Journal of Pharmacological and Toxicological Methods 69 (2014) 115–140

Contents lists available at ScienceDirect

Journal of Pharmacological and Toxicological Methods

journal homepage: www.elsevier.com/locate/jpharmtox

Review

Progress in computational toxicology

Sean Ekins *

Collaborations in Chemistry, 5616 Hilltop Needmore Road, Fuquay Varina, NC 27526, USA
Department of Pharmaceutical Sciences, University of Maryland, 20 Penn Street, Baltimore, MD 21201, USA
Department of Pharmacology, Rutgers University-Robert Wood Johnson Medical School, 675 Hoes Lane, Piscataway, NJ 08854, USA
Division of Chemical Biology and Medicinal Chemistry, UNC Eshelman School of Pharmacy, University of North Carolina at Chapel Hill, NC 27599-7355, USA

ARTICLE INFO

Article history:
Received 8 October 2013
Accepted 8 December 2013

Keywords:
Bayesian
Computational toxicology
Machine learning
Support Vector Machine

ABSTRACT

Introduction: Computational methods have been widely applied to toxicology across pharmaceutical, consumer product and environmental fields over the past decade. Progress in computational toxicology is now reviewed. **Methods**: A literature review was performed on computational models for hepatotoxicity (e.g. for drug-induced liver injury (DILI)), cardiotoxicity, renal toxicity and genotoxicity. In addition various publications have been highlighted that use machine learning methods. Several computational toxicology model datasets from past publications were used to compare Bayesian and Support Vector Machine (SVM) learning methods. **Results**: The increasing amounts of data for defined toxicology endpoints have enabled machine learning models that have been increasingly used for predictions. It is shown that across many different models Bayesian and SVM perform similarly based on cross validation data. **Discussion**: Considerable progress has been made in computational toxicology in a decade in both model development and availability of larger scale or 'big data' models. The future efforts in toxicology data generation will likely provide us with hundreds of thousands of compounds that are readily accessible for machine learning models. These models will cover relevant chemistry space for pharmaceutical, consumer product and environmental applications.

Contents

1. Introduction . . . 116
2. Simple models . . . 116
3. Hepatotoxicity . . . 116
   3.1. Drug induced liver injury models . . . 116
   3.2. Time dependent inhibition of cytochrome P450's . . . 117
   3.3. Transporters . . . 118
   3.4. Nuclear hormone receptor—pregnane X receptor . . . 119
4. Cardiac toxicity . . . 122
   4.1. Recent hERG studies . . . 123
5. Renal toxicity . . . 126
6. Genotoxicity . . . 127
7. Computational tools . . . 131
   7.1. Metabolism . . . 131
   7.2. Open source or free ADME/Tox methods and models . . . 131
   7.3. Bioclipse . . . 132
   7.4. OECD Toolbox . . . 133
   7.5. admetSAR . . . 133
   7.6. ADME/Tox models developed with different machine learning methods . . . 133

Abbreviations: ANN, artificial neural networks; BSEP, bile salt export pump; DILI, drug-induced liver injury; ECFC_6, extended connectivity functional class fingerprints of maximum diameter 6; hERG, human ether-a-go-go-related gene; HIAT, hepatocyte imaging assay technology; HLAED, Human Liver Adverse Effects Database; kNN, k-Nearest Neighbor; LDA, linear discriminant analysis; MIC, metabolic intermediate complex; MOE, Molecular Operating Environment; NTCP, Na(+)-dependent taurocholate cotransporting polypeptide; PXR, pregnane X receptor; PDB, Protein Data Bank; QSAR, quantitative structure activity relationship; RF, random forest; ROC, Receiver Operator Characteristic; RP, recursive partitioning; RTECS, Registry of Toxic Effects of Chemical Substances; SVM, Support Vector Machine; TDI, time dependent inhibition.

* 5616 Hilltop Needmore Road, Fuquay Varina, NC 27526, USA.
E-mail address: ekinssean@yahoo.com.

1056-8719/$ – see front matter © 2013 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.vascn.2013.12.003

8. Discussion . . . 133
Acknowledgments . . . 137
Appendix A. Supplementary data . . . 137
References . . . 137

1. Introduction

Computational toxicology aims to use rules, models and algorithms based on prior data for specific endpoints, to enable the prediction of whether a new molecule will possess similar liabilities or not. In some cases the models are derived from discrete molecular endpoints while in others they are quite broad in scope. Previous reviews have described in some detail the computational toxicology models for the human ether-a-go-go-related gene (hERG) potassium channel, cytochrome P450, and pregnane X receptor (PXR) as examples (Ekins, 2007b; Ekins & Williams, 2012; Ekins et al., 2012; Kortagere, Krasowski, & Ekins, 2012; Williams, Ekins, Spjuth, & Willighagen, 2013). While there are several books devoted to computational toxicology alone in which the reader can find more detail on algorithms and approaches developed or applied (Cronin & Livingstone, 2004; Ekins, 2007b; Helma, 2005). In the current review the aim is to provide examples of commonly encountered toxicity issues and the current status of models for their detection. This review will describe models for hepatotoxicity (e.g. for drug-induced liver injury (DILI)), cardiotoxicity, renal toxicity and genotoxicity and will describe recent literature on these topics. Models developed with publically available or published data will be used to illustrate how various computational methods compare. This review excludes areas like reactive metabolites, idiosyncratic toxicity, phototoxicity, and ocular toxicity (Solimeo, Zhang, Kim, Sedykh, & Zhu, 2012) which are likely no less important. The focus will also be predominantly on ligand-based approaches (e.g. pharmacophores, machine learning, and quantitative structure activity relationships). The goal is to show how these models are used in industry (pharmaceutical, consumer products and regulatory) and academia and their impact as well as recommendations for the future (Ekins & Williams, 2012).

2. Simple models

Since the Rule of Five described orally active compounds in terms of a few simple molecular properties (Lipinski, Lombardo, Dominy, & Feeney, 1997) there has been some development of other types of rules. For example, Pfizer scientists determined that with 245 preclinical compounds with ClogP <3 (a measure of hydrophobicity) and total polar surface area >75A^2 (surface sum of nitrogens, oxygens and hydrogens) there were fewer toxicity findings (Hughes et al., 2008). Many pharmaceutical companies have developed computational filters to remove reactive molecules from their screening datasets (Hann et al., 1999; Pearce, Sofia, Good, Drexler, & Stock, 2006; Walters & Murcko, 2002). Abbott developed an assay to detect thiol reactive molecules by NMR (ALARM NMR) (Huth et al., 2005, 2007) and this data was in turn used to create a Bayesian classifier model to predict reactivity (Metz, Huth, & Hajduk, 2007). It has also recently been suggested that molecules failing such reactivity filters may correlate (Ekins & Freundlich, 2011) with the number of violations of the Rule of Five (Lipinski et al., 1997). So this relationship appears to be cyclic, reactivity and toxicity may correlate with such rule violations. At the molecular level for a specific toxicity, there is however rarely a reliance on one particular descriptor. This therefore requires more sophisticated efforts to find a relationship between molecule descriptors and toxicity endpoint, using some of the methods that will be described in the following sections. Simple rules may give you a high level approximation of toxicity, but the deeper you go into a target or mechanism the murkier it gets. The amount of biological data being created due to high throughput screening and by compilation of published data creates a wealth of information for assessing relationships between physicochemical properties and compounds (Gleeson, Hersey, Montanari, & Overington, 2011) and for developing computational models although we must not lose sight of the importance of data quality (Ekins, Olechno, & Williams, 2013; Fourches, Muratov, & Tropsha, 2010; Williams & Ekins, 2011; Williams, Ekins, & Tkachenko, 2012). Using literature data compilations it appears that safety profile is inversely correlated with target potency (Gleeson et al., 2011). Pfizer used data for over 100,000 compounds extracted from the literature and measured in their own laboratories against many different assays. They used these data to develop a Bayesian model for predicting cytotoxicity (Langdon, Mulgrew, Paolini, & van Hoorn, 2010) with training Receiver Operator Characteristic (ROC) = 0.84. Other generic models of human toxicities have also been published.

3. Hepatotoxicity

There appear to be many factors that may come into play when describing hepatotoxicity and some of these have been computationally modeled to different extents. For example, computational quantitative structure activity relationship (QSAR) or machine learning methods have been used for predicting hepatotoxicity (Cheng & Dixon, 2003; Clark, Wolohan, Hodgkin, Kelly, & Sussman, 2004) or drug–drug interactions (Ekins et al., 2000; Marechal et al., 2006; Ung, Li, Yap, & Chen, 2007; Zientek et al., 2010). Drug metabolism in the liver can convert some drugs into highly reactive intermediates (Boelsterli, Ho, Zhou, & Leow, 2006; Kassahun et al., 2001; Park, Kitteringham, Maggs, Pirmohamed, & Williams, 2005; Walgren, Mitchell, & Thompson, 2005) and this may lead to DILI which is the major reason why drugs are not approved or are withdrawn from the market post-approval (Schuster, Laggner, & Langer, 2005). There are some relatively high level hepatotoxicity models like DILI which may operate via many mechanisms and there are also receptors and transporters that have been individually implicated in hepatotoxicity. Several studies make use of collated sets of compounds (like the Registry of Toxic Effects of Chemical Substances (RTECS) database) in order to build models. For example a group at the NIH compared Naïve Bayesian, sequential minimal optimization and weighted feature significance algorithms to predict hepatotoxicity and found the latter method performed the best with a test set (ROC = 0.67) (Huang et al., 2009). Another data source is the FDA's Human Liver Adverse Effects Database (HLAED) which has been used to build k-Nearest Neighbor (kNN) models which appeared to have reasonable predictivity when used to screen several external databases (Rodgers, Zhu, Fourches, Rusyn, & Tropsha, 2010).

3.1. Drug induced liver injury models

An early DILI study found that the concordance of an in vitro human hepatocyte imaging assay technology (HIAT) applied to about 300 drugs and chemicals is about 75% with regard to clinical hepatotoxicity, with very few false-positives (Xu et al., 2008). Human clinical DILI data can be used to create a computational model to be used as a prescreen before in vitro testing. A computational study used classification models based on linear discriminant analysis (LDA), and machine learning algorithms (OneR) with 74 molecules (Cruz-Monteagudo, Cordeiro, & Borges, 2007). These models were then tested with very small numbers