Complex versus simple models: ion-channel cardiac toxicity prediction (#21772)

First submission

Editor guidance

Please submit by **29 Nov 2017** for the benefit of the authors (and your $200 publishing discount).

Structure and Criteria
Please read the 'Structure and Criteria' page for general guidance.

Raw data check
Review the raw data. Download from the materials page.

Image check
Check that figures and images have not been inappropriately manipulated.

Privacy reminder: If uploading an annotated PDF, remove identifiable information to remain anonymous.

Files
Download and review all files from the materials page.

3 Figure file(s)
1 Table file(s)
1 Raw data file(s)

Structure and Criteria

2

Structure your review

The review form is divided into 5 sections.
Please consider these when composing your review:

1. **BASIC REPORTING**
2. **EXPERIMENTAL DESIGN**
3. **VALIDITY OF THE FINDINGS**
4. General comments
5. Confidential notes to the editor

You can also annotate this PDF and upload it as part of your review

When ready submit online.

Editorial Criteria

Use these criteria points to structure your review. The full detailed editorial criteria is on your guidance page.

BASIC REPORTING

- Clear, unambiguous, professional English language used throughout.
- Intro & background to show context. Literature well referenced & relevant.
- Structure conforms to PeerJ standards, discipline norm, or improved for clarity.
- Figures are relevant, high quality, well labelled & described.
- Raw data supplied (see PeerJ policy).

EXPERIMENTAL DESIGN

- Original primary research within Scope of the journal.
- Research question well defined, relevant & meaningful. It is stated how the research fills an identified knowledge gap.
- Rigorous investigation performed to a high technical & ethical standard.
- Methods described with sufficient detail & information to replicate.

VALIDITY OF THE FINDINGS

- Impact and novelty not assessed. Negative/inconclusive results accepted. *Meaningful* replication encouraged where rationale & benefit to literature is clearly stated.
- Data is robust, statistically sound, & controlled.
- Conclusions are well stated, linked to original research question & limited to supporting results.
- Speculation is welcome, but should be identified as such.

Standout reviewing tips

The best reviewers use these techniques

Tip	*Example*
Support criticisms with evidence from the text or from other sources	*Smith et al (J of Methodology, 2005, V3, pp 123) have shown that the analysis you use in Lines 241-250 is not the most appropriate for this situation. Please explain why you used this method.*
Give specific suggestions on how to improve the manuscript	*Your introduction needs more detail. I suggest that you improve the description at lines 57- 86 to provide more justification for your study (specifically, you should expand upon the knowledge gap being filled).*
Comment on language and grammar issues	*The English language should be improved to ensure that an international audience can clearly understand your text. Some examples where the language could be improved include lines 23, 77, 121, 128 – the current phrasing makes comprehension difficult.*
Organize by importance of the issues, and number your points	*1. Your most important issue* *2. The next most important item* *3. …* *4. The least important points*
Please provide constructive criticism, and avoid personal opinions	*I thank you for providing the raw data, however your supplemental files need more descriptive metadata identifiers to be useful to future readers. Although your results are compelling, the data analysis should be improved in the following ways: AA, BB, CC*
Comment on strengths (as well as weaknesses) of the manuscript	*I commend the authors for their extensive data set, compiled over many years of detailed fieldwork. In addition, the manuscript is clearly written in professional, unambiguous language. If there is a weakness, it is in the statistical analysis (as I have noted above) which should be improved upon before Acceptance.*

Complex versus simple models: ion-channel cardiac toxicity prediction

Hitesh Mistry Corresp. 1

1 Division of Pharmacy, University of Manchester, Manchester, United Kingdom

Corresponding Author: Hitesh Mistry
Email address: hitesh.mistry@manchester.ac.uk

There is growing interest in applying detailed mathematical models of the heart for ion-channel related cardiac toxicity prediction. However, a debate as to whether such complex models are required exists. Here an assessment in the predictive performance between two established cardiac models, *gold-standard* and *cardiac safety simulator*, and a simple linear model B_{net} was conducted. Three ion-channel data-sets were extracted from literature. Each compound was designated a cardiac risk category based on information within CredibleMeds. The predictive performance of each model within each data-set was assessed via a leave-one-out cross validation. In two of the data-sets B_{net} performed equally as well as the leading cardiac model, *cardiac safety simulator*, both of these outperformed the *gold-standard* model. In the 3rd data-set, which contained the most detailed ion-channel pharmacology, B_{net} outperformed both cardiac models. These results highlight the importance of bench-marking models but also encourage the development of simple models.

Complex versus simple models: ion-channel cardiac toxicity prediction

Hitesh B. Mistry[1]

[1]Division of Pharmacy, University of Manchester, UK

Email correspondence: hitesh.mistry@manchester.ac.uk

Abstract

There is growing interest in applying detailed mathematical models of the heart for ion-channel related cardiac toxicity prediction. However, a debate as to whether such complex models are required exists. Here an assessment in the predictive performance between two established cardiac models, *gold-standard* and *cardiac safety simulator*, and a simple linear model B_{net} was conducted. Three ion-channel data-sets were extracted from literature. Each compound was designated a cardiac risk category based on information within CredibleMeds. The predictive performance of each model within each data-set was assessed via a leave-one-out cross validation. In two of the data-sets B_{net} performed equally as well as the leading cardiac model, *cardiac safety simulator*, both of these outperformed the *gold-standard* model. In the 3rd data-set, which contained the most detailed ion-channel pharmacology, B_{net} outperformed both cardiac models. These results highlight the importance of benchmarking models but also encourage the development of simple models.

Introduction

There is a growing belief within the pharmaceutical industry that in order to improve predictions of future experiments more detailed mathematical models of biology are required (Peterson & Riggs, 2015; Knight-Schrijver et al., 2016). However by including more detail not only does the number of parameters that need to be estimated increase but so does the degree of structural uncertainty i.e. the degree of ~~confidence~~ in the actual structure of the equations (Beven, 2005). The objective of this study is to look at this issue within the field of drug induced ion-channel cardiac toxicity. This area has a well-defined question relating to prediction where a debate about the complexity of the model needed is ongoing.

The question of interest is: can high-throughput ion-channel screening data predict the propensity for a given type of arrhythmia, torsades de pointes (TdeP), in humans (Mirams & Noble, 2011). In order to answer this question the literature is divided in terms of the complexity of the modelling approach required (Mistry, 2017). The complex models used are based on biophysical models which describe the changes in ionic currents over time within a single cardiac cell (Trayanova, 2011). They contain 100s of parameters and 10s of differential equations. The drug input into these models involves scaling ion-channel conductance's by the amount of block at a given drug concentration.(Brennan, Fink & Rodriguez, 2009) Two biophysical models that have gained favour in the literature are the *gold-standard*, as described by Zhou *et al.* (Zhou et al., 2015), model by O'Hara *et al.* (O'Hara et al., 2011) which is being put forward for use by regulatory agencies (Colatsky et al., 2016) and another, by TenTusscher *et al.* (ten Tusscher & Panfilov, 2006), forms a key part of the *cardiac safety simulator* (Glinka & Polak, 2015). An alternative simpler model being put forward analyses the net difference, *via* a linear combination, in drug block of the ion-channels of interest, termed B_{net} (Mistry, 2017). Thus it is based on a higher level of abstraction than biophysical models and focusses on known biology/pharmacology.

Two previous studies have shown that the simple model is likely to give similar predictive performance to the more complex models (Mistry, Davies & Di Veroli, 2015; Mistry, 2017). However in those studies the definition of torsadegenic risk lacked consistency as each data-set used different criterion. Furthermore those studies were based only on 3 ion-channels, hERG, Cav 1.2 and Nav 1.5 peak, and so the dimensionality of ion-channel space can be considered narrow.

In this study we analyse the predictive performance of the *gold standard*, *cardiac safety simulator* and B_{net} models using a consistent and reliable definition of torsadegenic risk from CredibleMeds (Woosley, RL & Romero, KA; Woosley et al., 2017) across three literature data-sets (Mirams et al., 2011; Kramer et al., 2013; Crumb Jr. et al., 2016). Two of these data-sets, Mirams *et al.* (Mirams et al., 2011) and Kramer *et al.* (Kramer et al., 2013), measured drug effect against 3 ion-channels, hERG, Cav 1.2 and Nav 1.5 peak. The third and latest data-set, from Crumb *et al.* (Crumb Jr. et al., 2016), considers drug effect on 7 ion-channels, hERG (IKr), KCNQ1 + KCNE1 (IKs), Kv4.3 (Ito), Kir2.1 (IK1), Cav 1.2 (ICaL), Nav1.5 peak (INa) and Nav1.5 late (INaL), the largest number studied so far.

By using a consistent definition of torsdagenic risk across different data-sets the analysis conducted will provide a detailed view on the performance of each model. Thus enabling scientists to make a more informed decision about which modelling approach is likely to be the most useful for the prediction problem ~~considered.~~

Methods

Data

Ion-channel IC50 values, defined as concentration of drug the reduces the flow of current by 50%, were collected from three publications (Mirams et al., 2011; Kramer et al., 2013; Crumb Jr. et al., 2016). Compounds within those data-sets were classified as being TdeP positive or TdeP negative based on their classification by Credible Meds (Woosley, RL & Romero, KA; Woosley et al., 2017). A compound was classed TdeP positive if it was classified as known (KR) or ~~partial~~ risk (PR) on CredibleMeds which refers to whether there is substantial evidence the drug causes QT prolongation and/or TdeP. A compound is classed as TdeP negative if it was classified as conditional risk (CR), the risk of TdeP is conditional on other factors e.g. drug-drug interaction, or no risk if it wasn't listed (NR) as was done by Kramer *et al.*(Kramer et al., 2013). All data is provided in supplemental material.

Model input data

The percentage block against a given ion-channel inputted into all models was calculated based on the effective therapeutic concentration (EFTPC), which was provided in the original articles, using a pore block model,

$$\%\,Block = \frac{1}{1 + \frac{IC50}{EFTPC}} \quad (1)$$

Models

Single cell cardiac model simulations

The AP predict platform (Williams & Mirams, 2015) was used to simulate the gold-standard and cardiac safety simulator models in all cases except for one simulation study. A MATLAB version of the *d-standard* model available on the Rudylab website (http://rudylab.wustl.edu) was used when simulating the block of 7 ion-channels since that model on AP predict does not allow blocking of INaL – current measured in the Crumb *et al.* data-set. The default settings within the AP predict platform were used i.e. 1Hz pacing for 5 minutes with the APD90, time taken for the action potential to repolarise by 90%, recorded using the last cycle. The same protocol was applied in MATLAB when exploring the 7 ion-channels within the O'Hara model i.e. 1Hz pacing for 5 minutes with APD90 recorded using the last cycle. In all simulations drug block was initiated at the beginning of simulations.

B_{net}

We define the difference in block between repolarisation and depolarisation ion-channels as B_{net},

$$B_{net} = \sum_{i=1}^{n} R_i - \sum_{j=1}^{m} D_j$$

where R_i and D_j represent the percentage block against repolarisation and depolarisation ion-channels respectively for a specific drug.

Classification evaluation

For each compound the percentage change in APD90 compared to control (no block) from the biophysical model simulations was recorded as was the B_{net} value. These values were then placed within a logistic regression analysis to assess their correlative value to TdeP risk. This was done via a leave one out cross validation (LOOCV). This involves training a classifier to *n-1* compounds and testing on the n^{th}. Thus all compounds perform part of the test-set. The predicted probability of risk for each test compound is then used to generate a ROC AUC (area under the receiver operating characteristic curve) and is reported as was done previously (Cummins Lancaster & Sobie, 2016).

Results

Data

The total number of compounds that are TdeP positive (CredibleMeds known (KR) or ~~partial~~ (PR) risk) versus TdeP negative (CredibleMeds conditional risk (CR) or not listed (NR)) across the 3 data-sets of interest can be seen in Figure 1. Although the total number of compounds differs from one data-set to another the proportions that are KR or PR does not.

The distribution of block against each ionic current, at the effective therapeutic concentration (EFTPC), across all data-sets can be seen in Figure 2. The plots show that the activity of the compounds is greatest against IKr across all data-sets. After IKr, ICaL appears to be the next channel for which a substantial amount of activity is seen. A somewhat surprising result is the degree of activity against INaL but not INa in the Crumb *et al.* data-set. The amount of activity against INaL in that data-set mirrors that of ICaL activity.

Classification Evaluation

The results of the leave-one-out cross validation for each data-set using various models can be seen graphically in Figure 3 and also in Table 1. For the Mirams *et al.* data-set it's noticeable that the *gold-standard* model performs no better than using just block against hERG alone neither of which are better than random chance. Both the *cardiac safety simulator* and B_{net} show a similar improvement over using just hERG block.

Moving onto the Kramer *et al.* data-set the performance of all models improves dramatically over the Mirams *et al.* data-set. Here all 3 models show superior performance over just hERG block. Note that again the *gold-standard* model performance is not as high as B_{net} or the *cardiac safety simulator*. In addition the difference between B_{net} and the *cardiac safety simulator* is negligible.

Within the latest data-set by Crumb *et al.* the performance of all models, when using only 3 ion-channels, drops to a level similar to that seen within the Mirams *et al.* data-set. The key difference between the results between those two data-sets is that the *gold-standard* model now shows similar performance to the *cardiac safety simulator*. Furthermore neither biophysical model performs overly better than using hERG block. B_{net} however appears to give reasonable performance again and appears to show an improvement over using hERG block. Finally if we move onto using all the ion-channel data from the Crumb *et al.* data-set the difference in performance between the models is quite striking. B_{net}'s performance improves with the addition of more information whereas there is little improvement in either biophysical model. The cardiac safety simulator may even have regressed slightly.

In summary the results show that the performance of the models is data-set dependent. However, within each data-set the B_{net} model performs just as well if not better than leading biophysical models.

Discussion

There appears to be a strong belief within the field of ion-channel cardiac drug toxicity that large scale cardiac models are required to answer a well-defined question (Colatsky et al., 2016): can high-throughput ion-channels screening data predict the torsaedgenic risk of a drug in man? The evidence base, that suggests that large-scale models perform better than simpler models for this question, simply does not exist. As previous studies have shown that the performance of the large-scale cardiac models can be mirrored by simpler models (Mistry, 2017). Furthermore, the simpler model may have potential to out-perform large-scale cardiac models.

There were two major caveats in those previous studies. The first relates to the definition of torsadegenic risk, different databases were used, which has been debated within the literature (Wiśniowska & Polak, 2017). Within this study the classification was based on information from CredibleMeds (Woosley et al., 2017). Their classification is based on an extensive search of both the literature and public databases and are well known to the clinical community. Another advantage of the CredibleMeds classification is that they do not have a vested interest in the application of mathematical models within drug development.

The second caveat relates to the dimensionality of the ion-channel space, only 3 ion-channels were considered in previous studies (Mirams et al., 2011; Kramer et al., 2013). Therefore an understanding as to how generalizable the inferences were on those data-sets to larger dimensions was unknown. This caveat was addressed here by considering a data set by Crumb *et al.* which measured the drug affinity for 7 ion-channels (Crumb Jr. et al., 2016) in

addition to the previous data-sets using only 3 ion-channels (Mirams et al., 2011; Kramer et al., 2013).

Both of these caveats were addressed within this study. Three models were evaluated against the data-sets: 1) the *gold-standard* (Zhou et al., 2015) single cell model by O'Hara *et al.* (O'Hara et al., 2011); 2) the single cell model by TenTusscher *et al.* (ten Tusscher & Panfilov, 2006) which is used within the *cardiac safety simulator* (Glinka & Polak, 2015); 3) a linear model evaluating the net difference in block between ion-channels involved in repolarising and depolarising the action potential, B_{net} (Mistry, 2017). Each model was assessed via a leave-one-out cross validation. (Note that prospective assessment of models is not possible within this field since this would involve developing compounds with a TdeP risk which can be considered unethical.) In addition to using outputs from the aforementioned models within the classification exercise the amount of block against hERG channel was used as a naïve benchmark.

Overall the results showed that B_{net} was equal if not superior to the biophysical models. The key findings were as follows. Within the Mirams *et al.* data-set the *gold-standard* model was no better than hERG block neither of which was better than random chance. In the largest data-set, by Kramer *et al.*, all models show good performance and highlighted the benefit of measuring more than hERG. When using information on 7 ion-channels within the Crumb *et al.* data-set the performance of B_{net} was greater than that of the biophysical models. Both of which showed no improvement in performance when moving from 3 to 7 ion-channels unlike B_{net}. Furthermore the performance of the biophysical models was not all that superior to using only hERG block. In summary the only model which consistently showed the benefit of measuring more than hERG was B_{net}.

These results may appear surprising but are not uncommon in prediction problems in other fields (Makridakis & Hibon, 2000; Green & Armstrong, 2015). The key reason why complex models are not necessarily more predictive than simpler models is due to model error i.e. error in the structure of the model itself (Beven, 2005). The concept of model error has not been discussed at all within the cardiac modelling field. Thus the effect of model error on predictivity is largely unknown, although in other fields it tends to dominate prediction uncertainty (Orrell et al., 2001; Refsgaard et al., 2006).

A key caveat of the analysis conducted is that the data-sets used may be too small to understand how large a discrepancy there truly is between the different models. However it is hoped that by continuing to assess new data-sets as they become available that the community will eventually have a comprehensive compound list. Other caveats that relate to the B_{net} model itself are that it doesn't consider the kinetics of blocking which has been highlighted as an important factor (Di Veroli et al., 2014). However these studies have been on a small numbers of compounds and so a true assessment of the importance of kinetics cannot be determined from those studies alone. If sufficient evidence regarding the importance of drug kinetics does eventually become available the B_{net} model can ~~first~~ be adapted in ~~one of~~ two possible ways: i) make its variables time-dependent or ii) introduce a scaling factor which accounts for the type of modulation (e.g. slow versus fast etc.). Thus

there is a way to improve the model by considering kinetics of drug block if sufficient evidence suggests this will improve predictive/explanatory power.

Conclusion

In summary the study conducted here highlights the importance of benchmarking. Furthermore it highlights that simple mechanistic models can not only give similar performance to large-scale mechanistic models but can out-perform them. Finally it is hoped this study highlights that there is more than one solution to a problem and that although the question and quality of data dictates the modelling approach it should not dictate the size of the model.

References

Beven K. 2005. On the concept of model structural error. *Water Science and Technology: A Journal of the International Association on Water Pollution Research* 52:167–175.

Brennan T., Fink M., Rodriguez B. 2009. Multiscale modelling of drug-induced effects on cardiac electrophysiological activity. *European Journal of Pharmaceutical Sciences: Official Journal of the European Federation for Pharmaceutical Sciences* 36:62–77. DOI: 10.1016/j.ejps.2008.09.013.

Colatsky T., Fermini B., Gintant G., Pierson JB., Sager P., Sekino Y., Strauss DG., Stockbridge N. 2016. The Comprehensive in Vitro Proarrhythmia Assay (CiPA) initiative — Update on progress. *Journal of Pharmacological and Toxicological Methods* 81:15–20. DOI: 10.1016/j.vascn.2016.06.002.

Crumb Jr. WJ., Vicente J., Johannesen L., Strauss DG. 2016. An evaluation of 30 clinical drugs against the comprehensive in vitro proarrhythmia assay (CiPA) proposed ion channel panel. *Journal of Pharmacological and Toxicological Methods* 81:251–262. DOI: 10.1016/j.vascn.2016.03.009.

Cummins Lancaster M., Sobie EA. 2016. Improved prediction of drug-induced Torsades de Pointes through simulations of dynamics and machine learning algorithms. *Clinical Pharmacology and Therapeutics*. DOI: 10.1002/cpt.367.

Di Veroli GY., Davies MR., Zhang H., Abi-Gerges N., Boyett MR. 2014. hERG Inhibitors with Similar Potency But Different Binding Kinetics Do Not Pose the Same Proarrhythmic Risk: Implications for Drug Safety Assessment. *Journal of Cardiovascular Electrophysiology* 25:197–207. DOI: 10.1111/jce.12289.

Glinka A., Polak S. 2015. QTc modification after risperidone administration – insight into the mechanism of action with use of the modeling and simulation at the population level approach. *Toxicology Mechanisms and Methods* 25:279–286. DOI: 10.3109/15376516.2015.1025346.

Green KC., Armstrong JS. 2015. Simple versus complex forecasting: The evidence. *Journal of Business Research* 68:1678–1685. DOI: 10.1016/j.jbusres.2015.03.026.

Knight-Schrijver VR., Chelliah V., Cucurull-Sanchez L., Le Novère N. 2016. The promises of quantitative systems pharmacology modelling for drug development. *Computational and Structural Biotechnology Journal* 14:363–370. DOI: 10.1016/j.csbj.2016.09.002.

Kramer J., Obejero-Paz CA., Myatt G., Kuryshev YA., Bruening-Wright A., Verducci JS., Brown AM. 2013. MICE models: superior to the HERG model in predicting Torsade de Pointes. *Scientific Reports* 3:2100. DOI: 10.1038/srep02100.

Makridakis S., Hibon M. 2000. The M3-Competition: results, conclusions and implications. *International Journal of Forecasting* 16:451–476. DOI: 10.1016/S0169-2070(00)00057-1.

Mirams GR., Cui Y., Sher A., Fink M., Cooper J., Heath BM., McMahon NC., Gavaghan DJ., Noble D. 2011. Simulation of multiple ion channel block provides improved early prediction of compounds' clinical torsadogenic risk. *Cardiovascular Research* 91:53–61. DOI: 10.1093/cvr/cvr044.

Mirams GR., Noble D. 2011. Is it time for in silico simulation of drug cardiac side effects? *Annals of the New York Academy of Sciences* 1245:44–47. DOI: 10.1111/j.1749-6632.2011.06324.x.

Mistry H. 2017. Complexity vs. Simplicity: The Winner Is? *Clinical Pharmacology & Therapeutics* 101:326–326. DOI: 10.1002/cpt.503.

Mistry HB., Davies MR., Di Veroli GY. 2015. A new classifier-based strategy for in-silico ion-channel cardiac drug safety assessment. *Frontiers in Pharmacology* 6:59. DOI: 10.3389/fphar.2015.00059.

O'Hara T., Virág L., Varró A., Rudy Y. 2011. Simulation of the Undiseased Human Cardiac Ventricular Action Potential: Model Formulation and Experimental Validation. *PLoS Comput Biol* 7:e1002061. DOI: 10.1371/journal.pcbi.1002061.

Orrell D., Smith L., Barkmeijer J., Palmer TN. 2001. Model error in weather forecasting. *Nonlin. Processes Geophys.* 8:357–371. DOI: 10.5194/npg-8-357-2001.

Peterson M., Riggs M. 2015. FDA Advisory Meeting Clinical Pharmacology Review Utilizes a Quantitative Systems Pharmacology (QSP) Model: A Watershed Moment? *CPT: Pharmacometrics & Systems Pharmacology* 4:189–192. DOI: 10.1002/psp4.20.

Refsgaard JC., van der Sluijs JP., Brown J., van der Keur P. 2006. A framework for dealing with uncertainty due to model structure error. *Advances in Water Resources* 29:1586–1597. DOI: 10.1016/j.advwatres.2005.11.013.

Trayanova NA. 2011. Whole-Heart Modeling Applications to Cardiac Electrophysiology and Electromechanics. *Circulation Research* 108:113–128. DOI: 10.1161/CIRCRESAHA.110.223610.

ten Tusscher KHWJ., Panfilov AV. 2006. Alternans and spiral breakup in a human ventricular tissue model. *American Journal of Physiology. Heart and Circulatory Physiology* 291:H1088-1100. DOI: 10.1152/ajpheart.00109.2006.

Williams G., Mirams GR. 2015. A web portal for in-silico action potential predictions. *Journal of Pharmacological and Toxicological Methods* 75:10–16. DOI: 10.1016/j.vascn.2015.05.002.

Wiśniowska B., Polak S. 2017. Am I or am I not proarrhythmic? Comparison of various classifications of drug TdP propensity. *Drug Discovery Today* 22:10–16. DOI: 10.1016/j.drudis.2016.09.027.

Woosley, RL, Romero, KA www.Crediblemeds.org.

Woosley RL., Romero K., Heise CW., Gallo T., Tate J., Woosley RD., Ward S. 2017. Adverse Drug Event Causality Analysis (ADECA): A Process for Evaluating Evidence and Assigning Drugs to Risk Categories for Sudden Death. *Drug Safety* 40:465–474. DOI: 10.1007/s40264-017-0519-0.

Zhou X., Bueno-Orovio A., Orini M., Hanson B., Hayward MP., Taggart P., Lambiase PD., Burrage K., Rodriguez B. 2015. In Vivo and In Silico Investigation into Mechanisms of Frequency Dependence of Repolarization Alternans in Human Ventricular Cardiomyocytes. *Circulation Research*:CIRCRESAHA.115.307836. DOI: 10.1161/CIRCRESAHA.115.307836.

Figure Legends

Figure 1: Stacked bar-chart shows the proportion of compounds in each data-set that are TdeP positive (KR/PR on CredMeds database) or TdeP negative (CR on CredMeds database or not listed, NR).

Figure 2: Boxplots show the distribution of block for each ionic current across all 3 data-sets.

Figure 3: Bar-chart showing the performance of each model across all data-sets studied. The number in parentheses for the Crumb data-set refers to the number of ion-channels used 3 v 7.

Table Titles

Table 1: Results of the leave one out cross validation results across all data-sets for all models considered.

Figure 1(on next page)

Figure 1

Stacked bar-chart shows the proportion of compounds in each data-set that are TdeP positive (KR/PR on CredMeds database) or TdeP negative (CR on CredMeds database or not listed, NR).


No. Compounds
40
20
0
Crumb et al. (2016)
Kramer et al. (2013)
Mirams et al. (2011)
Study
CredMeds
CR/NR
KR/PR

Figure 2(on next page)

Figure 2

Boxplots show the distribution of block for each ionic current across all 3 data-sets.

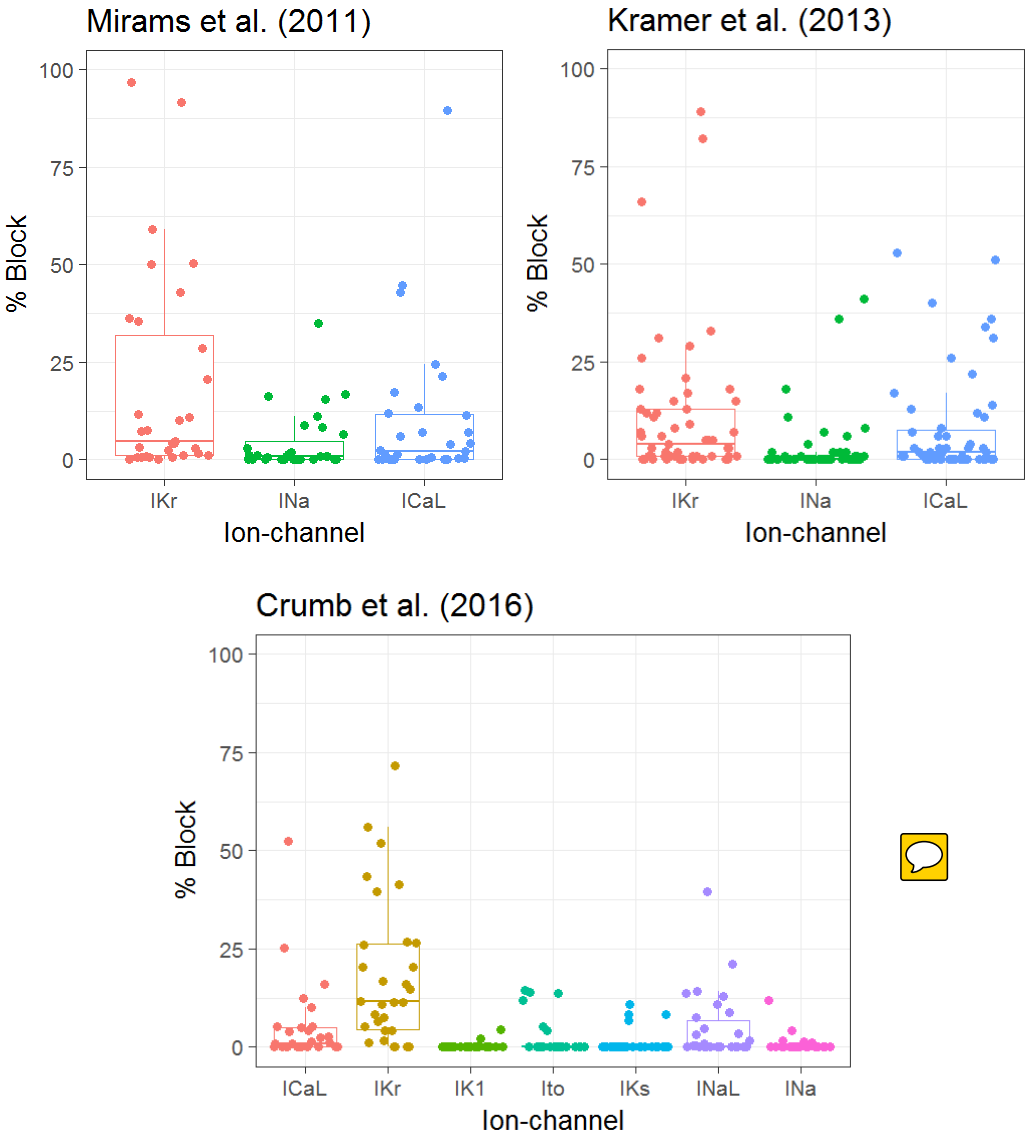
Mirams et al. (2011)
% Block
100
75
50
25
0
IKr
INa
ICaL
Ion-channel
Kramer et al. (2013)
% Block
100
75
50
25
0
IKr
INa
ICaL
Ion-channel
Crumb et al. (2016)
% Block
100
75
50
25
0
ICaL
IKr
IK1
Ito
IKs
INaL
INa
Ion-channel

Figure 3(on next page)

Figure 3

Bar-chart showing the performance of each model across all data-sets studied. The number in parentheses for the Crumb data-set refers to the number of ion-channels used 3 v 7.

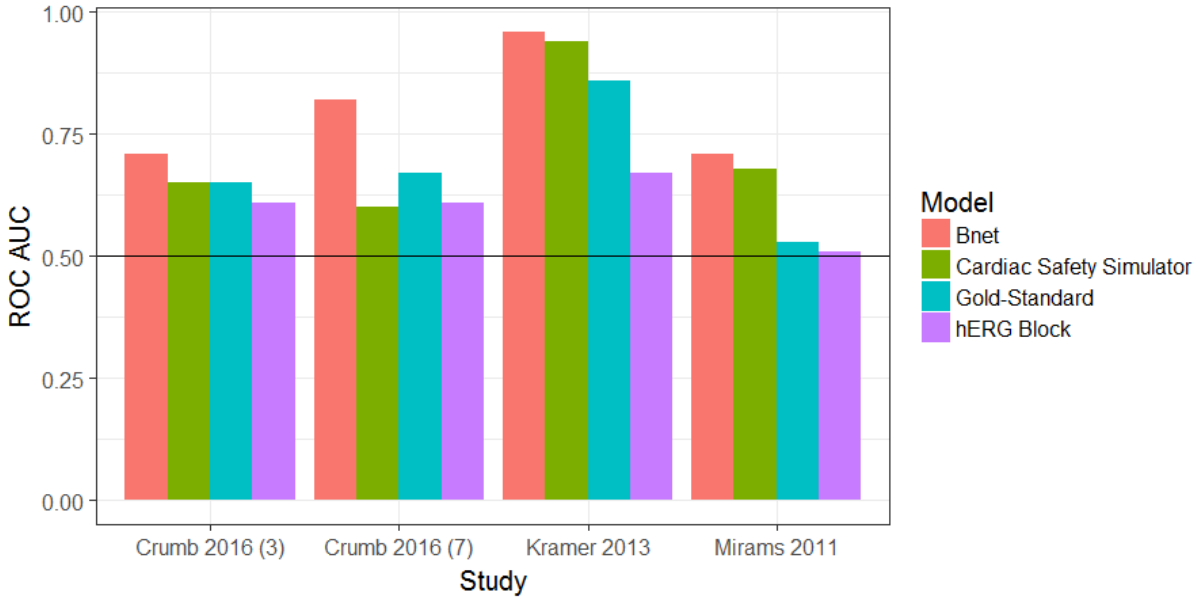
1.00
0.75
0.50
0.25
0.00
ROC AUC
Crumb 2016 (3)
Crumb 2016 (7)
Kramer 2013
Mirams 2011
Study
Model
Bnet
Cardiac Safety Simulator
Gold-Standard
hERG Block

Table 1(on next page)

Table 1

Results of the leave one out cross validation results across all data-sets for all models considered.

Table 1: Results of the leave one out cross validation results across all data-sets for all models considered.

Data-Set	Leave One Out Cross Validation ROC AUC			
		3 ion-channels		
	B_{net}	Gold-Standard: ΔAPD90	Cardiac Safety Simulator: ΔAPD90	hERG % Block IKr
Mirams (2011)	0.81	0.53	0.68	0.51
Kramer (2013)	0.96	0.86	0.94	0.67
Crumb (2016)	0.71	0.65	0.65	0.61
		7 ion-channels		
Crumb (2016)	0.82	0.67	0.60*	

*based on 6 ion-channels – INaL not modelled by TenTusscher et al.; ΔAPD90: percentage change in APD90