Springer: New York 2005.

CI900116Q