Biochemical and Biophysical Research Communications 457 (2015) 249–255

Contents lists available at ScienceDirect

Biochemical and Biophysical Research Communications

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

Systematic drug safety evaluation based on public genomic expression (Connectivity Map) data: Myocardial and infectious adverse reactions as application cases

Kejian Wang [a, *], Zuquan Weng [b], Liya Sun [a], Jiazhi Sun [c], Shu-Feng Zhou [c], Lin He [a, *]

[a] Bio-X Institutes, Key Laboratory for the Genetics of Developmental and Neuropsychiatric Disorders, Shanghai Jiao Tong University, Shanghai, China
[b] Japan National Institute of Occupational Safety and Health, Kawasaki, Japan
[c] Department of Pharmaceutical Sciences, College of Pharmacy, University of South Florida, Tampa, FL, USA

ARTICLE INFO

Article history:
Received 4 December 2014
Available online 6 January 2015

Keywords:
Adverse drug reaction
FDA boxed warning
Black box warning
Connectivity Map
Risk score model

ABSTRACT

Adverse drug reaction (ADR) is of great importance to both regulatory agencies and the pharmaceutical industry. Various techniques, such as quantitative structure–activity relationship (QSAR) and animal toxicology, are widely used to identify potential risks during the preclinical stage of drug development. Despite these efforts, drugs with safety liabilities can still pass through safety checkpoints and enter the market. This situation raises the concern that conventional chemical structure analysis and phenotypic screening are not sufficient to avoid all clinical adverse events. Genomic expression data following *in vitro* drug treatments characterize drug actions and thus have become widely used in drug repositioning. In the present study, we explored prediction of ADRs based on the drug-induced gene-expression profiles from cultured human cells in the Connectivity Map (CMap) database. The results showed that drugs inducing comparable ADRs generally lead to similar CMap expression profiles. Based on such ADR-gene expression association, we established prediction models for various ADRs, including severe myocardial and infectious events. Drugs with FDA boxed warnings of safety liability were effectively identified. We therefore suggest that drug-induced gene expression change, in combination with effective computational methods, may provide a new dimension of information to facilitate systematic drug safety evaluation.

1. Introduction

Adverse drug reactions (ADRs), also referred to as side effects, are a major public health concern. The number of deaths caused by ADRs is estimated to be comparable to that from common diseases, such as cancers and cardiovascular disorders [1]. ADRs also lead to failures of drug development and drug withdrawals from the market, thus elevating the overall cost of drug development and reducing the number of newly approved drugs [2].

Many safety assessment methods, such as quantitative structure–activity relationship (QSAR) modeling [3], preclinical animal tests, and phase 1 clinical studies in humans, are developed to set checkpoints at early stages of the drug development pipeline. However, some drugs with safety risks still enter the market, resulting in severe injuries and even loss of lives [4]. This situation suggests that conventional methods based on chemical structure [5,6] and human/animal phenotypes are necessary but not sufficient to detect all the various kinds of human safety risks [7], especially some latent reactions that are detectable only after long-term clinical use [8]. A series of novel computational models have therefore been established as alternative solutions to early prediction of ADRs. These models are based on various types of large-scale information, such as pharmacological networks [9,10], biochemical assays [11] and systematic chemical characteristics [12–14], but relatively less progress has been made regarding high throughput gene-expression data [15].

CMap is a transcriptomic data collection developed from 6100 human cell cultures treated with 1309 bioactive compounds and matched vehicle controls. For each compound, the expression changes of 22283 gene probes collectively constitute an expression profile. For each profile, a small set of gene probes with the highest fold change compared to vehicle controls were used as drug-specific signatures [16]. One potential limitation of the use of this vast

* Corresponding authors.
E-mail addresses: kejian.wang.bio@gmail.com (K. Wang), helin@Bio-X.com (L. He).

http://dx.doi.org/10.1016/j.bbrc.2014.12.096
0006-291X/

transcriptomic dataset is the so-called batch effect, in which gene expression changes due to a drug may be confounded by cell culture conditions and other uncontrolled variables [17,18]. Since cell cultures treated with the same compound but in different batches should differ more in batch variation and less in drug action, we used the expression profiles of these cultures to estimate and normalize batch variation. The signatures of normalized expression profiles were compared to each other using the Gene Set Enrichment Analysis (GSEA) algorithm [19], so the similarity between different drugs could be measured by Bridge Adjusted Expression Similarity (BAES) [20].

Since high similarity scores may reveal common actions of different drugs, CMap has been widely used to predict new drug indications (i.e., drug repositioning) [21–24]. But much less attention being given to the potential association between CMap and drug side effects [25]. Here we found drug-specific gene expression profiles as a guilt-by-association indicator of ADRs and established a 'risk score' model to computationally predict drug safety risks. The results showed the unique advantages of this CMap-based model, which included the prediction of ADRs that are difficult to detect with traditional methods.

2. Materials and methods

2.1. Selection of reference drugs according to boxed warning information

The CMap drugs subjected to records in DailyMed up to March 2014 were manually collected from the official website at http://dailymed.nlm.nih.gov/. And the box warning information was independently validated by 2 experts to determine the list of reference drugs. To avoid subjective bias of selection and keep only the most straightforward information, only the boxed warnings suggesting direct associations between one drug and one side effect were used to select reference drugs. Other warning information was strictly screened out (Table S1), such as concomitant administration of multiple drugs, exacerbation of preexisted conditions, consequence of discontinuance, etc. The ADRs with similar symptoms or correlated toxicology were pooled together as an ADR class (Table S2). For instance, 'acute liver failure' and 'hepatic necrosis' were both categorized as drug-induced liver injury (DILI). A total of 25 ADR classes were associated to at least one CMap drug (Data S1). To ensure statistical power, only those ADR classes linked to at least 10 drugs were addressed with risk score model.

2.2. The assessment model of risk score

The risk score of a certain drug was measured by its overall gene-expression similarity to a series of reference drugs with boxed warnings. If N ($N > 0$) reference drugs are used in the model, then the risk score of a test drug is estimated with this formula

$$Risk = \frac{\sum_{i=1}^{N} S_i}{N},$$

where S_i is the expression similarity, in terms of BAES score, between the test drug and the i-th reference drug.

3. Results

3.1. Drug side effects are characterized by genomic expression profiles

We primarily retrieved drug label information from the DailyMed database (http://dailymed.nlm.nih.gov/) [26]. The highest level of safety risks or adverse effects are usually suggested by the boxed warning (also referred to as 'black box warning'), which is the strongest warning that the FDA requires on drug labels. On the other hand, the risks of less concern are mostly indicated in the 'warnings and precautions' section. In the present study, all CMap drugs were queried in DailyMed for boxed warning information. A total of 293 drug-ADR relationships between all CMap drugs and 25 major ADR classes were identified as 'gold standard' (Data S1). Our basic hypothesis is that drugs showing high expression similarity were prone to induce similar side effects. Therefore, the drugs warned for direct association with specific clinical adverse events were used as 'reference drugs' (see Methods). Then the risk of a new test drug was estimated according to its expression similarity to reference drugs, i.e. the higher the similarity, the higher the likelihood of safety liability it may suggest (Fig. S1).

To verify our hypothesis, we primarily confirmed that with reduced batch variation in CMap data, drugs warned for the same class of ADRs showed generally higher expression similarity (in terms of BAES score) than random drugs (Fig. 1A–D). Moreover, we noticed that drugs associated to the same ADR class were enriched in the drug pairs with the highest similarity. The ADR-expression association suggested that not only therapeutic effects, but also side effects can be characterized by transcriptomic profiles. In particular, a variety of severe ADRs affecting different organs and tissues were highly associated with characteristic expression patterns (Fig. 1E). As a result, we developed computational models to detect potential drug safety risks, including drug-induced myocardial reactions and serious infections.

3.2. Case study: drug-induced myocardial adverse reactions

Myocardial toxicity is one type of life-threatening adverse reaction, and also one of the most common causes of drug withdrawals from worldwide markets [27]. Because of the unexpected myocardial reactions associated with the wide-scale use of Vioxx® (generic name rofecoxib) [28], damage to the heart muscle (e.g., myocardial infarction and heart failure) has become a major focus of drug safety. However, the toxicological mechanisms of myocardial adverse reactions remain poorly understood, which makes prospective safety assessment difficult.

Here we retrieved from DailyMed a set of reference drugs associated with adverse myocardial reactions, such as myocardial infarction, heart failure and congestive heart failure etc. Since drugs with comparable ADRs tend to produce correlated expression profiles in CMap, the likelihood of myocardial ADRs of a specific drug may be quantitatively characterized by its overall expression similarity (i.e., BAES score) to those of reference drugs warned for myocardial risks; the higher the similarity, the greater the odds of that specific drug causing ADRs (Fig. 2A). To verify this theory, we adopted a previously published naïve model [20] that equally weights expression similarity to individual reference drugs and translates ADR risk into a 'risk score' (see Methods).

The performance of risk score model was evaluated by leave-one-out cross validation (LOOCV) and visualized by receiver operating characteristic (ROC) curve. With reference and non-reference drugs all ranked in the order of risk score value, the efficiency of detecting reference drugs was quantified by the indexes of sensitivity and specificity (Fig. 2B). Setting the 98% quantile of risk score of non-reference drugs as the threshold (i.e., under the premise of 98% classification specificity), we identified 25% reference drugs above the threshold. That means drugs with boxed warning are 12.5 times more likely to be found among those with the highest risk scores than random inspection (Fisher's exact $p = 5.9 \times 10^{-7}$), suggesting the solid characteristics of warned drugs.

Given the guidance of risk score, we paid attention to not only the reference drugs but also the non-reference ones subjected to

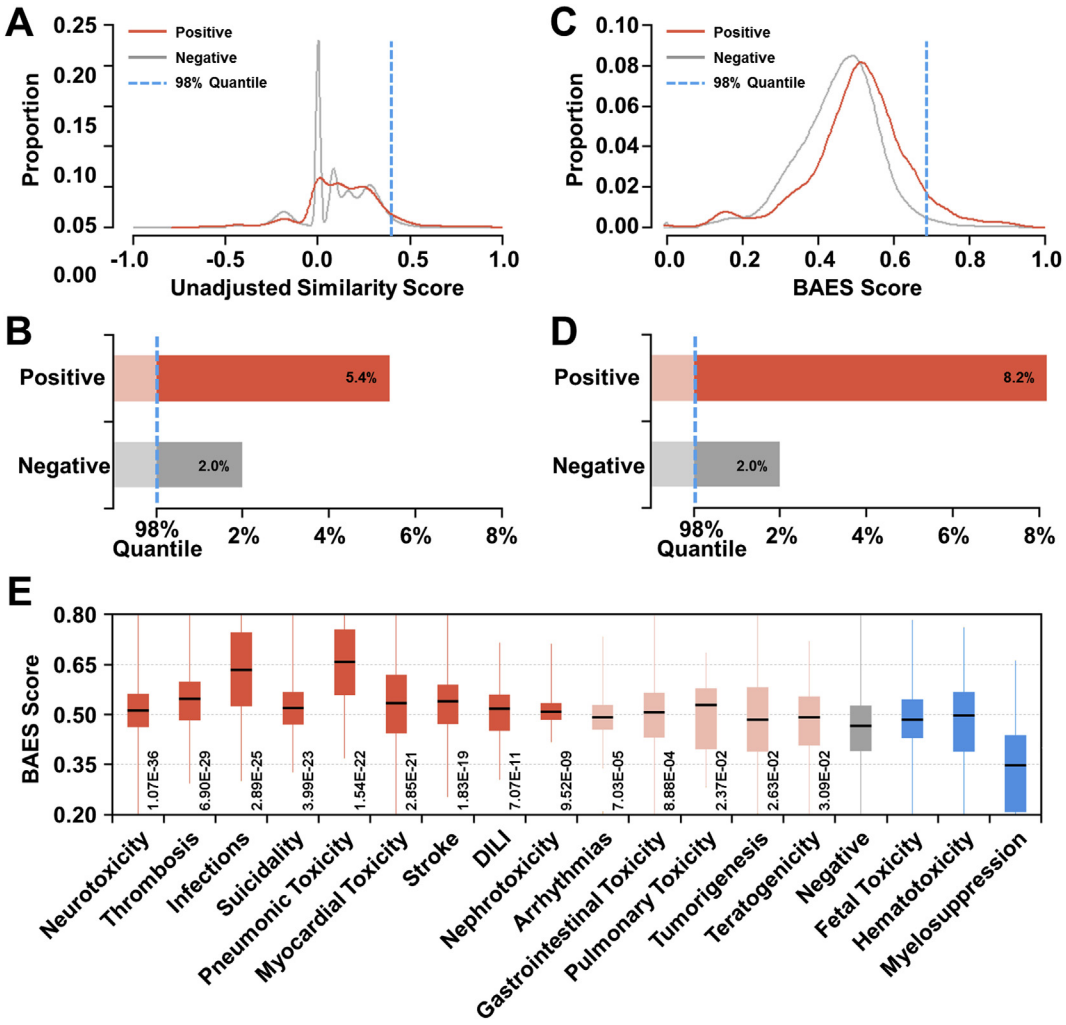

Fig. 1. Based on batch effect adjustment, drug pairs with common ADRs (i.e., positive pairs) exhibit higher expression similarity than other drug pairs (i.e., negative pairs). (A–B) In unadjusted CMap data, the difference of expression similarity between positive and negative pairs is relatively less clear. Only about 5.4% of positive pairs exceed the 98% BAES score quantile of negative pairs (indicated by blue dotted line). (C–D) After batch effect adjustment, the positive and negative pairs are better separated. About 8.2% positive pairs have higher BAES score than the 98% quantile of negative pairs, suggesting a significant improvement (Fisher's exact test $p = 0.0016$) (E) The box plot depicts the distribution (i.e., median and quartiles) of BAES scores. In comparison to negative pairs (gray box), drugs associated with ADRs mostly exhibit high (red boxes, with t-test p-value lower than 10^{-5}) or detectable (pink boxes, p-value lower than 0.05) expression similarity. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

high score. Another drug with boxed warning, the sex hormone estradiol (http://dailymed.nlm.nih.gov/dailymed/lookup.cfm?setid=0f943869-41a2-4f7e-943d-dd8c17996dd8#nlm34066-1), was surprisingly highlighted with the 5th highest risk score among all CMap drugs (Fig. 2C). Increased risk of myocardial infarction is attributed to concomitant administration of estradiol and other drugs, rather than estradiol alone. According to our strict criteria of selection (Table S1), therefore, estradiol was not used as a reference drug despite the boxed warning. However, due to the commonality of adverse reactions, estradiol still disturbs the same genes as other warned drugs do and produces a highly similar expression profile. This case study provided a typical example of risk score application. Following the same procedure, the drug candidates leading to high risk score (especially when higher than most reference drugs) can be screened out for further safety assessment.

3.3. Case study: drug-induced serious infections

As a type of serious adverse reaction, drug-induced infections have been increasingly concerned by clinical and pharmacological researchers [29–31]. Since the pathophysiological mechanisms of drug-induced infections are mostly unclear [32], early warning of drug infectious risk remains technically difficult. By using CMap drugs with boxed warning of infections as references, we built a prediction model similar to the one mentioned above concerning myocardial risks. Compared with non-reference, the reference drugs were 28-fold enriched among the high risk score drugs in LOOCV (Fig. 3A), indicating the strong association between infections incidence and genomic expression pattern (Fisher's exact test $p = 2.4 \times 10^{-11}$).

Since CMap drugs warned for infections are all included in reference drugs (unlike estradiol in the case study on myocardial events), we further scrutinize the 'warnings and precautions' in labels of the non-reference drugs above threshold (Fig. 3B). Ranked 7th in all CMap drugs, the anticonvulsant drug valproic acid has long been reported to stimulate the replication of virus [33–36], which has been addressed in the drug label (http://dailymed.nlm.nih.gov/dailymed/lookup.cfm?setid=6b4331f5-4475-417a-6a9d-09c2f8334235#nlm34066-1). And another high score drug, immunosuppressant azathioprine, also has precautions on serious

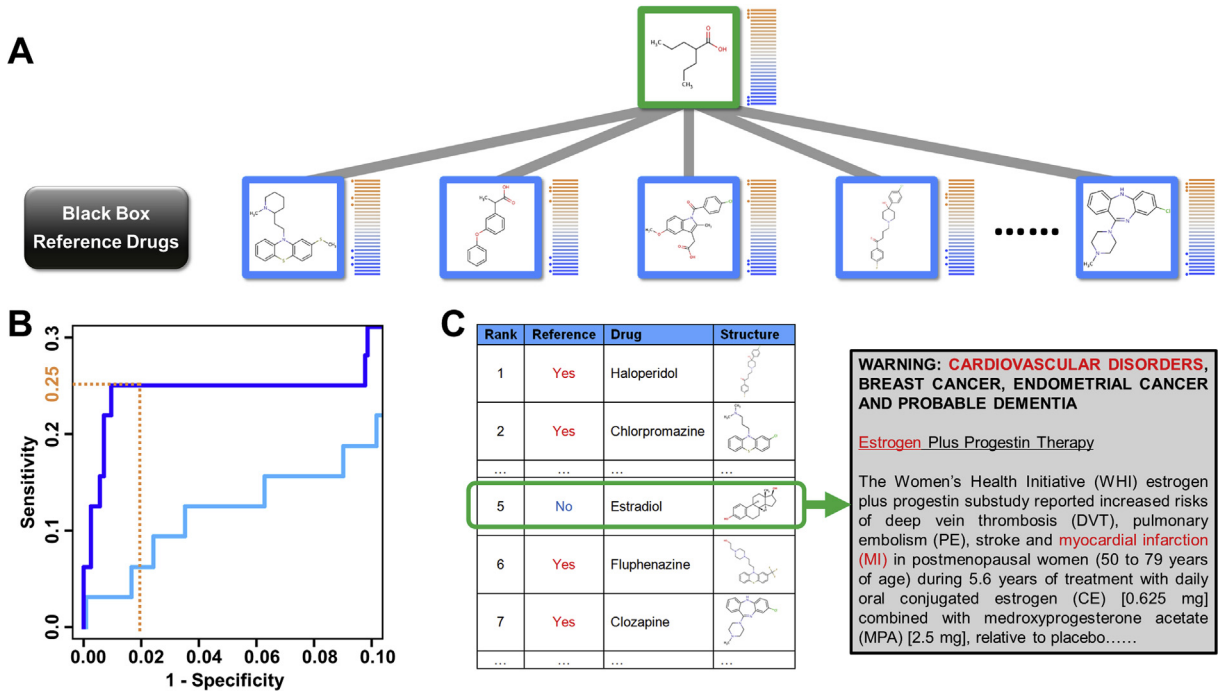

Fig. 2. The rationale and performance of the risk score model for myocardial risks. (A) Since side effect risks can be reflected by the genomic expression pattern, we hypothesize that drugs with potential myocardial reactions should particularly exhibit high expression similarity to reference drugs with boxed warning. (B) The risk score allowing 2% of the non-reference drugs is set as a threshold, which harvests 25% of the reference drugs. Blue and cyan lines correspond to models based on CMap data adjusted and unadjusted for batch effect, respectively. (C) Even though not one of the reference drugs, estradiol with boxed warning of myocardial risk is still found on the top of high risk score drugs. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

infections in drug label (http://dailymed.nlm.nih.gov/dailymed/lookup.cfm?setid=9050af7a-19c6-4670-937a-9445605de995), since it is suggested to induce infectious ADRs in patients by affecting immune system [37,38].

The effective detection of infectious risks again testified the ADR-expression association. Even though the relationship between ADR incidence (e.g. infections) and drug actions are not fully understood by human expert, it may still be biologically characterized by the genomic response of drug-treated cell cultures and computationally detected in early stage of new drug development.

3.4. Applicability to various types of ADRs

Following the same procedure of above case studies, we further assessed the performance of risk score model regarding other ADRs with characteristic expression patterns (Fig. 4). For example, drugs with tumorigenic concerns are highly distinguished from other drugs (Fig. 4J), with a 10-fold enrichment observed (i.e., 20% of reference drugs above threshold). For some other ADRs, such as stroke (Fig. 4G) and thrombosis (Fig. 4I), remarkable reference enrichment (over 3-fold) can still be achieved, if threshold is relaxed to 90% specificity. There are also some ADRs, such as nephrotoxicity (Fig. 4D) and neurotoxicity (Fig. 4E), exhibiting low efficiency of prediction. The reason may be that these ADRs occur through a variety of mechanisms, leading to diversified rather than consistent expression patterns. In this case, the consensus characteristics would hardly be found among reference drugs. Or probably, the genes associated with these ADRs are less active in the cell lines used in CMap. No matter which reason is true, some ADRs seemed to be less addressed than others by the risk score model.

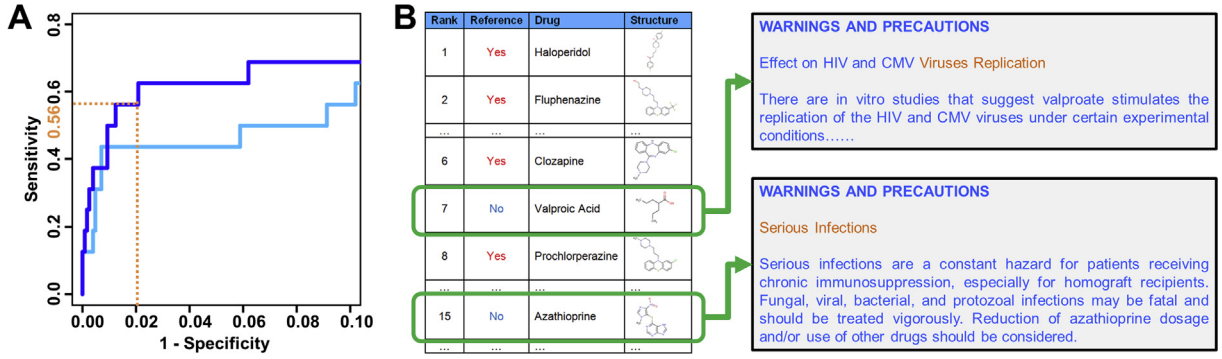

Fig. 3. The risk score model for drug-induced infections. (A) The risk score allowing 2% of the non-reference drugs is set as a threshold. And 56% of reference drugs are found above the threshold, leading to a significant enrichment. Blue and cyan lines correspond to CMap data adjusted and unadjusted for batch effect, respectively. (B) Some non-reference drugs associated with infectious concerns are further highlighted due to high risk score. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

In a previous study about predicting drug–target interactions, we have realized that even though ligands binding with common protein target generally showed higher expression similarity, the efficiency of ligand prediction varied widely from one target to another [20]. Similarly, here we found that risk score is not a one-size-fits-all indicator but highly dependent on the transcriptomic nature of specific ADR (Fig. 4K). Therefore, we recommend assessing the overall strength of ADR-expression association, by LOOCV or other statistical tools, before applying risk score model into predicting specific ADR.

4. Discussion

Predicting severe ADRs is one of the major objectives of pharmacology studies. Although many potential safety issues are effectively addressed by preclinical studies and clinical trials, ADRs still occur in marketed drugs. Since pharmacovigilance always lags behind the occurrence of adverse events, the health casualties and financial losses (including cost of drug development and compensation for injury) take place before safety warnings may be evident. Hence, there is an urgent need for effective methods to detect potential safety risks in advance. To meet this goal, numerous computational models have been established, according to the guilt-by-association principle, that correlate various drug 'features' (e.g., chemical structure, target binding, animal reaction, etc.) with specific ADRs [39,40]. On the other hand, the use of drug-induced *in vitro* response has not been adequately utilized. More often drug toxicology is studied with a handful of genes [41] than based on transcriptomic profile as a high dimensional feature.

CMap provided us with a unique opportunity to analyze a large number of drugs and genes at the same time. In the present study, we systematically explored ADR predictions based on the integration of CMap data and appropriate models. By controlling batch effects, the association between gene-expression pattern and ADR incidence was derived. This suggested the potential of applying transcriptomic data to safety evaluation. Especially in the early stage of drug development when biological characteristics (e.g., genomic expression) are extensively studied but clinical information is still limited, transcriptomic patterns may prospectively warn of potential risks in a cost-effective and rapid way. The success of applying CMap transcriptomic data to identify drugs with boxed warning implied the importance of utilizing other sources of

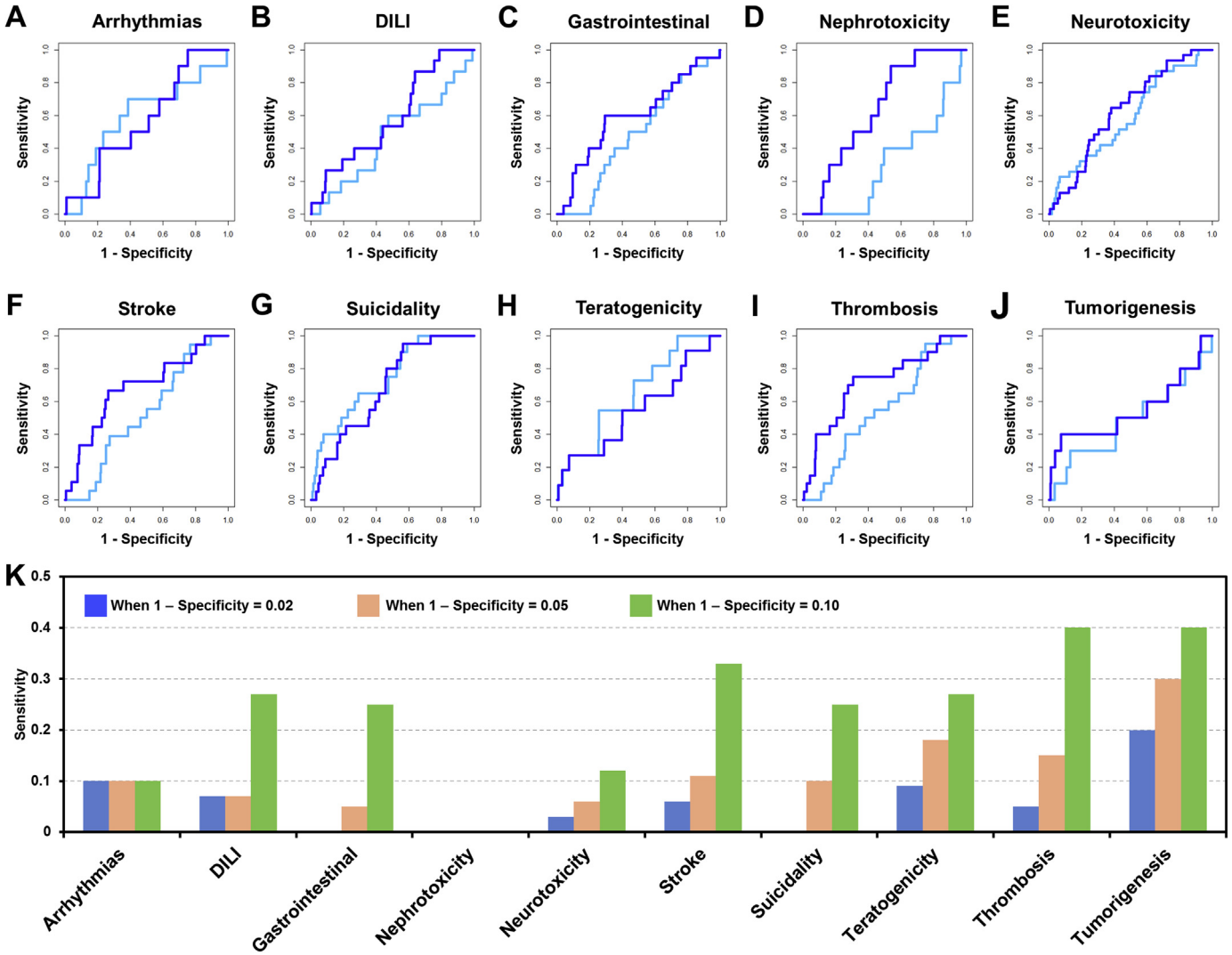

Fig. 4. The performance of risk score model regarding various ADRs (displayed in alphabetical order). (A–J) The ROC curves (blue and cyan lines stand for CMap data adjusted and unadjusted for batch effect, respectively) regarding individual ADRs. (K) The sensitivity of identifying reference drugs is presented given the thresholds of 98%, 95% and 90% classification specificity, respectively. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

genomic data such as those found in Gene Expression Omnibus (GEO) [42] and submitted to FDA by drug developers [43].

Given current results, there are several practical extensions of our work. First, not all drugs with boxed warning are enrolled in CMap, thus restricting the application range of risk score. Lactic Acidosis, for example, is a common ADR associated to quite a few drugs. However, only 2 of them are included in CMap, which makes risk score model inapplicable. Fortunately, the new generation of CMap (i.e., LINCS) is now in development [44], with much more drugs addressed than before. We expect the release of LINCS data in the near future to expand the boundary of ADRs prediction based on transcriptomic profiles. Second, the reference drugs are all equally considered in the current naïve model. However, these drugs may be associated with ADRs of different characteristics and severity [45]. Weighting the reference drugs accordingly may improve the ADR prediction, thus addressing not only boxed warnings, but also other ADR endpoints suggested by various sources of information. Finally, transcriptomic data should be combined with other types of information. Since ADR is a profound form of drug action, there may be many causes signaled by different high-dimensional information. We therefore expect the transcriptomic model described here to serve as not only a functional tool, but also a substantial part of an ADR prediction pipeline with even higher efficiency based on the integration of various data [46,47].

Conflict of interest

The authors declare no conflict of interest.

Acknowledgments

This work was supported by the National Natural Science Foundation of China, No. 81121001 (www.nsfc.gov.cn), the National Key Technology R&D Program, No. 2012BAI01B09 (program.most.gov.cn) and the 973 Program, No. 2010CB529600 (http://www.973.gov.cn/English/Index.aspx).

Transparency document

The transparency document associated with this article can be found in the online version at http://dx.doi.org/10.1016/j.bbrc.2014.12.096.

Appendix A. Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.bbrc.2014.12.096.

References

[1] K.M. Giacomini, R.M. Krauss, D.M. Roden, M. Eichelbaum, M.R. Hayden, When good drugs go bad, Nature 446 (2007) 975–977.
[2] M.L. Billingsley, Druggable targets and targeted drugs: enhancing the development of new therapeutics, Pharmacology 82 (2008) 239–244.
[3] D. Vidal, R. Garcia-Serna, J. Mestres, Ligand-based approaches to in silico pharmacology, Methods Mol. Biol. 672 (2011) 489–502.
[4] N.P. Tatonetti, P.P. Ye, R. Daneshjou, R.B. Altman, Data-driven prediction of drug effects and interactions, Sci. Transl. Med. 4 (2012) 125ra131.
[5] M.J. Keiser, B.L. Roth, B.N. Armbruster, P. Ernsberger, J.J. Irwin, B.K. Shoichet, Relating protein pharmacology by ligand chemistry, Nat. Biotechnol. 25 (2007) 197–206.
[6] E. Pauwels, V. Stoven, Y. Yamanishi, Predicting drug side-effect profiles: a chemical fragment-based approach, BMC Bioinform. 12 (2011) 169.
[7] A.A. Frid, E.J. Matthews, Prediction of drug-related cardiac adverse effects in humans-B: use of QS AR programs for early detection of drug-induced cardiac toxicities, Regul. Toxicol. Pharmacol. 56 (2010) 276–289.
[8] T.J. Moore, M.R. Cohen, C.D. Furberg, Serious adverse drug events reported to the Food and Drug Administration, 1998–2005, Arch. Intern. Med. 167 (2007) 1752–1759.
[9] A. Cami, A. Arnold, S. Manzi, B. Reis, Predicting adverse drug events using pharmacological network models, Sci. Transl. Med. 3 (2011) 114ra127.
[10] E.E. Schadt, S.H. Friend, D.A. Shaywitz, A network view of disease and compound screening, Nat. Rev. Drug Discov. 8 (2009) 286–295.
[11] Y. Pouliot, A.P. Chiang, A.J. Butte, Predicting adverse drug reactions using publicly available PubChem BioAssay data, Clin. Pharmacol. Ther. 90 (2011) 90–99.
[12] L. Yang, J. Chen, L. He, Harvesting candidate genes responsible for serious adverse drug reactions from a chemical-protein interactome, PLoS Comput. Biol. 5 (2009) e1000441.
[13] N.P. Tatonetti, T. Liu, R.B. Altman, Predicting drug side-effects by chemical systems biology, Genome Biol. 10 (2009) 238.
[14] L. Xie, J. Li, P.E. Bourne, Drug discovery using chemical systems biology: identification of the protein–ligand binding network to explain the side effects of CETP inhibitors, PLoS Comput. Biol. 5 (2009) e1000387.
[15] T.P. Reilly, M. Bourdi, J.N. Brady, C.A. Pise-Masison, M.F. Radonovich, J.W. George, L.R. Pohl, Expression profiling of acetaminophen liver toxicity in mice using microarray technology, Biochem. Biophys. Res. Commun. 282 (2001) 321–328.
[16] E.M. Airoldi, C. Huttenhower, D. Gresham, C. Lu, A.A. Caudy, M.J. Dunham, J.R. Broach, D. Botstein, O.G. Troyanskaya, Predicting cellular growth from gene expression signatures, PLoS Comput. Biol. 5 (2009) e1000257.
[17] M. Iskar, M. Campillos, M. Kuhn, L.J. Jensen, V. van Noort, P. Bork, Drug-induced regulation of target expression, PLoS Comput. Biol. 6 (2010).
[18] C. Chen, K. Grennan, J. Badner, D. Zhang, E. Gershon, L. Jin, C. Liu, Removing batch effects in analysis of expression microarray data: an evaluation of six batch adjustment methods, PLoS One 6 (2011) e17238.
[19] J. Lamb, E.D. Crawford, D. Peck, J.W. Modell, I.C. Blat, M.J. Wrobel, J. Lerner, J.P. Brunet, A. Subramanian, K.N. Ross, M. Reich, H. Hieronymus, G. Wei, S.A. Armstrong, S.J. Haggarty, P.A. Clemons, R. Wei, S.A. Carr, E.S. Lander, T.R. Golub, The Connectivity Map: using gene-expression signatures to connect small molecules, genes, and disease, Science 313 (2006) 1929–1935.
[20] K. Wang, J. Sun, S. Zhou, C. Wan, S. Qin, C. Li, L. He, L. Yang, Prediction of drug-target interactions for drug repositioning only based on genomic expression similarity, PLoS Comput. Biol. 9 (2013) e1003315.
[21] S.D. Kunkel, M. Suneja, S.M. Ebert, K.S. Bongers, D.K. Fox, S.E. Malmberg, F. Alipour, R.K. Shields, C.M. Adams, mRNA expression signatures of human skeletal muscle atrophy identify a natural compound that increases muscle mass, Cell. Metab. 13 (2011) 627–638.
[22] D.C. Hassane, M.L. Guzman, C. Corbett, X. Li, R. Abboud, F. Young, J.L. Liesveld, M. Carroll, C.T. Jordan, Discovery of agents that eradicate leukemia stem cells using an in silico screen of public gene expression data, Blood 111 (2008) 5654–5662.
[23] Y. Ishimatsu-Tsuji, T. Soma, J. Kishimoto, Identification of novel hair-growth inducers by means of connectivity mapping, FASEB J. 24 (2010) 1489–1496.
[24] F. Iorio, R. Bosotti, E. Scacheri, V. Belcastro, P. Mithbaokar, R. Ferriero, L. Murino, R. Tagliaferri, N. Brunetti-Pierri, A. Isacchi, D. di Bernardo, Discovery of drug mode of action and drug repositioning from transcriptional responses, Proc. Natl. Acad. Sci. U. S. A. 107 (2010) 14621–14626.
[25] S. Lee, K.H. Lee, M. Song, D. Lee, Building the process-drug-side effect network to discover the relationship between biological processes and side effects, BMC Bioinform. 12 (2011).
[26] J. de Leon, Highlights of drug package inserts and the website DailyMed: the need for further improvement in package inserts to help busy prescribers, J. Clin. Psychopharmacol. 31 (2011) 263–265.
[27] M. Fung, A. Thornton, K. Mybeck, J.H.-H. Wu, K. Hornbuckle, E. Muniz, Evaluation of the characteristics of safety withdrawal of prescription drugs from worldwide pharmaceutical markets-1960 to 1999, Drug Inf. J. 35 (2001) 293–317.
[28] S.E. Nissen, Adverse cardiovascular effects of rofecoxib, N. Engl. J. Med. 355 (2006) 203–204 author reply 203–205.
[29] A. Fisher, J.M. Seguel, A.N. de la Torre, D. Wilson, A. Merchant, R.K. Arora, B. Koneru, Effect of sirolimus on infection incidence in liver transplant recipients, Liver Transpl. 10 (2004) 193–198.
[30] J.S. Coblyn, Infections, drugs, and rheumatoid arthritis. What have we learned? J. Rheumatol. 35 (2008) 375–376.
[31] G. Trifiro, Antipsychotic drug use and community-acquired pneumonia, Curr. Infect. Dis. Rep. 13 (2011) 262–268.
[32] W. Knol, R.J. van Marum, P.A. Jansen, P.C. Souverein, A.F. Schobben, A.C. Egberts, Antipsychotic drug use and risk of pneumonia in elderly people, J. Am. Geriatr. Soc. 56 (2008) 661–666.
[33] G. Kuntz-Simon, G. Obert, Sodium valproate, an anticonvulsant drug, stimulates human cytomegalovirus replication, J. Gen. Virol. 76 (Pt 6) (1995) 1409–1415.
[34] C. Moog, G. Kuntz-Simon, C. Caussin-Schwemling, G. Obert, Sodium valproate, an anticonvulsant drug, stimulates human immunodeficiency virus type 1 replication independently of glutathione levels, J. Gen. Virol. 77 (Pt 9) (1996) 1993–1999.
[35] M. Michaelis, T. Suhan, A. Reinisch, A. Reisenauer, C. Fleckenstein, D. Eikel, H. Gumbel, H.W. Doerr, H. Nau, J. Cinatl Jr., Increased replication of human cytomegalovirus in retinal pigment epithelial cells by valproic acid depends

on histone deacetylase inhibition, Invest Ophthalmol. Vis. Sci. 46 (2005) 3451–3457.
[36] M. Michaelis, N. Kohler, A. Reinisch, D. Eikel, U. Gravemann, H.W. Doerr, H. Nau, J. Cinatl Jr., Increased human cytomegalovirus replication in fibroblasts after treatment with therapeutical plasma concentrations of valproic acid, Biochem. Pharmacol. 68 (2004) 531–538.
[37] A. Garrido Serrano, F. Perez Martin, F.J. Guerrero Igea, J. Galbarro Munoz, S. Palomo Gil, Fatal infectious mononucleosis during azathioprine treatment in Crohn's disease, Gastroenterol. Hepatol. 23 (2000) 7–8.
[38] D.I. Min, A.P. Monaco, Complications associated with immunosuppressive therapy and their management, Pharmacotherapy 11 (1991) 119S–125S.
[39] D.R. Abernethy, J.P. Bai, K. Burkhart, H.G. Xie, P. Zhichkin, Integration of diverse data sources for prediction of adverse drug events, Clin. Pharmacol. Ther. 90 (2011) 645–646.
[40] L. Yang, H. Luo, J. Chen, Q. Xing, L. He, SePreSA: a server for the prediction of populations susceptible to serious adverse drug reactions implementing the methodology of a chemical-protein interactome, Nucleic Acids Res. 37 (2009) W406–W412.
[41] R.P. Singh, K.V. Sastry, P.K. Dubey, R. Agrawal, R. Singh, N.K. Pandey, J. Mohan, Norfloxacin drug induces reproductive toxicity and alters androgen receptor gene expression in testes and cloacal gland of male Japanese quail (Coturnix Japonica), Environ. Toxicol. Chem. 32 (2013) 2134–2138.
[42] R. Edgar, M. Domrachev, A.E. Lash, Gene expression omnibus: NCBI gene expression and hybridization array data repository, Nucleic Acids Res. 30 (2002) 207–210.
[43] M.S. Orr, F. Goodsaid, S. Amur, A. Rudman, F.W. Frueh, The experience with voluntary genomic data submissions at the FDA and a vision for the future of the voluntary data submission program, Clin. Pharmacol. Ther. 81 (2007) 294–297.
[44] S. Santagata, M.L. Mendillo, Y.C. Tang, A. Subramanian, C.C. Perley, S.P. Roche, B. Wong, R. Narayan, H. Kwon, M. Koeva, A. Amon, T.R. Golub, J.A. Porco Jr., L. Whitesell, S. Lindquist, Tight coordination of protein translation and HSF1 activation supports the anabolic malignant state, Science 341 (2013) 1238303.
[45] N.P. Tatonetti, J.C. Denny, S.N. Murphy, G.H. Fernald, G. Krishnan, V. Castro, P. Yue, P.S. Tsao, I. Kohane, D.M. Roden, R.B. Altman, Detecting drug interactions from adverse-event reports: interaction between paroxetine and pravastatin increases blood glucose levels, Clin. Pharmacol. Ther. 90 (2011) 133–142.
[46] L.C. Huang, X. Wu, J.Y. Chen, Predicting adverse side effects of drugs, BMC Genomics 12 (Suppl. 5) (2011) S11.
[47] M. Liu, Y. Wu, Y. Chen, J. Sun, Z. Zhao, X.W. Chen, M.E. Matheny, H. Xu, Large-scale prediction of adverse drug reactions using chemical, biological, and phenotypic properties of drugs, J. Am. Med. Inf. Assoc. 19 (2012) e28–e35.