JOURNAL OF CHEMICAL INFORMATION AND MODELING

Article

pubs.acs.org/jcim

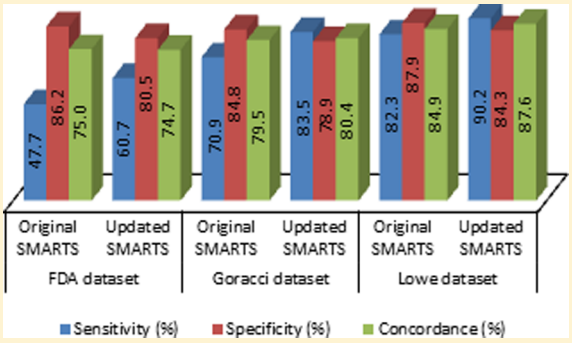
How Does the Quality of Phospholipidosis Data Influence the Predictivity of Structural Alerts?

Katarzyna R. Przybylak, Abdullah Rzgallah Alzahrani, and Mark T. D. Cronin*

School of Pharmacy and Chemistry, Liverpool John Moores University, Byrom Street, Liverpool, L3 3AF, England

Supporting Information

ABSTRACT: The ability of drugs to induce phospholipidosis (PLD) is linked directly to their molecular substructures: hydrophobic, cyclic moieties with hydrophilic, peripheral amine groups. These structural properties can be captured and coded into SMILES arbitrary target specification (SMARTS) patterns. Such structural alerts, which are capable of identifying potential PLD inducers, should ideally be developed on a relatively large but reliable data set. We had previously developed a model based on SMARTS patterns consisting of 32 structural fragments using information from 450 chemicals. In the present study, additional PLD structural alerts have been developed based on a newer and larger data set combining two data sets published recently by the United States Food and Drug Administration (US FDA). To assess the predictive performance of the updated SMARTS model, two publicly available data sets were considered. These data sets were constructed using different criteria and hence represent different standards for overall quality. In the first data set high quality was assured as all negative chemicals were confirmed by the gold standard method for the detection of PLD—transmission electron microscopy (EM). The second data set was constructed from seven previously published data sets and then curated by removing compounds where conflicting results were found for PLD activity. Evaluation of the updated SMARTS model showed a strong, positive correlation between predictive performance of the alerts and the quality of the data set used for the assessment. The results of this study confirm the importance of using high quality data for modeling and evaluation, especially in the case of PLD, where species, tissue, and dose dependence of results are additional confounding factors.



INTRODUCTION

Drug induced phospholipidosis (PLD) can lead to the impairment of lipid metabolism and the accumulation of phospholipids and drugs in cells, especially lysosomes.[1] Lysosomes have an important role in the degradation of lipids. Lipids, including phingolipids, cholesterolesters, and glycerophospholipids, arrive in lysosomes for degradation by phospholipases as side products of organelles' autophagy and endocytosis.[2] PLD is a very complex effect caused mostly by cationic amphiphilic drugs (CADs), which exhibit the property of lysosomotropism. CADs share similar structural properties—a hydrophobic structural domain consisting of an aromatic and/or aliphatic ring system and a hydrophilic cationic side chain with a primary, secondary, or tertiary amine.[3] Lysososmotropic compounds are weak bases, which, in unionized forms, are membrane permeable and partition across the lysosomal membrane based on concentration gradient. In the acidic environment of the lysosome, CADs are protonated and are trapped inside the organelle. As a consequence, the concentration of free base within the lysosome decreases gradually, which subsequently drives additional unionized drug to diffuse into the lysosome and then be protonated. This causes an increase in the pH of the lysosome toward neutrality, which is less favorable for lysosomal hydrolases.[4] There are several possible mechanisms by which CADs can induce PLD, including the inhibition of lysosomal phospholipases[5,6] and the formation of complexes between CADs and phospholipids destined for degradation.[7,8] Recent toxicogenomics studies by Sawada et al. proposed four possible mechanisms for PLD: (1) inhibition of lysosmal phospholipase activity; (2) inhibition of lysosmal enzyme transport; (3) enhanced phospholipid biosynthesis; and (4) enhanced cholesterol biosynthesis.[5]

The morphological hallmark of PLD is the presence of lysosomal lamellar bodies (LLB), which are secondary lysosomes filled with electron-dense, membranous, lamellar material.[9,10] PLD is not organ specific and can be induced in many tissues in the body such as the lung, liver, brain, kidney, ocular tissue, heart, adrenal glands, hematopoietic tissue and circulating lymphocytes.[10] No species, gender, or age group appears to be excluded from their susceptibility to the induction of phospholipidosis.[11] The sites of induction vary among, and within, species and cells/tissues and one drug may induce PLD in certain tissues in one species but not another. A recent study showed that dog hepatocytes are more sensitive to some lysososmotropic compounds than those from the rat.[12]

Received: April 15, 2014
Published: July 25, 2014

Moreover, PLD can be observed in vivo and also in cell culture. However, drugs can induce PLD in vivo, while the effect may not be demonstrated in vitro, or vice versa. For example, gentamicin, known to induce PLD in both animals and humans, has been reported to be a false negative in some cell-based assays.[13] As a result of these factors, it is currently not possible to predict in what species, tissues or cells a CAD will induce PLD, if at all.

Due to the complicated nature and possible multiple mechanisms of PLD, it is difficult to develop reliable methods to identify PLD inducers. To date, the determination of PLD relies on transmission electron microscopy (EM) for confirmation of the presence of multilamellar bodies,[14] examination of peripheral blood lymphocytes[15] and direct chemical analysis of phospholipids.[16] EM is considered as the gold standard method for the detection of PLD; however, this technique is a time-consuming and labor intensive methodology and does not allow for the screening of a large number of chemicals. In addition, drug-induced PLD detected by EM can be identified only after subchronic to chronic intake of compounds.[17] Therefore, a number of in vitro and in silico methods to identify PLD have been developed in the past decade.[18]

It is well-known that the potential to induce PLD depends strongly on the physicochemical and structural futures of the chemical. The first in silico methods mostly employed simple physicochemical descriptors, such as the pK_a of the most basic group and the logarithm of octanol/water partition coefficient (log P), to classify compounds as potential PLD inducers.[19,20] However, further analyses showed that these parameters are not sufficient to differentiate the PLD inducers from noninducers for certain chemical classes.[18,21] Therefore, other physicochemical parameters, such as the net charge at lysosomal pH, retention on HPLC columns and Langmuir surface balance,[19,22−24] or pharmacokinetic parameters - the volume of distribution[25] or parameters expressing binding to phospholipids, inhibition of enzymatic degradation of phospholipids and metabolic stability of CADs[26] have been utilized to predict the potential of PLD induction. In addition, the PLD-inducing potential of chemicals was evaluated by applying quantitative structure−activity relationships (QSAR) using the MC4PC, MDL-QSAR,[27] Symmetry and ADMET Predictor software[28] as well as machine learning algorithms (Random Forest and Support Vector Machine)[29] or Bayesian models.[20]

At the same time as model development, a number of in vitro cell based assays for PLD induction have been developed using various cell systems coupled with either electron microscopy[30−32] or the fluorescent probes Nile red[33,34] or *N*-(7-nitrobenz-2-oxa1,3-diazol-4-yl)-1,2-dihexadecanoyl-*sn*-glycero-3-phospholipid (NBD), *N*-(7-nitrobenz-2-oxa-1,3-diazol-4-yl)-1,2-dihexadecanoyl-*sn*-glycero-3-phosphoethanolamine, triethylammonium salt (NBD-PE),[13,35−41] or HCS LipidTOX Green phospholipidosis detection reagent to confirm phospholipid or lipid accumulation.[42] In addition, toxicogenomic assays have been adopted to measure cellular responses to the presence of phospholipidosis.[5,43] However, the results from the cell based in vitro assays can be contradictory, as the potential to induce PLD depends on the concentration[43] and the cell line used.[37] In vitro assays differ in their detailed methodology which can affect the sensitivity or specificity of the methods, for example erythromycin, loratadine, and quinidine were identified as positive in a LipidTox assay using HepG2cells[43] but are negative in the NBD-phosphocholine assay using CHO-K1 cells.[37]

Other complications in using in vitro data arise from differences between in vivo and in vitro results. For instance, this could be explained by metabolism in the in vivo studies. Goracci et al. discussed the effects of drug biotransformation on the ability to induce PLD and observed that for some chemicals, such as benzamide and ketoconazole, it is the metabolites which have the potential to induce PLD.[44]

Similarly, noncell based in vitro methods to assess phospholipidosis potential have been developed. Vitovič et al. estimated the PLD potency of compounds by measuring drug−phospholipid formation based on the critical micelle concentration (CMC) of short-chain acidic phospholipids using the surface tension activity.[24] Further, Zhou et al. described an alternative approach for the determination of PLD by measuring drug−lipid formation via changes in CMC of a short-chain lipid using fluorescent dye probe.[45] Other noncell in vitro assays have applied immobilized artificial membrane (IAM) chromatography and electrokinetic chromatography (EKC) with surfactant unilamellar vesicles as the pseudostationary phase[46] and the pH-gradient permeability (PAMPA) model.[47] The latter assay mimics one-way transport of CADs from the cytosol to the lysosome, while the former noncell based models characterize drug−lipid interactions.

High-throughput screening in vitro assays to detect PLD provide larger data sets compared to in vivo methods, especially those derived from EM. However, the possibility of discrepancy between the in vitro and in vivo PLD-inducing potential must be considered. Sun et al. compared cell-based PLD activities with in vivo data for 164 chemicals.[21] A set of 136 compounds exhibited similar PLD potential and provided reasonably good correlation. However, higher consistency was observed for noninducers (91%), with relatively low agreement along the PLD actives (56%). This shows that there is good relationship between in vivo and in vitro results for PLD noninducers only, whereas almost 50% of PLD results for inducers are inconsistent.

All in vivo and in vitro methods detecting PLD inducers produce data of differing quality. It is well-known that the performance of models developed to predict and/or identify adverse effects or properties of chemicals depends on the quality of data on which models were developed.[48,49] In the case of PLD, it is especially important as the occurrence of PLD can vary among different species and tissues. The situation is even more complicated by the significant discrepancy between in vivo and in vitro testing. Therefore, ideally for modeling, and especially to develop structural alerts, in vivo data confirmed by gold standard EM should be used as they are considered most reliable. However, the scarcity of such data, especially for noninducers drives the use of data from different methods. That type of data was used in a previous study from the authors, where a set of 32 most characteristic structural alerts, captured using SMARTS (SMILES arbitrary target specification) patterns, were developed.[18] The aim of this paper was to update the original SMARTS model by developing new structural alerts using a larger data set and then assess its predictive performance applying other publicly available data sets which were constructed differently and thus represent varying quality. Such analysis of the SMARTS model using data sets with different quality demonstrates explicitly the impact of the type of data on the predictability of the model. As such, this was not an exercise to compare the performance of different in

silico models for PLD, which has been undertaken previously by the authors,[18] rather to further develop structural alerts for PLD and assess the impact of the underlying data.

MATERIALS AND METHODS

Data Sources. To develop structural alerts capable of identifying PLD, two data sets published recently by the US FDA have been used.[27,50] The first contains 583 chemicals and the more recent consists of 743 compounds. After removing in-house (393) and duplicate chemicals (197), a final set of 736 compounds was obtained with 214 PLD inducers and 522 noninducers. The data were collected from various species, including rats, dogs, mice, monkeys, and humans as well as from a number of different tissue types, with the lungs and liver being common targets.

To assess the SMARTS model, two data sets representing different standards for overall quality were chosen. The first consists of 185 compounds with all negative chemicals being confirmed by EM to ensure the high quality of data.[29] The second was compiled from seven publicly available data sets and then curated by removing compounds with conflicting PLD results giving final set of 331 chemicals.[44]

Development of Structural Fragments. The strategy to develop structural alerts was described previously.[18] In general, they were developed based on the structural characteristics of PLD inducers: a hydrophobic ring system and a hydrophilic amine group. Then the structural features and fragments associated with the induction of PLD were captured using SMARTS patterns. After testing the sensitivity of a number of different structural fragments capable of identifying PLD inducers, the seven most characteristic and desirable SMARTS patterns were created and added to the previously developed 32 structural alerts. The structural alerts coded as SMARTS string can be implemented into a KNIME platform to provide automated and transparent workflows for the rapid screening of potential PLD inducers.[51]

Statistical Analysis. The SMARTS patterns were analyzed for their ability to assign the potential of each compound to induce PLD. A chemical was considered to have been assigned correctly if there was agreement between the decision made from the SMARTS patterns and the PLD observation for the compound considered.

Statistical performance was quantified by performing sensitivity and specificity analyses. Sensitivity measures the proportion of actual positives which are correctly identified, whereas specificity measures the proportion of negatives correctly assigned.[52]

RESULTS AND DISCUSSION

Phospholipidosis Data for Modeling and Assessment of Their Consistency. The accuracy of predictions and acceptance of the models depend greatly on the quality of data used to develop these models. For model development, a number of data sets have been compiled from different sources (literature, reports, study reports) from the past decade.[20,27,29,42,44,50] Usually the authors have compiled data from different assays including in vivo and in vitro procedures. For instance, Pelletier compiled 201 compounds from the literature and Pfizer's internal database.[20] Eighty-five of these chemicals were identified as in vivo PLD inducers, and 116, as noninducers. However, 84 chemicals were taken from Pfizer's internal database which means that no structure or name were available; hence they cannot be used for further modeling. The positive compounds were first identified through keyword searching for PLD synonyms and subsequently confirmed by EM. The noninducers were classified as chemicals with no evidence of phospholipidosis throughout the extensive preclinical testing at multiple doses and time points in at least two species. In 2008, the US FDA Centre for Drug Evaluation and Research (CDER) Phospholipidosis Working Group (PLWG) published a database of 583 compounds with 190 inducers and 393 noninducers.[27] Ninety-eight of these compounds were in-house chemicals, again without name or structure. Only 39 of the negative chemicals were confirmed by EM, the remainder noninducers were classified based on the absence of positive reported data. Similarly, 76 of the PLD inducers were confirmed by EM and 114 were considered positive due to other evidence. The data were taken from in vivo experiments in different species, such as rats, mice, dogs, monkeys and humans. This database was expanded in 2012 resulting in 743 compounds of which 385 were positive and 358 were negative.[50] This was achieved based on a rigorous keyword search for PLD synonyms across the published literature, internal US FDA archives of Investigational New Drug (IND) and New Drug Application (NDA) submissions and PharmaPendium. Again, the presence of a defined subset of keywords, such as phospholipid accumulation or foamy macrophages, indicated a PLD inducer, while the absence of these keywords suggested a noninducer. The chemicals were further classified into high or medium confidence categories depending of the type of keywords and the source of the literature or other report. These two data sets published by US FDA represent the largest, publicly available PLD databases; therefore they have been chosen for further research on the development of structural alerts to identify PLD.

The absence of PLD related keywords does not always mean the lack of PLD-inducing potential, as the chemical could not be tested at all. Therefore, the identification of noninducers, based only on the absence of specific keywords, is potentially erroneous. As a consequence, more efforts have recently been put into compiling data sets with PLD results, mostly negative, confirmed by the PLD gold standard EM in order to develop a more robust data set. Lowe at al. created such a data set of 185 chemicals (102 positive and 83 negative) from various literature sources.[29] All noninducers were confirmed by EM, as well as 68 inducers. The remaining 34 positive chemicals were identified by the presence of foamy macrophages. A similar approach was undertaken by Choi et al., who generated a database of 379 chemicals with 147 positives confirmed by EM and 232 negatives with no evidence of phospholipidosis from the data sources analyzed.[28] As the primary source, the FDA database of 743 chemicals[50] was utilized to collect the final 379 compounds. However, around half of them (the authors do not state exact number) were in-house chemicals with no information on structure and name.

There is a lot of discrepancy in PLD activity between different databases, e.g. Cloforex and Felbamate are referred to as being positive with high confidence in the CDER PLWG database from 2012[50] but as negative in the Lowe data set.[29] Recently, Goracci et al. investigated seven publicly available PLD databases published in the literature.[44] Initially, the authors compiled a database of 466 chemicals using all information available in these selected databases. After the analysis of the PLD activity, 24 chemicals were found to have contradictory activity from the different literature sources,

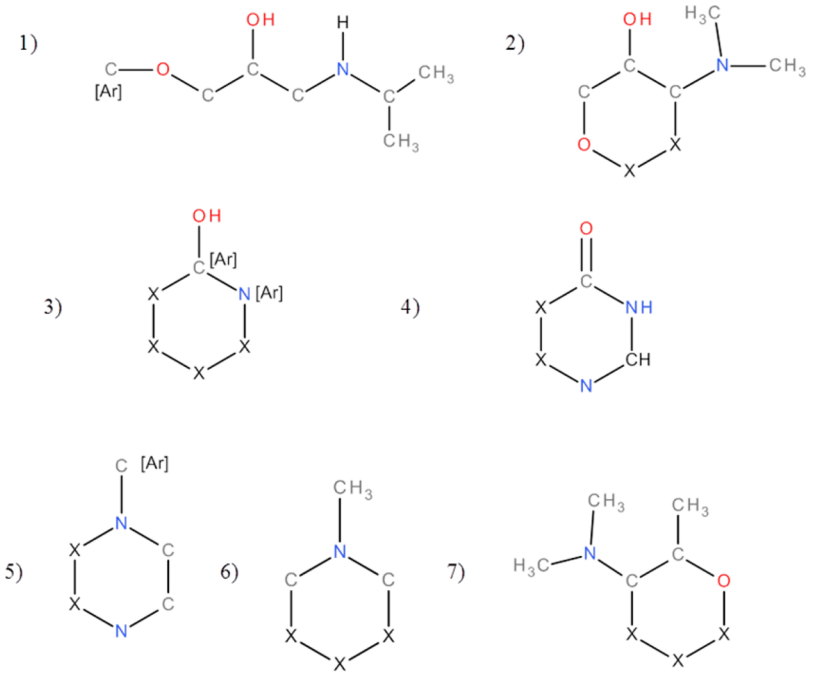


Figure 1. Structural alerts added to the updated SMARTS model. [Ar] aromatic atom. X = any atom not defined by a SMARTS string.

therefore they were further studied. As a result, 11 compounds were removed because of the lack of clear evidence of them inducing or not inducing PLD. The authors then used the VolSurf+ software to generate partial least squares (PLS) models which allowed for the identification of a number of doubtful PLD positive and PLD negative assignments. Finally, to explain some false negative predictions, metabolism was taken into account this helped to clarify some negative predictions, e.g. for benzamide and ketoconazole. Consideration of all these factors led to the removal of 135 compounds with uncertain PLD activity and gave a final, curated database of 331 chemicals.

The discrepancy in reported activity may result from using different assays to assess PLD potential. As was mentioned previously, there can be a lack of correlation between in vivo and in vitro induced PLD, mostly caused by the involvement of drug metabolites, rather the drug itself inducing PLD. For example, in the case of amiodarone, phospholipidosis may be caused by desethylamiodarone, its major metabolite.[53] Therefore, this effect may have been overlooked in some earlier in vitro studies. Moreover, the pharmacological effects of CADs are strongly dose-dependent; thus, variation in dosing regimens and in the duration of drug administration may also have an impact on the assay results.[54] Also, tissue specificity can influence the PLD potential due to drug distribution and pharmacokinetics.

Updating the SMARTS Model. The original SMARTS model for PLD[18] was developed based on 450 chemicals taken from the US FDA database, consisting of 583 compounds; in work by Kruhlak et al.[27] The model comprises 32 structural alerts, presented as SMARTS patterns, which are the most characteristic and desirable structural features for inducing PLD. The SMARTS patterns were divided into five main chemical groups, those for: primary amines, secondary amines, tertiary amines, cyclic amines and aromatic systems. Additionally, a group of SMARTS patterns, to assign to the presence of ring systems, was established to allow for improved identification of PLD inducers. The fragments that characterized the ring systems could not be used in combination with the other 32 SMARTS patterns as they would lead to many false positives (many PLD noninducers possess a ring system). Moreover, two undesirable structural features (carboxylic acid and nitro groups) for potential PLD inducers were also established. Most of the non-PLD inducers from the 450 chemicals were classified based on the absence of a reported positive result and only 39 chemicals from the original 393 negative chemicals were confirmed by EM. The positive classification was performed based on the presence of lamellar bodies confirmed by EM or the presence of foamy macrophages. In 2012 the US FDA database was updated resulting in a total of 743 chemicals (296 were in-house chemicals).[50] Combining these two data sets and removing the in-house and overlapping chemicals gave data set of 736 compounds (214 PLD inducers and 522 noninducers). It is one of the largest data set publicly available for PLD. Initially, this data set was used to test the original SMARTS model. As a result, 102 PLD inducers and 450 noninducers were identified correctly, with a low sensitivity of 48.1% but a reasonably high specificity of 85.6%. Closer inspection of PLD inducers showed that six chemicals: flunisolide, lovastatin, lubiprostone, megestrol acetate, pravastatin sodium and ursodiol do not possess a nitrogen atom which is necessary for the protonation in the acidic environment of the lysosomes. Moreover, there are 22 PLD inducers with undesirable fragments: carboxylic acid or nitro group motifs, which classified them directly as noninducers.

Taking into account the presence of 286 new chemicals in this newly compiled PLD data set and, more importantly, due to the poor predictability of the original SMARTS model, an analysis of the new structures within the combined PLD data set was undertaken to investigate novel structural features able to distinguish PLD inducers from noninducers. As a result, seven new structural fragments capable of identifying potential PLD-inducers have been developed (see Figure 1). Two of the

dx.doi.org/10.1021/ci500233k | J. Chem. Inf. Model. 2014, 54, 2224−2232

Table 1. Updated SMARTS Patterns for the Identification of PLD Inducers and Noninducers

structural group	SMARTS pattern
primary amine	[NH2][CX4;!R][CX4]
	[NH2][CX4](C)(C)(C)
	NH2][C;R]([C;R][OH])[C;R][OH]
	[NH2][C;R]([CH;R])[C;R]O
	[NH2]CC1OCCCC1
	[NH2][C;R]([C;R])[C;R](O)[O;R]
	[NH2]c1c(Br)cc(Br)cc1
	[NH]=C([NH2])c1ccccc1
	[NH]=C([NH2])[NH]C[NH]
secondary amine	c[CX4;!R][CX4;!R][NH][CH2][CH3]
	[C;!R][NH][C;R]([C;R])c
	[C;!R][C;!R][C;!R][NH][C;!R][C;!R]c
	c1[cH1]c[cH1][cH1]c1[NH]c1[cH1][cH1][cH1][cH1]c1
	c[NH][C;!R][C;R]
	[CH3][NH][CH2;!R][CH2;!R]
	[CH3][NH][C;R]([C;R][OH])[C;R][OH]
	[CX4;!R][NH][CX4;!R][CX4;!R][CX4;!R]Oc1c2ccccc2ccc1 [CX4][NH][CX4;R]
	c[OX2][CX4;!R][CX4;!R]([OH])[CX4;!R][NH][CH]([CH3])[CH3][a]
tertiary amine	[CX4;!R][N;!+]([CX4;!R])[CX4;!R][CX4]
	[CH3][N;!+]([CH3])[CH2][CH]=C(c)c
	[CH3][N;!+]([CH3])[CX4;R][CX4;R]([OH])[CX4;R][OX2;R][a]
	[CH3][N;!+]([CH3])[CX4;R][CX4;R]([CH3])[OX2;R][a]
cyclic amine	[nH]1n = ccc1
	cO[CH2][CH2][N;!+]1CCCC1
	[CH3][N;!+]1CC[N;!+]([C;!R])CC1
	[NH;R][C;R](C)[C;R]cc[C;R]
	[NH]1C(C)CCCC1
	[N;!+]1[CH2][CH2][CH]([NH])[CH2][CH2]1
	[N;!+]1[CH2][CH2]C(c2ccc(Cl)cc2)[CH2][CH2]1
	[n]c[OH][a]
	[#6;r6](=O)[#7H;r6][#6H;r6][#7;r6][a]
	c[N;R][CX4;R][CX4;R][N;R][CX4][a]
	[CH3][N;R!+]([CX4;R])[CX4;R][a]
aromatic system	cN([CX4][CX4][CX4][NX3;!R])c
	cN([CX4][CX4][NX3])c
	cN([CX4][CX4][CX4;R][N;R])c
	c1ccccc1[CH2]c1ccccc1
	c[CH]([N;R])c
ring system	[R;a]
	[R;!a]
acidic groups	[#6,#1]C(=O)[OH]
	[CH](=O)[OH]
	[#6]N(=O)=O
	[#6][N+](=O)[O-]

[a]New SMARTS patterns.

new alerts describe the keto−enol tautomeric forms of fragment found in allopurinol (structural alerts 3 and 4 in Figure 1). According to the original division of SMARTS patterns into five main chemical groups (primary amines, secondary amines, tertiary amines, cyclic amines, and aromatic systems), four new fragments are assigned as cyclic amine (two of them describe the keto−enol tautomerism), two as tertiary amine, and one describes the secondary amine. Additionally, one of the original structural alerts capturing the presence of a secondary amine and first developed based on the structure of AY-9944 has been updated due to error in structure in the first version of the US FDA database.[27] This mistake is also repeated in the Lowe database[29] which was corrected for the further analysis. Table 1 presents the original 32 SMARTS patterns together with the seven new patterns (flagged with [a]). Application of the updated structural fragments allows for the correct identification of 131 inducers and 418 noninducers, which gives sensitivity and specificity of 60.7 and 80.5%, respectively. It shows the significant improvement in the identification of PLD-inducers as the sensitivity increases from 47.7% to 60.7%. However, at the same time, the identification of noninducers decreases about 6% (number of negative chemicals predicted correctly dropped from 450 to 418). Therefore, the overall concordance stays almost on the same level (74.7% and 75.0% for the updated and original SMARTS model, respectively). Despite this, the substantial increase in the correct identification of PLD-inducers, which is very

Table 2. Predictive Statistics of the Original and Updated SMARTS Model Used for Three Data Sets

		sensitivity (%)		specificity (%)		concordance (%)	
data set	Nr chemicals	original model	updated model	original model	updated model	original model	updated model
combined US FDA	736 214 inducers 522 noninducers	47.7	60.7	86.2	80.5	75.0	74.7
Goracci	331 127 inducers 204 noninducers	70.9	83.5	84.8	78.9	79.5	80.4
Lowe	185 102 inducers 83 noninducers	82.3	90.2	87.9	84.3	84.9	87.6

important in the drug development process, supports the inclusion of the seven new structural fragments.

Assessment of the Predictability of the Updated SMARTS Model. To develop the structural alerts captured as SMARTS patterns to identify potential PLD inducers, the two largest, publicly available data sets compiled by US FDA were used. The choice of these databases was driven mainly by their size and diversity, not necessarily by their quality. Considering how these databases were constructed, it could be assumed that many of the compounds within them, especially the non-inducers identified based only on the lack of evidence of induction PLD, could be classified incorrectly. However, it was decided that for the investigation of structural features potentially responsible for inducing PLD, the larger and more versatile databases are useful. Conversely, it is well-known that the performance of any in silico model depends strongly on the quality of data used to develop it as well as the quality of data used to make such an assessment of performance.[48,49] Therefore, the goal of this study was to investigate how the quality of data set impacts on the predictability of the SMARTS model. For this purpose, two databases of different quality were chosen to assess the predictability of SMARTS model: the refined Goracci data set[44] and the Lowe data set.[29] The results from this external evaluation were then compared with the goodness of fit of the training set of 736 chemicals.

The Goracci data set of 331 chemicals was compiled after careful curation taking into account different factors, such as inconsistency of PLD potential in original literature sources and metabolism. However, there are still chemicals whose PLD potential is still uncertain. For example, for 127 inducers there is still one chemical: ursodiol, which does not contain a nitrogen atom, which is crucial to fulfill structural criteria (basic amine group able to be protonated in lysosome) for inducing PLD. Moreover, there are five inducers with undesirable acidic groups: bepotastine besilate, bilirubin, levodopa, ranitidine, and ursodiol. These chemicals have been identified directly as noninducers by the SMARTS model. Some of other false negative predictions can be explained by the biotransformation of parent compound. For example, phenacetin is metabolized to 4-ethoxyaniline, which demonstrates much greater permeability through the lysosome membrane than the parent compound, therefore it has higher PLD-inducing potential.[47,55] Among the 204 noninducers, 45 compounds incorporate acidic motifs, which assign them immediately as noninducers. Here 43 PLD noninducers were identified as positive. One of them, methapyrilene, possesses the CADs motifs, but it was determined that this compound is metabolized to methapyrilene-*N*-oxide, which lacks the possibility of protonation of a nitrogen atom.[56] However, there are in vitro studies in CHO-K1 cells, which identified methapyrilene as a PLD inducer.[41] This is another example of contradictory PLD activity determined by in vivo and in vitro assays. Overall, the updated SMARTS model has identified 106 PLD inducers and 161 noninducers correctly, giving a sensitivity of 82.7%, specificity of 78.9%, and concordance of 80.4%.

The second evaluation data set was compiled by Lowe at al. and represents data of high quality, as all negative results were confirmed by EM.[29] Among the 102 PLD inducers only two chemicals, bilirubin and chloramphenicol, have an undesirable acidic motif which were therefore directly identified as noninducers. Additionally, another eight positive chemicals were classified falsely as noninducers. Two of the false negative chemicals were removed from the final Goracci database. Chloramphenicol was removed due to the variability of PLD potential among the different literature sources and suramin was eliminated because of the lack of strong evidence to induce PLD in the literature. Loratadine was reported to undergo hepatic biotransformation into descarboethoxyloratadine (DCL), also called desloratadine, which is the active metabolite to induce PLD.[44] Therefore, loratadine was replaced by desloratadine in the Goracci final data set. Interestingly, in the extended version of the FDA data set both chemicals: parent and metabolite are present as PLD inducers with medium confidence. Another false negative chemical, paraquat, exists in the protonated form, which cannot be identified by any structural alerts, as it is assumed that only neutral compounds can pass through the lysosomal membrane and then be protonated in the acidic environment. Analysis of the 83 noninducers showed that 20 of them possess undesirable acidic motifs. Thirteen chemicals without PLD-inducing potential were identified incorrectly as inducers. Four of them, buspirone, clarithromycin, physostigmine, and rifampin, were identified by the new structural fragments. It is worth mentioning that clarithromycin is an analog to the PLD inducer erythromycin and in the FDA database is classified as an inducer with medium confidence. Because of that discrepancy, that chemical has been removed from the Goracci database, similar to buspirone, which is classified as a PLD inducer by Orogo at al. with medium confidence.[50] Overall, the updated SMARTS model identified 162 chemicals correctly giving a concordance of 87.6%, with sensitivity and specificity of 90.2 and 84.3%, respectively.

The predictive performance of the updated SMARTS model for the three data sets: the combined US FDA (training data set), Lowe, and Goracci is shown in Table 2. The identification of PLD inducers increases significantly when databases with higher quality were assessed. Also the specificity of the SMARTS model rises for the database, where all noninducers

were identified by the gold standard method for PLD. This is the clear evidence that the prediction performance greatly depends on the quality of data used for such assessment.

Finally, the assessment of the predictive performance of the original SMARTS model was performed for these three databases (see Table 2) and then compared with the predictability of the updated SMARTS model. First, the same trend of predictability was noted, when three data sets with different quality were used. The highest correct identification of PLD inducers, as well as noninducers, was obtained for the database with best quality—the Lowe data set. At the same time, for the biggest and most versatile database used as training data set for developing SMARTS model, the lowest sensitivity and concordance were obtained. Comparing the predictive statistics of both models, it was shown that the new structural alerts allowed for the significant increase in sensitivity of more than 10%. On the other hand, the correct identification of noninducers was reduced, but only by about 5%. However, overall concordance of both models was at the same level for two of the studied databases. Only for the Lowe data set was there a 2.7% increase of concordance when the updated SMARTS model was applied.

CONCLUSIONS

PLD still remains a complex phenomenon with unclear mechanisms. Recently, many models have been developed to identify potential PLD inducers. Our previous study showed that PLD is linked directly to molecular (sub)structures: hydrophobic, cyclic moieties with hydrophilic, peripheral amine groups. Based on that concept, a set of SMARTS patterns capable of identifying potential PLD inducers has been developed. In this study, the structural alerts have been updated using larger data set of 736 chemicals and as a result seven additional SMARTS patterns have been established. These new structural fragments significantly increase the correct identification of PLD inducers; however at the same time, there is a rise in false positive prediction. The validation of the SMARTS model using different PLD data sets showed a strong correlation between the predictive performance of the model and the quality of database used for the assessment. The best predictive statistics were obtained for the data set with the highest reliability, where all negative chemicals were confirmed by the gold standard method—transmission electron microscopy—and not only by the lack of evidence of PLD. This study showed the importance of the type of data used for modeling and evaluation, especially in the case of phospholipidosis, which is species, tissue, and dose dependent. In the ideal scenario, the most reliable data should be used for developing in silico models, as well as for the validation process. However, the scarcity of high quality PLD data determines the use of more diverse data obtained from different types of assays, species, and tissues for developing models, especially structural alerts capturing the molecular fragments associated with inducing of PLD. However, the assessment of predictability of such model strongly depends on the quality of data used for that purpose.

ASSOCIATED CONTENT

Supporting Information

Three data sets used in this study for the development of new structural alerts and for the assessment of the predictability of SMARTS models. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

*E-mail: m.t.cronin@ljmu.ac.uk.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

The funding provided by the eTOX project, grant agreement number 115002, under the Innovative Medicines Initiative Joint Undertaking (IMI-JU), the European Community's Seventh Framework Program (FP7/2007-2013) COSMOS project under grant agreement number 266835, and from Cosmetics Europe is gratefully acknowledged.

REFERENCES

(1) Reasor, M. J.; Hastings, K. L.; Ulrich, R. G. Drug-induced phospholipidosis: issues and future directions. *Expert Opin. Drug Saf.* **2006**, *5*, 567−583.

(2) Johnson, W. J.; Warner, G. J.; Yancey, P. G.; Rothblat, G. H. Lysosomal metabolism of lipids. *Subcell. Biochem.* **1996**, *27*, 239−293.

(3) Kodavanti, U. P.; Mehendale, H. M. Cationic amphiphilic drugs and phospholipid storage disorder. *Pharmacol. Rev.* **1990**, *42*, 327−54.

(4) Goldman, S. D.; Funk, R. S.; Rajewski, R. A.; Krise, J. P.Mechanisms of amine accumulation in, and egress from, lysosomes. *Bioanalysis* **2009**, *1*, 1445−59.

(5) Sawada, H.; Takami, K.; Asahi, S. A toxicogenomic approach to drug-induced phospholipidosis: analysis of its induction mechanism and establishment of a novel in vitro screening system. *Toxicol. Sci.* **2005**, *83*, 282−292.

(6) Abe, A.; Hiraoka, M.; Shayman, J. A. A role for lysosomal phospholipase A2 in drug induced phospholipidosis. *Drug Metab. Lett.* **2007**, *1*, 49−53.

(7) Lullmann-Rauch, R. Drug-induced lysosomal storage disorders. *Front. Biol.* **1979**, *48*, 49−130.

(8) Nonoyama, T.; Fukuda, R. Drug-induced Phospholipidosis-Pathological Aspects and Its Prediction. *J. Toxicol. Pathol.* **2008**, *21*, 9−24.

(9) Reasor, M. J. A review of the biology and toxicologic implication of the induction of lysosomal lamellar bodies by drugs. *Toxicol. Appl. Pharmacol.* **1989**, *97*, 47−56.

(10) Halliwell, W. H. Cationic amphiphilic drug-induced phospholipidosis. *Toxicol. Pathol.* **1997**, *25*, 53−60.

(11) Reasor, M. J.; Kacew, S. Drug-induced phospholipidosis: are there functional consequences? *Exp. Biol. Med.* **2001**, *226*, 825−30.

(12) Morelli, J. K.; Ciaccio, P. J. Differential cytotoxicity responses by dog and rat hepatocytes to phospholipogenic treatments. *J. Toxicol.* **2013**, 308594.

(13) Morelli, J. K.; Buehrle, M.; Pognan, F.; Barone, L. R.; Fieles, W.; Ciaccio, P. J. Validation of an ex vivo screen for phospholipidosis using a high-content biology platform. *Cell Biol. Toxicol.* **2006**, *22*, 15−27.

(14) Hook, G. E. R. Alveolar proteinosis and phospholipidoses of the lungs. *Toxicol. Pathol.* **1991**, *19*, 482−513.

(15) Lullmann, H.; Lullmann-Rauch, R.; Wassermann, O. Drug induced phospholipidosis. *Crit. Rev. Toxicol.* **1975**, *2*, 185−218.

(16) Reasor, M. J.; Ogle, C. L.; Walker, E. R.; Kacew, S. Amioadrone-induced phospholipidosis in rat alveolar macrophages. *Am. Rev. Respir. Dis.* **1988**, *137*, 510−518.

(17) Atienzar, F.; Gerets, H.; Dufrane, S.; Tilmant, K.; Cornet, M.; Dhalluin, S.; Ruty, B.; Rose, G.; Canning, M. Determination of phospholipidosis potential based on gene expression analysis in HepG2 cells. *Toxicol. Sci.* **2007**, *96*, 101−114.

(18) Przybylak, K. R.; Cronin, M. T. D. *In silico* studies of the relationship between chemical structure and drug induced phospholipidosis. *Mol. Inf.* **2011**, *30*, 415−429.

(19) Ploemen, J. P.; Kelder, J.; Hafmans, T.; van de Sandt, H.; van Burgsteden, J. A.; Saleminki, P. J.; van Esch, E. Use of physicochemical calculation of pKa and CLogP to predict phospholipidosis- inducing

potential: A case study with structurally related piperazines. *Exp. Toxicol. Pathol.* **2004**, *55*, 347−355.
(20) Pelletier, D. J.; Gehlhaar, D.; Tilloy-Ellul, A.; Johnson, T. O.; Greene, N. Evaluation of a published in silico model and construction of a novel Bayesian model for predicting phospholipidosis inducing potential. *J. Chem. Inf. Model.* **2007**, *47*, 1196−1205.
(21) Sun, H.; Shahane, S.; Xia, M.; Austin, C. P.; Huang, R. Structure based model for the prediction of phospholipidosis induction potential of small molecules. *J. Chem. Inf. Model.* **2012**, *52*, 1798−1805.
(22) Tomizawa, K.; Sugano, K.; Yamada, H.; Horii, I. Physicochemical and cell-based approach for early screening of phospholipidosis inducing potential. *J. Toxicol. Sci.* **2006**, *31*, 315−324.
(23) Valco, K.; Du, C. M.; Bevan, C. D.; Reynolds, D. P.; Abraham, M. H. Rapid-gradient HPLC method for measuring drug interactions with immobilized artificial membrane: comparison with other lipophilicity measures. *J. Pharm. Sci.* **2000**, *89*, 1085−1096.
(24) Vitovic, P.; Alakoskela, J. M.; Kinnunen, K. J. Assessment of drug-lipid complex formation by a high throughput Langmuir-Balance and correlation to phospholipidosis. *J. Med. Chem.* **2008**, *51*, 1842−1848.
(25) Hanumegowda, U. M.; Wenke, G.; Regueiro-Ren, A.; Yordanova, R.; Corradi, J. P.; Adams, S. P. Phospholidosis as a function of basicity, liphophilicity and volume of distribution of compound. *Chem. Res. Toxicol.* **2010**, *23*, 749−55.
(26) Kuroda, Y.; Saito, M. Prediction of phospholipidosis-inducing potential of drugs by in vitro biochemical and physicochemical assays followed by multivariate analysis. *Toxicol. in vitro* **2010**, *24*, 661−668.
(27) Kruhlak, N. L.; Choi, S. S.; Contrera, J. F.; Weaver, J. L.; Willard, J. M.; Hastings, K. L.; Sancilio, L. F. Development of a phospholipidosis database and predictive quantitative structure-activity relationship (QSAR) models. *Toxicol. Mech. Methods* **2008**, *18*, 217−27.
(28) Choi, S. S.; Kim, J. S.; Valerio, L. G., Jr.; Sadrieh, N. *In silico* modeling to predict drug-induced phospholipidosis. *Toxicol. Appl. Pharmacol.* **2013**, *269*, 195−204.
(29) Lowe, R.; Glen, R. C.; Mitchell, J. B. Predicting phospholipidosis using machine learning. *Mol. Pharmaceutics* **2010**, *7*, 1708−14.
(30) Van Bambeke, F.; Gerbaux, C.; Michot, J. M.; d'Yvoire, M. B.; Montenez, J. P.; Tulkens, P. M. Lysosomal alterations induced in cultured rat fibroblasts by long-term exposure to low concentrations of azithromycin. *J. Antimicrob. Chemother.* **1998**, *42*, 761−767.
(31) Hostetler, K.; Vande Berg, J.; Aldern, K.; Brophy, G. Effect of amiodarone on the phospholipid and lamellar body content of lymphoblasts in vitro and peripheral blood lymphocytes in vivo. *Biochem. Pharmacol.* **1991**, *41*, 1007−1013.
(32) Pospischil, A.; Walter, P.; Dingemanse, J. Phospholipidosis in healthy subjects participating in clinical studies: ultrastructural findings in white blood cells. *Exp. Toxicol. Pathol.* **2010**, *62*, 567−571.
(33) Casartelli, A.; Bonato, M.; Cristofori, P. A cell-based approach for the early assessment of the phospholipidogenic potential in pharmaceutical research and drug development. *Cell Biol. Toxicol.* **2003**, *19*, 161−176.
(34) Xia, Z.; Appelkvist, E. L.; Depierre, J. W.; Nassberger, L. Tricyclic antidepressant-induced lipidosis in human peripheral monocytes in vitro, as well as in a monocyte-derived cell line, as monitored by spectrofluorimetry and flow cytometry after staining with Nile red. *Biochem. Pharmacol.* **1997**, *53*, 1521−1532.
(35) Gum, R. J.; Hickman, D.; Fagerland, J. A.; Heindel, M. A.; Gagne, G. D.; Schmidt, J. M.; Michaelides, M. R.; Davidsen, S. K.; Ulrich, R. G. Analysis of two matrix metalloproteinase inhibitors and their metabolites for induction of phospholipidosis in rat and human hepatocytes(1). *Biochem. Pharmacol.* **2001**, *62*, 1661−1673.
(36) Ulrich, R. G.; Kligore, K. S.; Sun, E. L.; Cramer, C. T.; Ginsberg, L. C. An in vitro fluorescence assay for the detection of drug-induced cytoplasmic lamellar bodies. *Toxicol. Meth.* **1991**, *1*, 89−105.
(37) Kasahara, T.; Tomita, K.; Murano, H.; Harada, T.; Tsubakimoto, K.; Ogihara, T.; Ohnishi, S.; Kakinuma, C. Establishment of an in vitro high-throughput screening assay for etecting phospholipidosis-inducing potential. *Toxicol. Sci.* **2006**, *90*, 133−141.
(38) Sun, M.; Zhao, C.; Gfesser, G. A.; Thiffault, C.; Miller, T. R.; Marsh, K.; Wetter, J.; Curtis, M.; Faghih, R.; Esbenshade, T. A.; Hancock, A. A.; Cowart, M. Synthesis and SAR of 5-Amino- and 5-(Aminomethyl)benzofuran Histamine H3 Receptor Antagonists with Improved Potency. *J. Med. Chem.* **2005**, *48*, 6482−6490.
(39) Mesens, N.; Steemans, M.; Hansen, E.; Peters, A.; Verheyen, G.; Vanparys, P. A 96-well flow cytometric screening assay for detecting in vitro phospholipidosis-induction in the drug discovery phase. *Toxicol. in vitro.* **2009**, *23*, 217−226.
(40) Bhandari, N.; Figueroa, D. J.; Lawrence, J. W.; Gerhold, D. L. Phospholipidosis assay in HepG2 cells and rat or rhesus hepatocytes using phospholipid probe NBD-PE. *Assay Drug Dev. Technol.* **2008**, *6*, 407−19.
(41) van de Water, F. M.; Havinga, J.; Ravesloot, W. T.; Horbach, G. J. M. J.; Schoonen, W. G. E. J. High content screening analysis of phospholipidosis: Validation of a 96-well assay with CHO-K1 and HepG2 cells for the prediction of in vivo based phospholipidosis. *Toxicol. in vitro.* **2011**, *25*, 870−1882.
(42) Muehlbacher, M.; Tripal, P.; Roas, F.; Kornhuber, J. Identification of Drugs Inducing Phospholipidosis by Novel in vitro Data. *Chem. Med. Chem.* **2012**, *7*, 1925−1934.
(43) Nioi, P.; Perry, B. K.; Wang, E. J.; Gu, Y. Z.; Snyder, R. D. *In vitro* detection of drug-induced phospholipidosis using gene expression and fluorescent phospholipid based methodologies. *Toxicol. Sci.* **2007**, *99*, 162−173.
(44) Goracci, L.; Ceccarelli, M.; Bonelli, D.; Cruciani, G. Modeling phospholipidosis induction: reliability and warnings. *J. Chem. Inf. Model.* **2013**, *53*, 1436−1446.
(45) Zhou, L.; Geraci, G.; Hess, S.; Yang, L.; Wang, J.; Argikar, U. Predicting phospholipidosis: a fluorescence noncell based in vitro assay for the determination of drug−phospholipid complex formation in early drug discovery. *Anal. Chem.* **2011**, *83*, 6980−6987.
(46) Jiang, Z.; Reilly, J. Chromatography approaches for early screening of the phospholipidosis-inducing potential of pharmaceuticals. *J. Pharm. Biomed. Anal.* **2012**, *61*, 184−190.
(47) Balogh, G. T.; Müller, J.; Könczöl, Á. pH-gradient PAMPA-based in vitro model assay for drug-induced phospholipidosis in early stage of drug discovery. *Eur. J. Pharm. Sci.* **2013**, *49*, 81−89.
(48) Cronin, M. T. D. Finding the Data to Develop and Evaluate (Q)SARs and populate categories for toxicity prediction. In *In silico Toxicology: Principles and Applications*; Cronin, M. T. D., Madden, J. C., Eds.; The Royal Society of Chemistry: Cambridge, United Kingdom, 2010; pp 31−56.
(49) Cronin, M. T. D.; Schultz, T. W. Pitfalls in QSAR. *J. Mol. Struc.-THEOCHEM.* **2003**, *622*, 39−51.
(50) Orogo, A. M.; Choi, S. S.; Minnier, B. L.; Kruhlak, N. L. Construction and consensus performance of (Q)SAR models for predicting phospholipidosis using a dataset of 743 compounds. *Mol. Inf.* **2012**, *31*, 725−739.
(51) Berthold, M. R.; Cebron, N.; Dill, F.; Gabriel, T. R.; Kötter, T.; Meinl, T.; Ohl, P.; Sieb, C.; Thiel, K.; Wiswedel, B. KNIME: The Konstanz Information Miner. In *Studies in Classification, Data Analysis, and Knowledge Organization*; Preisach, C., Burkhardt, H., Schmidt-Thieme, L., Decker, R., Eds.; Springer: Berlin Heidelberg, 2008; pp 319−326.
(52) Rowe, P. H. Statistical Methods for Continuous Measured Endpoints in *In silico* Toxicology. In *In silico Toxicology: Principles and Applications*; Cronin, M. T. D., Madden, J. C., Eds.; The Royal Society of Chemistry: Cambridge, United Kingdom, 2010; pp 252−273.
(53) Seki, S.; Itagaki, S.; Kobayashi, M.; Hirano, T.; Iseki, K. Amiodarone increases the accumulation of DEA in a human alveolar epithelium-derived cell line. *Biol. Pharm. Bull.* **2008**, *31*, 1449−1452.
(54) Kapatou, E.; Skyrlas, A.; Agelaki, M. G.; Pantos, C.; Kolettis, T. M.; Malamou-Mitsi, V. Amiodarone attenuates apoptosis, but induces phospholipidosis in rat alveolar epithelial cells. *J. Physiol. Pharmacol.* **2010**, *61*, 671−677.
(55) Kudo, S.; Umehara, K.; Hosokawa, M.; Miyamoto, G.; Chiba, K.; Satoh, T. Phenacetin deacetylase activity in human liver

microsomes: distribution, kinetics, and chemical inhibition and stimulation. *J. Pharmacol Exp Therap.* **2000**, *294*, 80−88.
(56) Kelly, D. W.; Holder, C. L.; Korfmacher, W. A.; Getek, T. A.; Lay, J. O., Jr.; Casciano, D. A.; Shaddock, J. G.; Duhart, H. M.; Slikker, W., Jr. Metabolism of methapyrilene by Fischer-344 rat and B6C3F1 mouse hepatocytes. *Xenobiotica* **1992**, *22*, 1367−1381.