



A simple model to solve a complex drug toxicity problem†

Cite this: *Toxicol. Res.*, 2019, **8**, 157

Vaibhav A. Dixit 

Linear drug toxicity models like therapeutic index (TI), physicochemical rules (rule of five, 3/75), ligand efficiency indices (LEI), ideal pharmacokinetic (PK) and pharmacodynamic (PD) profiles are widely used in drug discovery and development. In spite of this, predicting drug toxicity at various stages remains challenging and the overall productivity (<20%) and ultimate benefit to the patients remain low. A simple drug toxicity model, “Drug Toxicity Index” (DTI), is developed here using 711 oral drugs. DTI redefines drug toxicity as scaled biphasic and exponential functions of PD, PK and physicochemical parameters. PD, PK and physicochemical toxicity contributions were estimated from the on and off target IC_{50} , maximum unbound plasma drug concentration (free C_{max}), and $\log D$ values, respectively. These contributions are then scaled by molar dose and oral bioavailability and the logarithm of the sum of scaled contributions is DTI. Drugs with DTI above the WHO ATC drug category specific average values consistently have toxic profiles, while drugs with DTI below this average are relatively safe. DTI performs better than standard rules for lead optimization, LEI and exposure based TIs in identifying safe and toxic drugs. DTI classifies 392 drugs reported in the US-FDA’s Liver Toxicity Knowledge Base (LTKB) with an AUC for ROC curves of 0.91–0.64 for different WHO ATC categories. DTI has been used to predict network meta-analysis results on relative toxicity within/across eight different therapeutic areas. It is useful in understanding PD, PK and physicochemical toxicity contributions and identifying potentially toxic drugs and the toxicity of recently approved drugs. Decision trees are proposed for applying the DTI concept in preclinical drug discovery and clinical trial settings. DTI can potentially reduce failure in drug discovery and might be useful in therapeutic drug monitoring and in xenobiotic and environmental toxicity studies.

Received 30th September 2018,

Accepted 28th November 2018

DOI: 10.1039/c8tx00261d

rs.c.li/toxicology-research

1. Introduction

Drug safety is estimated in drug discovery by the measurement of the therapeutic index (TI) and its variants *e.g.* exposure based therapeutic indices (IC_{50}/C_{max}).^{1,2} But investigational drugs continue to fail in clinical trials due to low safety and efficacy.^{3,4} TI assumes simplified linear relationships between receptor affinity, maximum unbound plasma drug concentration (C_{max}) and toxicity. But high TI does not guarantee safety. For drugs metabolized by cytochrome P450 (CYP450), estimating TI based on target potencies alone is insufficient.

Toxicity may be due to the accumulation in a specific organ/tissue (*e.g.* bosentan), the co-administration of other drugs affecting ADMET (absorption, distribution, metabolism, elimination and toxicity),⁵ C_{max} reaching off target IC_{50} , or the high C_{max} required for therapeutic effects, see ref. 1.

A 2D “TI heatmap grid” was developed earlier to compare efficacy and safety data.¹ The limitations of this approach are: a lack of a universal TI value, large variations and difficulty of direct comparison across chemical or therapeutic classes. TI is used for risk–benefit assessment of known toxicities and is not applicable to rare and idiosyncratic toxicities. For example, the Valdecoxib TI values for COX-1 are 7 times those for CYP2C19 inhibition, but its toxicity is due to on target activity in cardiomyocytes, seldom considered in TI calculations.^{1,6} Using the same C_{max} value for calculating TI at different doses seems erroneous.² The major off target effects involve proteins closely related to the therapeutic target, CYP450s, hERG and BSEP activities.¹ The use of 44 proteins (from GPCR, ion channel, neurotransmitter transporter, nuclear hormone receptors, AChE, COX, MAO, phosphodiesterase and kinase families) for off target profiling has been recommended.⁷ Pharmaceutical companies have been screening this protein

Department of Pharmacy, Birla Institute of Technology and Sciences Pilani (BITS Pilani), Vidya Vihar Campus, Street number 41, Pilani, 333031 Rajasthan, India. E-mail: vaibhav.dixit@pilani.bits-pilani.ac.in, vaibhavdixit@gmail.com; Tel: +91 1596 255652, +91-7709129400

† Electronic supplementary information (ESI) available: Word file with ANOVA analysis for WHO ATC drugs, discussion on potentially toxic drugs from the remaining categories, detailed methodology for data collection and analysis, tables with relative efficacy/toxicity data, average values and performance of the different parameters considered, and pairwise *t*-test data for different drug categories. Microsoft Excel files containing raw data, DTI analysis, comparison with US-FDA’s LTKB database and ROC analysis. See DOI: 10.1039/c8tx00261d

set for decades, but data for most drugs are not available in the public domain. When such data are available the interpretation of complex TI heatmap grids could still present practical difficulties.

Physicochemical property and toxicity relationships have been extensively reviewed.^{8,9} Trends between oral bioavailability, drug-like properties and hepatotoxicities, *e.g.* the RO5,¹⁰ 3/75, RO3,¹¹ RO2,¹² RO2-RM,¹³ and 4/400 rule,⁸ have been useful in lead optimization. Typically, compound libraries are filtered using predicted physicochemical properties, and drug-likeness rules, and subjected to biological activity (IC_{50} , K_i , K_d or K_m) predictions. Alternatively, high throughput, fragment, focused, physiological or even NMR screening, ligand and/or structure based design approaches can be used to find lead compounds.^{9,14} Predictions are then tested and structure–activity and/or structure–metabolism relationships are built and an iterative process may lead to candidates suitable for clinical testing. Major challenges are the simultaneous optimization of lead properties for on/off target, metabolic, toxicological, and pharmaceutical (*e.g.* solubility) issues.¹⁵ Current state-of-the-art for lead optimization includes multi-parameter optimization with different forms of predictive modelling, utility functions, weighed desirability and spread design.¹⁶ Typically compounds with the highest binding affinity or potency are selected for the further optimization of ADME (absorption, distribution, metabolism and elimination), solubility, stability and other pharmaceutical properties with the hope that necessary structural changes will modify potency within acceptable limits. Although these methods avoid rigid rules and have improved success rates in lead optimization cycles, they assume a linear relationship between parameter values and desirability. Thus quantitative predictions and differentiating drug toxicity potential within/across therapeutic classes are challenging at this stage.

The effect of C_{max} on the observed safety and toxicity of 245 drugs was reported earlier.⁹ Fig. 3 of that reference suggests a non-linear relationship of total C_{max} with toxicity, but the relationship was not discussed. Drugs with an average affinity of 10 nM for the target and 30 nM to 1 μ M C_{max} were considered safe. Wenlock has studied the estimation of C_{max} and C_{min} for 215 oral drugs.¹⁷ Ninety percent of the safe drugs had C_{max} below 10.4 μ M. C_{max} at the non-observable adverse effect level (NOAEL) correlated with C_{max} at the maximum dose. A maximum C_{max} of 50 μ M for oral drugs was suggested, but no lower bound was given. This suggests that for a drug-like molecule, C_{max} , C_{min} and pharmacologically effective concentration (C_{eff}) should be minimum. A $C_{max}/C_{min} \leq 30$ is desirable for oral drugs.

IC_{50} , $\log P$, bioavailability, % plasma protein binding (PPB), molecular weight, dose and C_{max} have important roles in drug toxicity. Since oral drugs have molecular weight (~150–750 Da), daily dose (~0.1 mg to 1 g), $\log P$ (~–5 to 6), IC_{50} (on and off target: ~0.001 to 10 μ M), oral bioavailability (0–1), and free C_{max} (0.001 to 10 μ M), predicting toxicity is a complex problem. Thus a novel method for estimating the clinical tox-

icity of drugs and combinations based on physicochemical, PD and PK properties is urgently required.

In this paper, drug toxicity is redefined as the logarithmic sum of scaled biphasic and exponential toxicity contributions from the basic *in vitro*, *in vivo* and physicochemical parameters, thus giving Drug Toxicity Index (DTI). DTI is a significant advancement as a simple model combining toxicity contribution from different parameters is presented for the first time. DTI values are compared between drugs and categories to estimate relative toxicity potentials. PD, PK and physicochemical contributions to DTI offer scope to gain mechanistic insights into drug toxicity. The DTI concept is applied for understanding drug (liver) toxicity, differentiating drug categories, toxic and safe drugs, and identifying potentially toxic drugs. Decision trees like thinking with potential applications of the DTI concept in various drug discovery phases are presented followed by the limitations of the current approach.

2. Methodology

Data on IC_{50} (on/off target), C_{max} at a specified dose, % plasma protein binding, oral bioavailability, molecular weight, ACD $\log D$, $\log P$ for 711 oral drugs were collected and analyzed using a methodology detailed in the ESI PDF file.†

3. Results and discussion

3.1 Pharmacodynamic (PD) toxicity and on target affinity (potency)

Maximizing LE may increase the drug-likeness, reduces dose and C_{max} required for efficacy.¹⁸ Compounds with high affinity ($IC_{50} < 0.01 \mu$ M) are less likely to cause off target toxicities. But they have the potential to cause on target toxicity *e.g.* bupivacaine¹⁹ or toxicity due to activation/inhibition of related targets.¹ Such toxicities are hereby referred to as pharmacodynamic (PD) toxicity. Highly potent drugs (*e.g.* opioids) may cause PD toxicity when administered either alone or in combination with other drugs. This PD toxicity may precipitate due to small changes in the ADME of the highly potent drug induced by factors like the fluctuating patient's physiological parameters or due to the influence of co-administered drugs. For such drugs low doses are usually prescribed, but this doesn't eliminate toxicity risks. For some other highly potent drugs, large doses are required (mostly due to poor bioavailability *e.g.* bosentan), further increasing the PD toxicity risk due to ADME changes. Compounds with low affinity for the primary target generally require the administration of larger doses, achieve high C_{max} , and have higher risk for off target activation and PD toxicity and are generally eliminated early in the drug discovery process. Thus compounds with $IC_{50} > 10$ or $< 0.01 \mu$ M at target are not drug-like. Therefore, PD on target (x_t) toxicity contribution is modelled as a biphasic function of IC_{50} (Fig. 1 and eqn (1)). For molecules with very high binding affinity the toxicity potential is modelled as $1/(10 \times x_t)$,

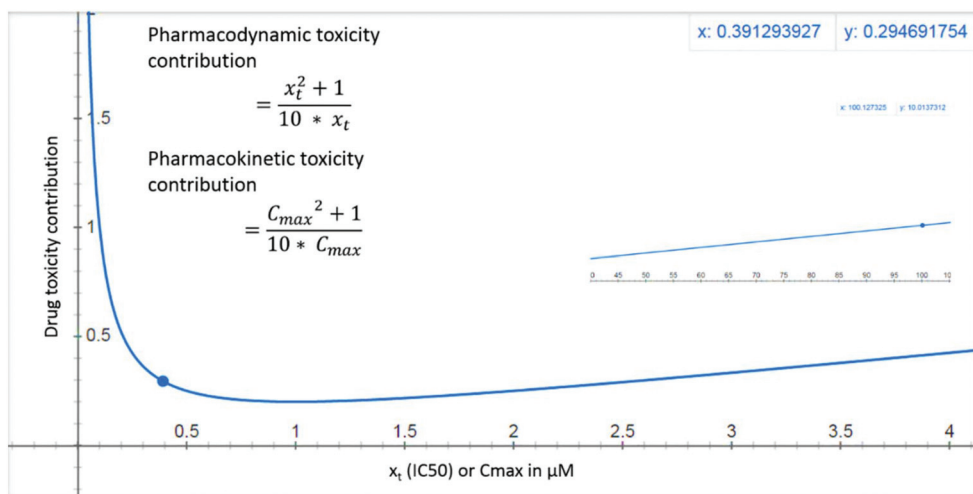


Fig. 1 PD and PK toxicity of drug-like molecules is shown as a function of *in vitro* on target IC_{50} (x_t) or C_{max} (*in vivo* or extrapolated from *in vitro* permeability and metabolism assays) respectively. The inset shows that for a drug with an IC_{50} or C_{max} value of 100 μM , the toxicity reaches a value of 10 units.

while for molecules with low binding affinity the toxicity potential is modelled as $x_t/10$. The combined term $(x_t/10 + 1/10x_t)$ then models the toxicity potential for both high and low affinity molecules.

3.2 Pharmacokinetic (PK) toxicity and maximum plasma concentration (C_{max})

DRUGeff proposed earlier accounts for ADME, but ignores target potency.²⁰ Drug Efficiency Index (DEI), proposed later, doesn't include PK and off target toxicity contributions.²¹ The lack of an upper bound on the unbound concentration (C_{free}) for efficacy seems erroneous. When C_{free} is comparable to dose ($mg\ kg^{-1}$), exposure to the target organ may reach high levels, causing target over-activation, and PD toxicity. As toxicity correlates with C_{max} , using steady state concentrations or V_d may not capture such toxicities. Utilizing the time to C_{max} or AUC for toxicity prediction is complicated by factors like significant differences in frequency, duration of sampling/treatment, and ethical constraints in allowing drug concentrations to reach subtherapeutic levels, among others. Thus to keep the model simple C_{max} was preferred.

In developing the DTI concept the relationship of free C_{max} and toxicity is assumed to be similar to IC_{50} and toxicity (eqn (1), and Fig. 1). C_{max} toxicity contribution is referred to here as "PK toxicity". Drugs with $C_{max} < 0.01\ \mu M$ required for efficacy are more likely to cause PK toxicity due to small variations in ADME *e.g.* bosentan.²² The toxicity potential for such compounds is modelled as $1/(10 \times C_{max})$. Compounds with $C_{max} > 10\ \mu M$, required either for efficacy (due to high IC_{50} , V_d or protein binding, or low clearance), may cause off target activation and PK toxicity and are modelled as $C_{max}/10$. Thus, toxicity risk increases with very high or low IC_{50} and C_{max} for drugs. Eqn (1) models the PD (on target) and PK toxicity contributions assuming that on target IC_{50} (x_t) and C_{max} between 0.01 and 10 μM represent safe drugs.

Pharmacodynamic PD (x_t , on-target) and pharmacokinetic (PK) contributions to toxicity:

$$\begin{aligned} \text{PD } (x_t) \text{ toxicity contribution} &= \left(\frac{x_t^2 + 1}{10 \times x_t} \right) \\ \text{PK toxicity contribution} &= \frac{C_{max}^2 + 1}{10 \times C_{max}} \end{aligned} \quad (1)$$

3.3 Pharmacodynamic (PD) toxicity and off target potency

Compounds with high binding affinity at off targets ($IC_{50} < 10\ \mu M$) are likely to have PD (off target) toxicity,^{1,7,20} while drugs with off target $IC_{50} > 10\ \mu M$ are usually considered further in the drug discovery process. One can model this fact using a negative exponential function, but a purely negative exponential function would always give a PD off target contribution of less than one ($e^{-x} < 1$, for positive x and >1 for negative x values). Thus to keep the off target contributions aligned with the on target contributions an inverse square-root function $10/x^{0.5}$ was added. This gives the off target PD toxicity contribution of $[(e^{-x_{ot}} \times (x_{ot})^{0.5} + 10)/(x_{ot})^{0.5}]$ where x_{ot} is the off target IC_{50} (eqn (2)). Fig. 2 shows the PD (off target) toxicity contribution as a function of off target IC_{50} . For a candidate with off target $IC_{50} = 1, 10$ and $100\ \mu M$, the off target PD contributions are 10.37, 3.16, and ~ 1 respectively. Although negative exponential functions are commonly used in modeling PKPD properties (*e.g.* plasma concentrations, see chapter 2 in ref. 2), inverse square-root functions have been applied rarely. Inverse square-root relationships have been found useful in other fields like permeability (see the section on passive diffusion in chapter 5 of ref. 2), overall connectivity (Randić's inverse-square-root function),²³ and radiotherapy.²⁴ This paper represents the first ever application of an inverse square-root relationship to toxicity prediction.

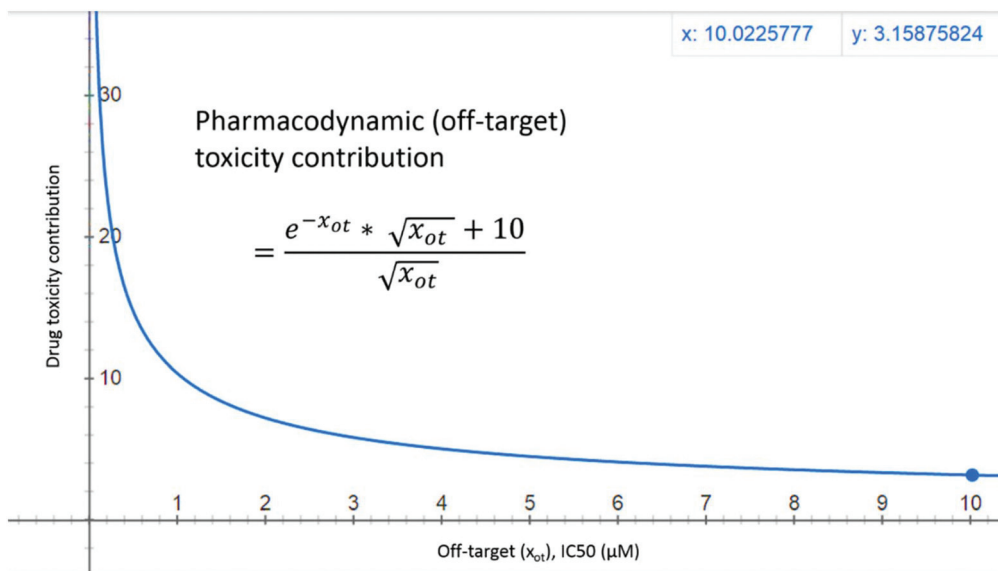


Fig. 2 PD toxicity of drug-like molecules is shown as a function of *in vitro* off-target IC_{50} (x_{ot}) values. For a compound with an off-target IC_{50} value of $1 \mu\text{M}$, the toxicity contribution has a value of 10.37.

Pharmacodynamic PD (x_{ot} , off-target) contributions to toxicity: An off target IC_{50} of $10 \mu\text{M}$ gives a PD toxicity contribution of 3.16.

$$\text{PD } (x_{ot}) \text{ toxicity contributions} = \frac{e^{-x_{ot}} \times \sqrt{x_{ot}} + 10}{\sqrt{x_{ot}}} \quad (2)$$

3.4 Physicochemical contribution to toxicity and log D

A small increase in $\log D$ value causes significantly larger amounts of the drug to easily cross membranes, and accumulate in the liver, brain, and adipose tissues. Thus lipophilic compounds generally have off target activity, extensive metabolism, and hepatotoxicity, and are linked with adverse outcomes.⁹ A $\text{Clog } P$ cut-off of 3 was found useful in categorizing drugs as toxic or non-toxic.²⁵ But a mathematical formulation of the general relationship between these parameters and toxicity has not been proposed.

$\log D$ is the ratio of the amount of drug present in the polar (water/cytoplasm) to that in the non-polar (octanol/cell membrane) compartments at physiological pH 7.4.²⁶ Higher values for this ratio mean that a larger fraction of the drug would get absorbed, transported and/or metabolized. Because toxicity in most cases is likely to be related to this fraction of the drug entering into different organs, it is reasonable to expect that toxicity would correlate better with this fraction (rather than its logarithm). To recalculate this fraction of drug while keeping the physicochemical toxicity contribution aligned with PD contributions, the exponential of $\log D$ was chosen (Fig. 3 and eqn (3)). Considering the RO_3 ,¹¹ and RO_5 ¹⁰ concepts, this contribution was corrected with a $\log D \times e$ term such that a molecule with $\log D = 3, 5$ get values close to 10 and 100 respectively. This gives eqn (3) for estimating physicochemical ($\log D$) toxicity contributions. Polar drugs

($\text{PSA} > 75 \text{ \AA}^2$) inhibit anionic transporter systems with frequencies slightly higher than non-polar compounds (Fig. 3 of ref. 27) and may have similar toxicity potential. This highlights the need for the right balance between polar and non-polar characteristics. $\log D$ contributions to DTI model this aspect as moderately higher contributions for polar compounds (3.09, 8.20, and 13.59 for $-1, -3$ and $-5 \log D$ values respectively). To keep the model simple and aligned with PD contributions, additional overfitting and antilog functions (e.g. $10^x - 10 \times x$) with complex behavior (negative contributions between $x = 0.1371$ and 1.) were avoided.

Physicochemical ($\log D$, pH = 7.4) contribution to toxicity:

$$\text{Physicochemical contribution to toxicity} = (e^{\log D} - (\log D \times e)) \quad (3)$$

3.5 Drug toxicity index (DTI)

Although correlations between toxicity and physicochemical, PK and PD parameters are known, none of the toxicity prediction methods combines them into a single parameter. Thus a method using physicochemical, PD and PK data to estimate toxicity contributions for oral drugs was developed and tested. The individual PD, PK and physicochemical toxicity contributions discussed above estimate the intrinsic toxicity potentials of the drug/xenobiotics. Additionally, these intrinsic toxicity potentials are modulated by the exposure to the candidate toxicant, which in turn is determined by the dose and bioavailability²⁸ (<https://tox.tutor.nlm.nih.gov/03-002.html>). For instance, very high or very low potency or C_{max} values need not necessarily translate into high PD or PK toxicities respectively. These should rather be considered as one of the risk factors for toxicity and a realistic estimation of toxicity contributions requires scaling by oral bioavailability and molar dose. For

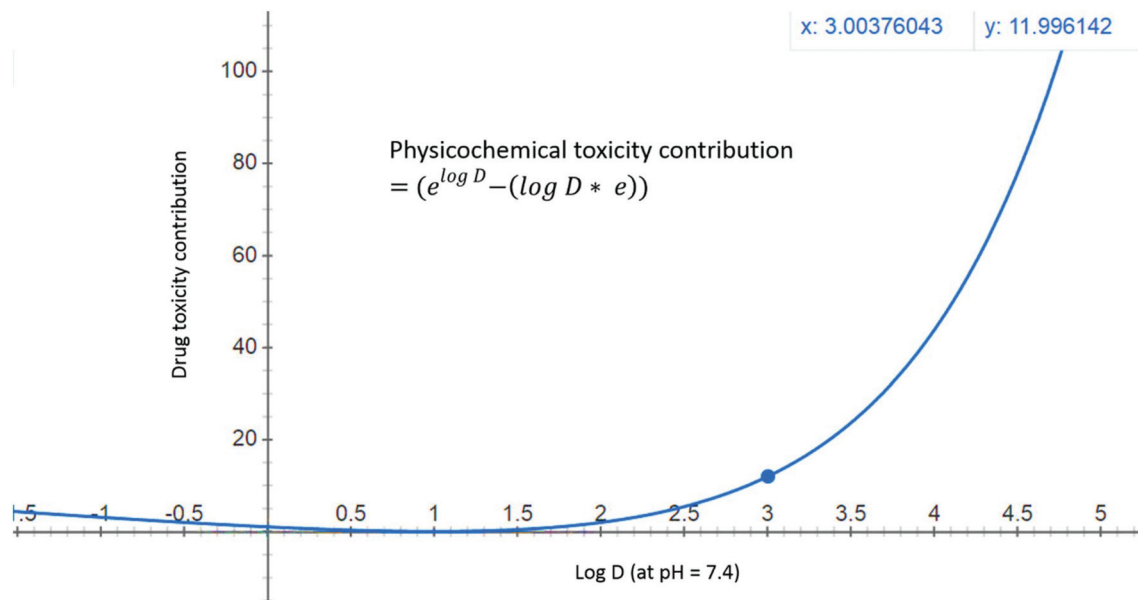


Fig. 3 Physicochemical toxicity of drugs is modelled as an exponential function of $\log D$ (at pH = 7.4) values. A compound with a $\log D$ value of zero will have a toxicity contribution of 1, while compounds with $\log D = 1$ will have zero physicochemical contribution to toxicity, and candidates with $\log D > 3$ are expected to have a high toxicity contribution (>12 units).

example, Calcitriol has the lowest C_{\max} value (0.22 nM) in the dataset considered here and thus has the largest PK toxicity contribution (462 963). Scaling with molar dose (0.0000024 mmol) and oral bioavailability (0.61) gives a realistic PK toxicity contribution of 0.68 (and DTI = -0.17), which is in agreement with the generally safe clinical profile for this drug at recommended doses. Phentolamine with a high C_{\max} (45.8 μM) gives an unscaled PK toxicity contribution of 4.5, suggesting moderate potential for toxicity. Scaling with molar dose and bioavailability reduces this PK toxicity contribution to 0.17, in agreement with the well-known clinically safe profile of this drug.²⁹

Although it is customary to avoid candidates with high off target affinity in drug discovery projects, a low IC_{50} value at the off target alone need not necessarily translate into a clinically relevant drug interaction and toxicity.³⁰ For example, Pergolide Mesylate is a potent CYP2D6 inhibitor and has an unscaled PD (off target) toxicity contribution of 163 but has been found safe in patients with comorbidities (Parkinson's and heart diseases), and co-administered drugs (causes only mild liver enzyme elevations).^{30,31} Scaling with molar dose and oral bioavailability reduces the PD (off target) toxicity contribution to 0.31 (DTI = -0.20), thus correctly predicting a safe clinical profile.

Categorizing drugs based on simple $\log P/D$ cut-off values is an oversimplification of the drug toxicity and lipophilicity relationship. For example, Cefdinir has the lowest $\log D$ value (-5.48) in this dataset and thus should be safe according to simplified rules and cut-off criteria. But Cefdinir has been reported to cause liver toxicity.³² Unscaled $\log D$ toxicity contribution (14.9) predicts high potential for toxicity, whereas scaled contribution reduces to 3.4 (DTI = 0.63), suggesting a moderate toxicity potential in agreement with rare literature

reports on hepatotoxicity. Similarly, Nabilone is very lipophilic ($\log D = 7.25$) and, according to RO3, is expected to have high potential for hepatotoxicity. The physicochemical $\log D$ toxicity contribution for this drug is 1388. But scaling with molar dose and oral bioavailability significantly reduces the toxicity contribution to 0.75 (DTI = -0.06). This is in agreement with the absence of liver injury and safe clinical profile of this drug.^{31,32}

The individual PD, PK and physicochemical ($\log D$) contributions are thus scaled by molar dose and oral bioavailability, summed and the logarithmic sum is defined as the Drug Toxicity Index (DTI, eqn (4)). In contrast to classical TI, DTI does not consider drug toxicity to simply decrease (or increase) with a decrease (or increase) in on target IC_{50} , plasma free C_{\max} , or $\log D$. The proposal of higher response (toxicity) at very low or high doses has been tested, and debated under the Hormetic model of dose–response curves.³³ The DTI concept is different and novel as it expresses the toxicity as a function of different physicochemical, PD and PK parameters of a molecule instead of just dose.

The utility of the DTI concept and values for understanding (i) differences in WHO ATC drug classes, (ii) classification of drugs for drug-induced liver injury (DILI) concern, (iii) relative toxicity potential within a therapeutic area, (iv) physicochemical, PK and PD toxicity contributions, (v) identifying potentially toxic drugs, and (vi) recently approved drugs is discussed below. A decision tree approach is presented describing the potential use of the DTI concept in a typical drug discovery project followed by limitations and scope for improvements.

The Drug Toxicity Index (DTI) is defined as the logarithmic sum of physicochemical ($\log D$), pharmacokinetic (PK) and pharmacodynamic (PD) contributions to toxicity scaled by

absolute oral bioavailability (OBA), and molar dose (MD). Here, X_t is potency (IC_{50} or MIC) at the target protein, C_{max} is the maximum unbound plasma drug concentration, X_{ot} is the binding affinity at the off-target (CYP450 isoforms, hERG channel and BSEP transporter), and $\log D$ is the ACD $\log D$ at pH = 7.4. PD and PK contributions to DTI are defined only when X_t , C_{max} and X_{ot} are greater than zero.

$$DTI = \log \left(MD \times OBA \sum \text{contributions to toxicity} \right)$$

$$DTI = \log [MD \times OBA (PD \text{ (on and off target)} + PK + \text{Log } D \text{ contributions to toxicity})]$$

$$DTI = \log \left(OBA \times MD \left[\left(\frac{x_t^2 + 1}{10 \times x_t} \right) + \left(\frac{C_{max}^2 + 1}{10 \times C_{max}} \right) + \left(\frac{e^{-x_{ot}} \times \sqrt{x_{ot}} + 10}{\sqrt{x_{ot}}} \right) \right] + [OBA \times MD \times (e^{\log D} - (\log D \times e))] \right)$$

3.6 DTI and WHO ATC drug classes

The DTI for WHO ATC drug classes (711 drugs) was analyzed. DTI values (for full data see Excel files in the ESI†) for antiviral, antibiotic, antineoplastic (anticancer), and withdrawn drugs are higher than those for other drug classes (Fig. 4). As small datasets are used, the average DTI for antiparasitic, systemic and hormonal, and blood categories should be interpreted cautiously.

Fig. 5 shows scatter plot for DTI and LEI. Molecules with $LEI > 0.3$ and $LLE > 5$ are generally considered good candidates for drug development.¹⁸ Withdrawn drugs are expected to have $LEI < 0.3$, $LLE < 5$. Of 88 withdrawn drugs, 53% (Table 1) satisfy LEI criteria.¹⁸ DTI performs similarly (50%, at 0.67 cut-off). Only 43, 24 and 43% withdrawn drugs have dose > 100 mg and $\log D > 3$. On average 36 and 27% approved drugs satisfy RO2 criteria (73% for systemic and hormonal drugs) and 13 and 20% satisfy $\log D$ criteria (see Table S11†). Although DTI and LEI (LLE, 67%) have similar performances for withdrawn drugs, DTI offers scope for interpreting toxicity contributions.

DTI performs better than LEI, LLE, and the doses for Musculoskeletal, Antiparasitic, Metabolic and GIT, CNS,

Systemic and Hormonal, Gento-Urinary and Sex, Respiratory, Blood, and Cardiovascular categories (Table 1). DTI one cut-off performs better than other parameters for Cardiovascular, Respiratory, Blood, Gento-Urinary and Sex, Systemic and Hormonal categories. For Antiparasitic, Antibiotic, Antiviral, and withdrawn drug categories the average DTI cut-off value performs better than other parameters. The use of the average

cut-off value is justified considering the known higher toxicity for these drug classes. Across all categories the average percentage of drugs satisfying parameter cut-off values shows that DTI is the best parameter (69%) considering its interpretability. $\log D$ gives minor improvement but has limited scope for understanding toxicity contributions. PK, physicochemical, PD on and off target, contributions to DTI also show excellent performance at identifying safe drugs (see ESI Tables S12 and S13† for the averages). Contributions to DTI have moderate performance for withdrawn drugs but DTI itself performs better. DTI is better than exposure based TI at estimating safety. Only 26, 20 and 23% drugs across different WHO ATC drug categories satisfied exposure based TI of ≥ 30 for CYP450, hERG and BSEP inhibition (Table S11†). At efficacious doses the free C_{max} is generally within 0.5–2 times that of IC_{50} .³⁴ Against this criterion DTI is 6.6 times better at estimating drug safety. Within the dataset considered here, $\sim 27\%$ of drugs have negative DTI values (see Table S12†). These should not be interpreted as having zero or no potential for toxicity. Rather negative DTI values indicate a total contribution to the toxicity of less than 1 (eqn (4), $DTI = \log(MD \times OBA \sum \text{toxicity contributions})$).

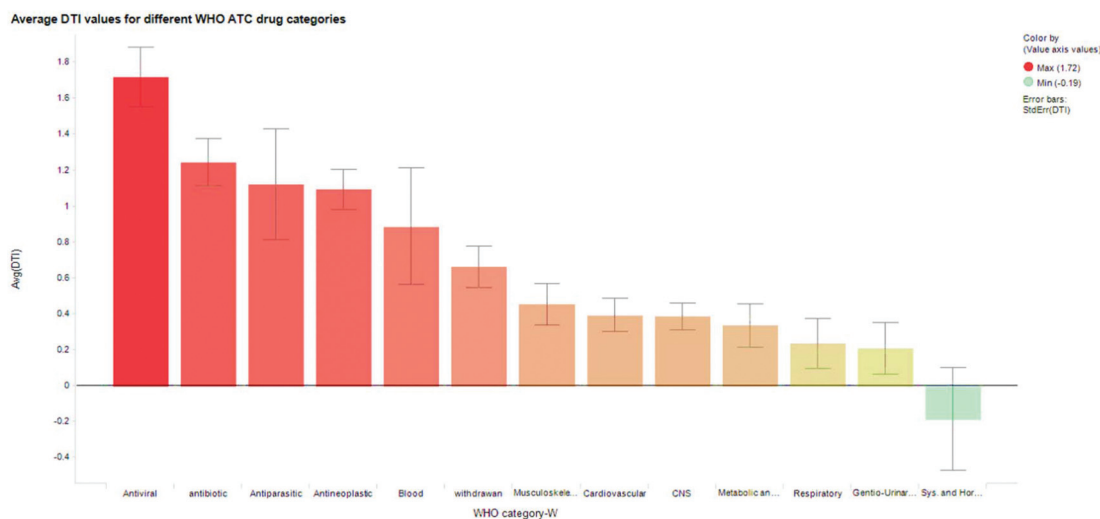


Fig. 4 Average DTI values for different WHO ATC classes of drugs and withdrawn drugs considered in this work. Error bars represent standard errors.

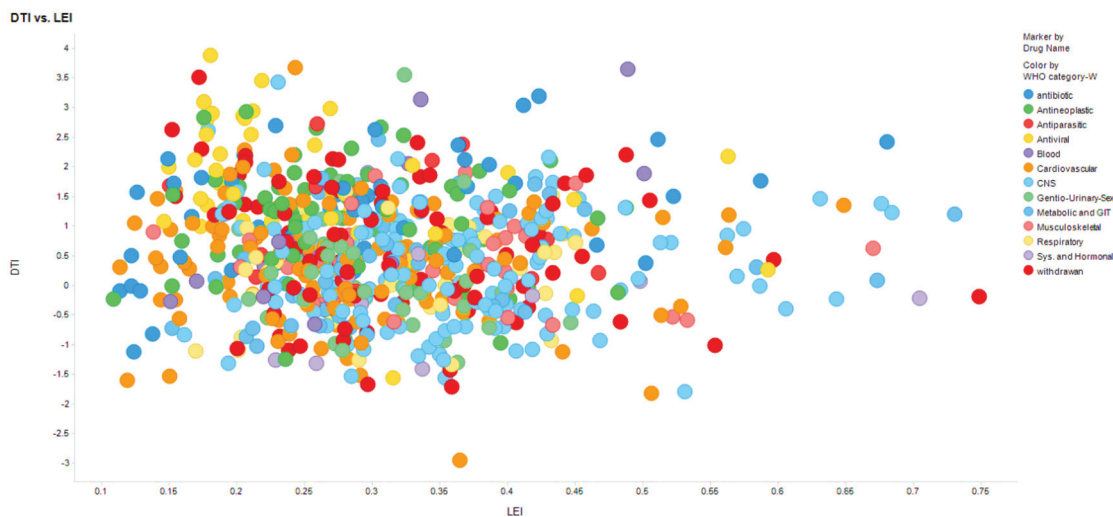


Fig. 5 Scatter plots of LEI and DTI values for different WHO ATC drug classes show that LEI and DTI values have no correlation.

Table 1 Percentage of drugs satisfying the different efficiency (LEI, LLE), dose and physicochemical parameter criteria recently recommended. Comparison of average DTI values against these widely used parameters for lead optimization suggests that a DTI criterion of <1 is better in many cases and equal in others (except for antineoplastic and antiviral drugs). For PD on target/off target, PK and physicochemical contributions to DTI, the respective average values have been used as cut-off criteria (see ESI)

WHO ATC drug category	LEI > 0.3	LLE > 5	PD on target	PK	Physico-chemical Log <i>D</i>	PD off target			DTI < 1	Exposure based TI, % of IC ₅₀ /C _{max} values between 0.5–2	Dose < 100 mg	ACD log <i>D</i> < 3	log <i>P</i> < 3
						CYP450	hERG	BSEP					
Musculoskeletal	58	24	71	89	84	84	95	89	^a 84	13	34	97	55
Antiparasitic	45	18	73	82	91	100	91	82	^a 55	0	27	82	73
Antibiotic	36	54	94	86	80	84	86	76	^b 54	0	0	90	80
Metabolic and GIT	46	63	83	96	83	83	83	90	77	17	69	90	79
CNS	69	44	94	86	84	90	85	90	73	21	68	84	62
Systemic and hormonal	45	55	82	91	91	82	100	82	^c 91	9	82	82	64
Withdrawn	53	67	18	14	01	11	18	9	^c 50	0	43	24	43
Gentio-Urinary and Sex	40	38	98	98	90	90	88	90	^d 85	20	80	48	38
Antiviral	24	61	80	73	93	76	85	76	^d 51	7	17	56	39
Respiratory	61	31	89	78	83	89	83	92	^d 83	17	75	75	47
Blood	33	27	93	87	87	87	80	87	73	13	53	73	40
Cardiovascular	37	31	86	98	89	86	89	88	74	11	63	70	48
Antineoplastic	35	56	82	85	90	85	81	87	43	10	35	62	50
Average across therapeutic area	45	44	80	82	80	81	82	80	^d 69	11	50	72	55

The following average DTI values for the respective therapeutic areas were used as the cut-off: ^a1.12, ^b1.4, ^d1.72. ^cFor withdrawn category criteria of LEI < 0.3, LLE < 5, and DTI > 0.67 (avg for withdrawn), ACD log *D* and log *P* > 3 were used. For exposure based TI, the criteria were IC₅₀/C_{max} values between 0.5 and 2. For classical TI (CYP450, hERG, BSEP) the criterion was ≥ 30 . None of the antibiotic classes of drugs considered here had a dose < 100 mg. The average values used as the cut-off for PD on/off target, PK and physicochemical parameters are given in Tables S12–13.

These can be interpreted (cautiously) as having lower toxicity potential (see sections 3.8.1 and 3.13), arising from different physicochemical (Log *D*), PD (on and off target) and PK toxicity contributions (discussed in sections 3.1–3.4) considered for calculating the DTI. It is worth noting that, although DTI values can be negative, the individual and total toxicity contributions (as defined in eqn (1)–(3)) can be either positive or zero, but not negative.

DTI differentiates WHO ATC drug categories (ANOVA, Table S1†). Antibiotic and antiparasitic drugs are similar in

their toxicity potential (similar DTI values avg; 1.12, 1.24, max; 3.2, 2.74, min; –1.12, –0.08 respectively, Fig. 4). The withdrawn drugs (avg; 0.67) are differentiated from nine drug classes (antibiotic, metabolic and GIT, nervous system (CNS), systemic and hormonal, gentio-urinary and sex, antiviral, respiratory, cardiovascular and antineoplastic). LEI, molar dose, dose, OBA, IC₅₀, free C_{max}, PPB, ACD log *D*, and log *P* differentiate only 4, 6, 6, 4, 5, 6, 4, 5 and 3 drug category pairs respectively (Tables S11–28†). Antiviral drugs have the highest average DTI, suggesting high toxicity. Although ~64% correct differen-

tiation among WHO ATC drugs may sound ordinary, it is better than random prediction and has been attempted for the first time. Thus differentiating the toxicity potential between the different WHO ATC drug classes is not possible with simple PD, PK and physicochemical parameters. This is possible with the DTI concept as it is able to differentiate withdrawn drugs from 9 (out of 12; 75%) other approved drug classes.

3.7 US-FDA's liver toxicity knowledge base (LTKB) and DTI

The utility of DTI to correctly classify drugs with drug induced liver injury (DILI) concern was tested with the US-FDA's LTKB containing 1039 drugs.¹³ DTI could be calculated for 392 drugs. Eighty-five, 233 and 74 had most-DILI-concern, less-DILI-concern and no-DILI-concern. This proportion of drugs with different types of concern annotations is similar to that present in the larger database with 1039 drugs. The area under the receiver operating characteristic curve (AUC, Table 2) was generated by considering different DTI cut-off values. The AUC was calculated using two types of categorization (see Table 2 and ESI Excel file†).

DTI gave a good AUC of 0.72 for all 392 drugs (categorization A), while for CNS and cardiovascular drugs excellent classification was achieved (AUC, 0.84). Respiratory and antibiotic drugs were classified with good AUC, 0.72 and 0.69. Categorization B gave an excellent AUC of 0.89 and 0.91 for antiviral and musculoskeletal drugs respectively. These drug categories are more likely to cause liver injury and even a low hepatotoxicity risk should be considered seriously. These results agree with the toxicity trends found earlier.¹³

The AUC for the ROC curves estimated with categorization A correlates well ($r = 0.77$) with the number of drugs within each therapeutic area. This suggests that predictions can be improved with the inclusion of additional drugs within each therapeutic area if the missing data (*in vitro/in vivo*) become available. Lack of correlation for categorization B ($r = -0.09$) suggests that this calcification is specific to antiviral and musculoskeletal drugs. But due to a small number of drugs considered here, data for additional drugs would be required

to confirm this suggestion. DTI's ability to predict the potential for hepatotoxicity and other forms of toxicity is discussed in the following sections.

3.8 Relative toxicity potential within a therapeutic area

Assessing the relative drug efficacy and toxicity is important for medicinal chemists, pharmacologists, pharmacists, physicians, and regulators.³⁵ As multiple treatment options are available for many diseases, relative toxicity assessment is necessary. Difficulty in direct clinical trial comparisons forces network meta-analyses for estimating the relative toxicity.³⁶ The application of DTI for estimating network meta-analysis outcomes is discussed (Table 3).

3.8.1 NSAIDs and DTI. DTI for a set of NSAIDs were compared with the clinical risk of gastrointestinal complications and gave a good correlation.³⁷ Table S2† shows that the relative risk of gastrointestinal toxicity correlates well with DTI (r ; 0.85, Diclofenac and Piroxicam excluded). The Diclofenac DTI is negative because all off-target effects are not considered. Addition of COX-1 IC₅₀ (0.003 μM) and PD toxicity contribution (5.06) makes DTI 0.84, explaining Diclofenac's toxicity. Piroxicam's toxicity has been linked to reactive metabolite formation (see "section 3.13 Limitations of the current DTI methodology"). Thus, DTI gives reliable estimates of relative NSAID toxicity.

3.8.2 Antibiotics and DTI. The relative efficacy and adverse effects of 5 antibiotics (Ciprofloxacin, Norfloxacin, Amoxicillin-Clavulanate, Gatifloxacin, and Trimethoprim/Sulfamethoxazole) give a moderate correlation of 0.64 with DTI (Table S3†).³⁸ DTI correlates with the relative adverse effects (except for Gatifloxacin). For the trimethoprim/sulfamethoxazole combination, a modification of the Hill function to predict the response to the combination has been used.³⁹ A larger clinical dataset may confirm this utility of DTI.

3.8.3 CNS drugs and DTI. Leucht *et al.* performed a network meta-analysis to estimate the odds ratio (OR) for all-cause discontinuation, sedation, weight gain, QTc prolongation and extrapyramidal effects for fifteen antipsychotics.⁴⁰ DTI correlates well with all-cause discontinuation OR

Table 2 AUC values for the ROC curves calculated using different DTI cut-off values for the classification of toxic and safe drugs reported in the US-FDA LTKB database

WHO ATC drug category	Number of drugs	AUC value for ROC curve	
		(A) Toxic = Most-DILI concern safe = No-DILI + Less-DILI concern	(B) Toxic = Most-DILI + Less-DILI concern safe = No-DILI concern
Drugs used from LTKB	392	0.72	0.62
Gentio-urinary & sex	20	0.61	0.64
Antiviral	19	0.53	0.89
Respiratory	16	0.73	0.46
Metabolic and GIT	26	0.52	0.39
Musculoskeletal	24	0.48	0.91
CNS	100	0.84	0.64
Cardiovascular	76	0.84	0.58
Antibiotics	34	0.69	NA ^a

^a None of the drugs in the antibiotics class for which DTI could be calculated had a No-DILI-concern annotation. For other drug categories, the no. of drugs with LTKB annotations and DTI values was below 10 and thus these were not analysed.

Table 3 Summary of the correlations of DTI values with clinical outcome measures reported in various network meta-analyses of clinical trials performed to analyse the relative safety and efficacy of different treatment options in following therapeutic areas

Therapeutic area	Number of drugs	Clinical toxicity/efficacy outcome (odds ratio)	Correlation of DTI values with clinical toxicity/efficacy outcomes
Musculo-skeletal (NSAIDs)	9	GIT side effects	0.85
Antibiotics	5	Overall adverse effects	0.64
CNS-antipsychotics	13	All cause discontinuation	0.75
CNS-antidepressants	9	Response rate odds ratio	-0.72
Metabolic and GIT-antidiabetics	15	Hypoglycemia	-0.68
Gentio-urinary and sex	5	Frequency of adverse effects	0.84
Antiviral	10	Discontinuation due to AE	0.63
		Viral suppression	-0.86
Blood and blood forming agents	5	All types of strokes	0.69
		Myocardial infraction	0.86

(13 drugs, r ; 0.75, Table S4†). Clozapine has been associated with many liver injuries. Clozapine's DTI captures liver toxicity *via* the contribution of $\log D$, 22.5. A re-evaluation of Clozapine cardiotoxicity has been suggested (scaled hERG DTI contribution, 21.8).⁴¹

Relative tolerability was reported for ten antidepressants using OR for the response, remission rates, and withdrawal rates due to adverse effects.⁴² The OR for response rates correlated well (negatively, r ; -0.72) with DTI (Table S5†). Agomelatine showed the highest OR for the response rate (low DTI). Thus DTI can predict the relative safety of CNS drugs.

3.8.4 Antidiabetic drugs and DTI. The relative hypoglycemia risk and body weight change estimated for 15 antidiabetic drugs correlated (negatively, r ; -0.68, -0.60, Table S6†) with DTI.⁴³ Relative risk for UTI and genital tract infections showed weak and decent correlations respectively with DTI. The significance of the negative correlations with adverse effects requires further investigation.

3.8.5 Gentio-urinary and sex drugs and DTI. The relative efficacy and adverse event frequencies reported for seven PDE5 inhibitors show good correlation (r ; 0.84, Table S7†) with DTI.⁴⁴ An earlier study comparing Tadalafil, Vardenafil, and Avanafil did not find a significant difference in their safety profiles⁴⁵ and thus these results should be interpreted cautiously.

3.8.6 Antiviral drugs and DTI. Patel *et al.* estimated OR and percentage values for the efficacy, safety and discontinuation due to adverse effects (Table S8†) for Dolutegravir and seven other drugs.⁴⁶ OR for efficacy parameters and triglyceride levels have good correlation with DTI. The correlation of OR for discontinuation due to AE with DTI is 0.63. Poor correlations for combined-DTI values can be due to different and undisclosed doses used in different combinations. This can significantly alter the individual (PD, physicochemical and PK) contributions and the DTI.

3.8.7 Blood and blood forming (anticoagulant) drugs and DTI. OR for myocardial infraction and all-cause-strokes for acetylsalicylic acid plus Clopidogrel (ASA + C), Apixaban, Rivaroxaban, Dabigatran, and Edoxaban against Warfarin as a reference have excellent correlations with DTI (Table S10†).⁴⁷ These suggest that DTI might be useful in estimating the relative toxicity of antithrombotics.

Although the relative toxicity/efficacy for cardiovascular,⁴⁸ antiemetic,⁴⁹ and antiulcer drugs is available, such studies often compare drug sub-classes (*e.g.* beta-blocker, ACE, proton-pump inhibitors, H2-receptor antagonists, and 5HT-agonists). Molecule specific clinical risks associated with these and anti-parasitic drugs were not found in the literature. Thus the use of the DTI concept in estimating relative drug safety/efficacy for these categories remains to be tested.

3.9 Physicochemical, PK and PD toxicity contributions

Efficacy (52%) and safety (24%) are top reasons for Phase II drug discovery failures.⁵⁰ Thus estimating toxicity potential in Phase I studies is a high priority. The following examples demonstrate the utility of DTI in estimating toxicity potential from Phase I PK, PD, and physicochemical data.

Reactive metabolite formation has been proposed for Troglitazone toxicity.⁵¹ DTI estimates a PD on target contribution of 0.08 (IC_{50} at PPAR γ , 0.98 μ M), and PD (off target) contributions of 2.36, 0, and 2.51 (CYP450, hERG, BSEP, see Table 4). For a dose of 400 mg (MD, 0.91) and an oral bio-availability of 45%, the free C_{max} value is 0.036 μ M (PPB, 99%) and the PK toxicity contribution is 1.12.⁵² The largest contribution is from the physicochemical parameter ($\log D$, 3.65) of 11.64, making DTI 1.25. Calculations for Pioglitazone (non-hepatotoxic) give DTI 0.45 (Table 4). Reducing the Troglitazone dose to 200 mg decreases these contributions to 0.04, 1.18, 0.0, 1.25, 1.00 and 5.82, marginally decreasing DTI (0.97) and suggesting a similar toxicity risk. Thus reducing the dose doesn't reduce the toxicity potential proportionally. Rosiglitazone cardiotoxicity is most likely the result of K_{ATP} or kinase inhibition.⁵³ Thus, Rosiglitazone toxicity cannot be predicted using the current DTI formulation (-0.17), as it excludes other off target effects.

Terfenadine has PD on target (4.24), off target (1.49, hERG; 4.05, CYP450; 0.0, BSEP), PK (0.08) and physicochemical (6.91) toxicity contributions giving DTI 1.22 (Table 4). The physicochemical contribution is the largest; hERG and CYP450 inhibition contributes significantly to correct toxicity prediction. Fexofenadine (non-toxic)⁵⁴ has PD on target (0.56), PD off target (0.10, 0.0, 0.0), PK (0.05) and physicochemical (0.01) toxicity contributions with DTI -0.15 (Table 4).

Table 4 Individual toxicity contributions and DTI for drugs discussed in the main text

Drug	Toxicity contributions						DTI
	PD on target	PD off target			PK	Log <i>D</i> (physicochemical)	
		CYP450	hERG	BSEP			
Troglitazone (400 mg)	0.08	2.36	0.00	2.51	1.12	11.64	1.25
Troglitazone (200 mg)	0.04	0.00	1.25	1.00	1.18	5.82	0.97
Pioglitazone	0.08	0.34	0.00	1.99	0.28	0.14	0.45
Terfenadine	4.24	4.05	1.49	0.00	0.08	6.91	1.22
Fexofenadine	0.56	0.00	0.10	0.00	0.05	0.01	-0.15
Telotristat ethyl	0.01	0.96	0.99	0.00	28.99	3.54	1.54
Enasidenib	0.02	0.00	0.40	0.00	0.03	2.84	0.52
Brivaracetam	7.56	0.00	0.00	0.00	12.79	0.50	1.32
Cariprazine	1.25	0.00	0.00	0.00	15.24	0.74	1.24

3.10 Identifying potentially toxic and relatively safe drugs

The LiverTox website categorizes drugs based on reports of clinically apparent liver injuries.³¹ This categorization along with the literature reports was used to assess predictions made with DTI. Since the categorizations in LiverTox are still in active development and all drugs do not have a likelihood score assignment or equal experience in the clinic, the assessment of toxicity potentials presented here is of qualitative and suggestive nature and is meant for retrospective analysis (<https://livertox.nih.gov/DrugCategory.html> accessed on 11/11/2018). DTI values and different PD, PK and Log *D* contributions are with liver and other forms of toxicities. Representative examples with high, low and average DTI values are discussed below for musculoskeletal drugs.

Musculoskeletal category (DTI max, 1.85; min, -0.66; avg, 0.41), niflumic acid has the highest DTI value, but toxicity information is limited. Mesalazine (DTI, 1.73) has been associated with many cases of hepatotoxicity and has a likelihood score of C. Probenecid and Eperisone (DTI, 1.62) have been associated with hypersensitivity reactions and Eperisone's safety is questionable.⁵⁵ Indomethacin (DTI, 0.46) has been associated with rare idiosyncratic hepatotoxicity and has a likelihood score of C. Meloxicam (DTI, -0.32) is generally found to be non-hepatotoxic and shows only rare serum enzyme elevations. Teriflunomide has a low DTI (-0.61), does not show liver injuries and show only mild serum enzyme elevations. With a likelihood score of D, it nevertheless has a black-boxed warning due to less clinical experience. Addition of off target (COX-1, IC₅₀, 140 μM)¹ PD contribution for Valdecoxib increases the DTI value to 0.53. With a toxicity cut-off >1, the musculoskeletal drugs were predicted with high accuracy, sensitivity, and specificity (0.84). A comparison with withdrawn drugs (DTI avg, 0.68, *n* = 9) shows that approved drugs tend to have lower DTI. A similar discussion on other drug categories is given in the ESI PDF file.†

3.11 Recently approved drugs

Telotristat ethyl has PD (on and off target) and physicochemical contributions to toxicity of 0.01, 0.96 (CYP450), 0.99 (hERG), 0.0 (BSEP) and 3.54 (log *D*), a PK contribution of

28.99, and DTI 1.54 (Table 4). This agrees with the safety data (US-FDA approval package) where 8.4% of the patients experience serum enzyme elevations, and 66.9% showed GIT side effects. Enasidenib, approved August 2017, has PD (on and off target), PK and physicochemical toxicity contributions of 0.02, 0.0 (CYP450), 0.40 (hERG), 0.0 (BSEP), 0.03 and 2.84 respectively and DTI 0.52 (Table 4). The US-FDA label warns of differentiation syndrome and embryo-fetal toxicity. Off target inhibition (UGT1A1, not considered here) has been suspected for this toxicity.⁵⁶

Brivaracetam, approved February 2018, has PD (on and off target), PK and physicochemical toxicity contributions of 7.56, 0.0 (CYP450, hERG, BSEP), 12.79 and 0.50 respectively and DTI 1.32 (Table 4). The US-FDA label mentions hypersensitivity, hepatic injury and adverse neurological effects, but the mechanisms are unclear. PD (off target) and physicochemical contributions to DTI correctly predict a lack of hepatotoxicity. Cariprazine, approved September 2015, has PD (on and off target), PK and physicochemical toxicity contributions 1.25, 0.0 (CYP450, hERG, BSEP), 15.23 and 0.74 respectively and DTI 1.24 (Table 4). Increased death risk, weight gain, orthostatic hypotension, hypersensitivity, cerebrovascular, tardive dyskinesia, and bipolar mania were observed during clinical trials (US-FDA approval package). Thus the toxicity of Cariprazine is mostly related to complex pharmacological interactions; CNS disorder management and DTI capture these toxicities.

3.12 Decision trees for using DTI in a typical drug discovery project

Since the drug discovery usually starts with a large number of potential candidates it is important throughout this process to filter less efficacious or toxic candidates using a set of well understood rules or guidelines. At present none of the available rules or guidelines can be used from the start until the end *i.e.* each set of rules has its applicability domain within different drug discovery phases. For example, Lipinski's rule of five, ligand/structure based methods, and LEI are extensively used along with others in virtual screening and lead to optimization stages. But these have a limited role once a candidate enters preclinical or clinical stages where PK parameters and

efficacy/toxicity endpoints are utilized for screening (on smaller sets). A comprehensive approach, like the DTI presented here, that allows a parameter's efficacy/toxicity prediction value to be used beyond the original application domain is likely to increase accuracy and offer scope for deeper insights as one moves ahead in drug discovery phases. A flowchart like decision tree summarizing the potential application of the DTI concept in a typical lead identification/optimization to Phase I drug discovery project is given in Fig. 6 and its caption.

For example, let us consider Troglitazone (discussed in section 3.9) as a potential candidate for development. One begins with an unscaled $\text{Log } D$ (3.65) contribution to DTI, which is 28.55. Thus if a highly bioavailable formulation for Troglitazone ($\text{OBA} \approx 1$) is targeted, an oral dose of <20 mg (0.05 mmol) would be required to keep $\text{Log } D$'s contribution below 1.3. Higher doses *e.g.* 100 mg (0.23 mmol) lead to higher $\text{Log } D$ contributions, effectively reducing the flexibility

and tolerance available with other parameter ranges (since DTI should be <1). Then the IC_{50} (on target; $0.98 \mu\text{M}$) gives an unscaled PD (on target) contribution of 0.2. Thus, as per the DTI concept, Troglitazone's moderate potency is not a problem in terms of its toxicity potential. But its unscaled PD off target contributions would add a total of 11.93. This also constrains the dose to <20 mg to keep DTI < 1 . Thus the design question to move such a candidate into the next (pre-clinical) phase is: can we achieve good exposure (C_{free}) with these doses and OBA constraints and maintain PK contribution to DTI below 10 and ideally DTI below average for the therapeutic area? For Troglitazone the answer is no because larger doses are required for good exposure and a therapeutic effect, thus increasing the molar dose and PD, PK and physicochemical contributions to DTI and toxicity potential. Thus, using the DTI concept, Troglitazone can be dropped from further development at different stages if other compounds

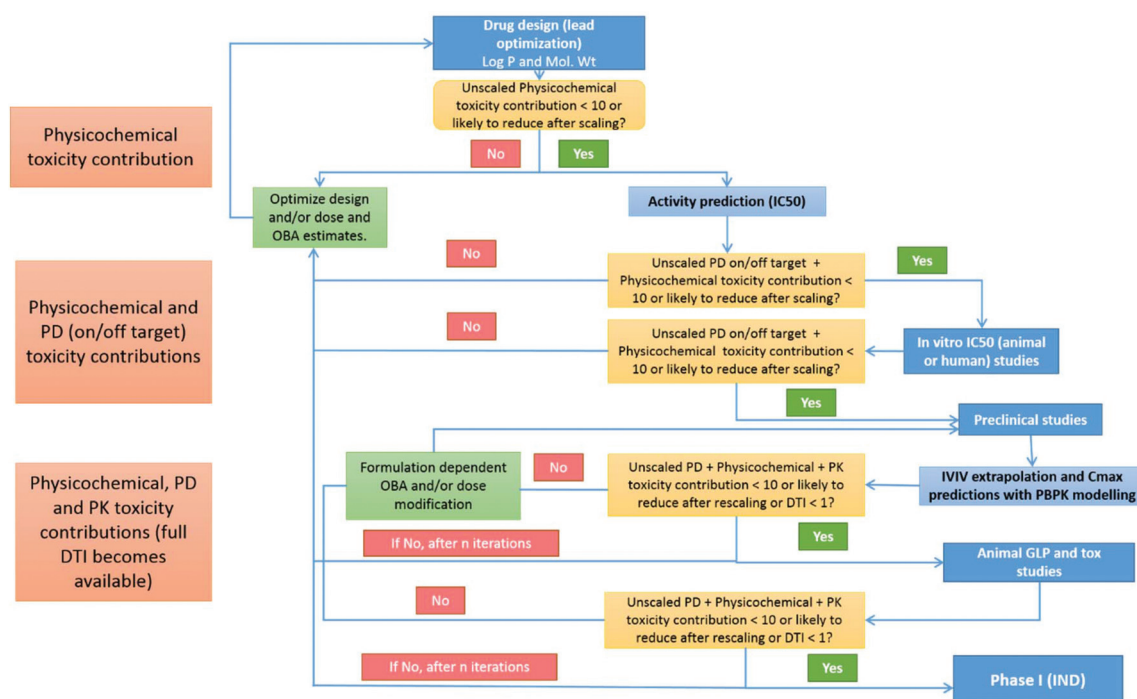


Fig. 6 A decision tree proposed for potential application of Physicochemical, PD (on/off target) and PK toxicity contributions and the DTI concept in a typical small molecule drug discovery project. Ideally one would start with the structure and $\text{log } D/P$ values for potential candidates obtained using different hit/lead identification procedures. Candidates with physicochemical toxicity contributions > 10 can be filtered out unless scaling with molar dose and/or oral bioavailability is likely to reduce them significantly. *In silico* prediction of selected compounds then gives an estimate of the PD (on/off target) toxicity contributions. Candidates with physicochemical + PD toxicity contributions > 10 can be filtered out unless scaling is likely to reduce them. Selected candidates can be tested *in vitro* and PD toxicity contributions re-estimated and candidates again filtered on the same criteria. *In vitro*–*in vivo* (IVIV) extrapolations for animal or human exposure of selected candidates using PBPK modelling will give estimates of C_{max} (dose and oral bioavailability) and PK toxicity contributions. At this stage the first total DTI estimates will become available and candidates with DTI > 1 should be re-evaluated for progression to the next phase. Follow-up toxicity and efficacy studies may be useful based on a mechanistic understanding derived from toxicity contributions, and DTI. Preclinical studies performed on selected candidates will allow first experimental measurements of DTI in animals. Candidates with DTI < 0.5 can be considered for Phase I, while those with DTI $\geq 1 \pm 0.5$ should be carefully reassessed using general FDA guidelines, therapeutic area specific knowledge and risk/benefit ratio for patients. DTI and its physicochemical, PD and PK contributions are intended to complement existing state-of-the-art in drug discovery and any disagreement between DTI and other approaches should be assessed carefully considering the limitations of DTI (and other methods) before making a final decision. Blue filled boxes on the right indicate experimental studies. Light orange color filled boxes indicate decision points based on toxicity contributions (<10) or DTI values (<1). Light green filled boxes on the left indicate the requirement of additional cycle(s) of design and optimizations. Light red boxes on the left indicate the type of data required and toxicity contributions that can be estimated at different time points within a drug discovery project.

(with a similar/lower molecular weight) have lower lipophilicity, similar on target potency, lower off target potencies ($>10 \mu\text{M}$), good (predicted) exposures coupled with lower estimates for the PD, PK contributions and DTI.

Based on the correlations discussed in “section 3.8 Relative toxicity potential within a therapeutic area”, applications of the DTI concept beyond Phase 1 are clear and a flowchart like decision tree is proposed in Fig. 7 and its caption.

3.13 Limitations of the current DTI methodology

Current DTI formulation has some limitations which should be considered during interpretation. DTI considers hepatotoxicity arising only from lipophilicity, CYP450 and BSEP inhibition. Interaction with non-CYP450 metabolic enzymes might lead to drug–drug interactions and toxicity.⁵⁷ Additional off target⁷ contributions may improve the predictive performance (as demonstrated for diclofenac, section 3.8.1), especially for drugs with negative DTI values. The PK contribution may indirectly compensate, but this needs to be tested. C_{free} at the target has been suggested for *in vivo* efficacy but such information is not available for the majority of drugs.³⁴ Oral bioavailability data for

59 (~8%) compounds were not from human studies. This is unlikely to affect the major findings, but the prediction accuracy can improve with human data for these compounds.

Active and reactive metabolites play important roles in drug efficacy and toxicity. But accounting for this is non-trivial since information on the extent of metabolism, identity, and exposure to metabolites, and their on and off target effects is unknown. Currently, DTI doesn't include reactive metabolite contributions to toxicity.

The DTI concept was applied to orally administered drugs since C_{max} and bioavailability of intravenously, arterially, intrathecally, and intracranially administered drugs may depend on the infusion rate and the site of administration. Topical, inhalational, rectal, vaginal, intramuscular, peritoneal, and other routes of administrations also present difficulty and thus were not considered. Since the PK and PD data for investigational compounds is generally not released before approval, the DTI concept could not be tested for them. Nevertheless, DTI and its PK, PD and physicochemical contributions are able to explain the observed clinical safety/toxicity profiles of recently approved drugs.

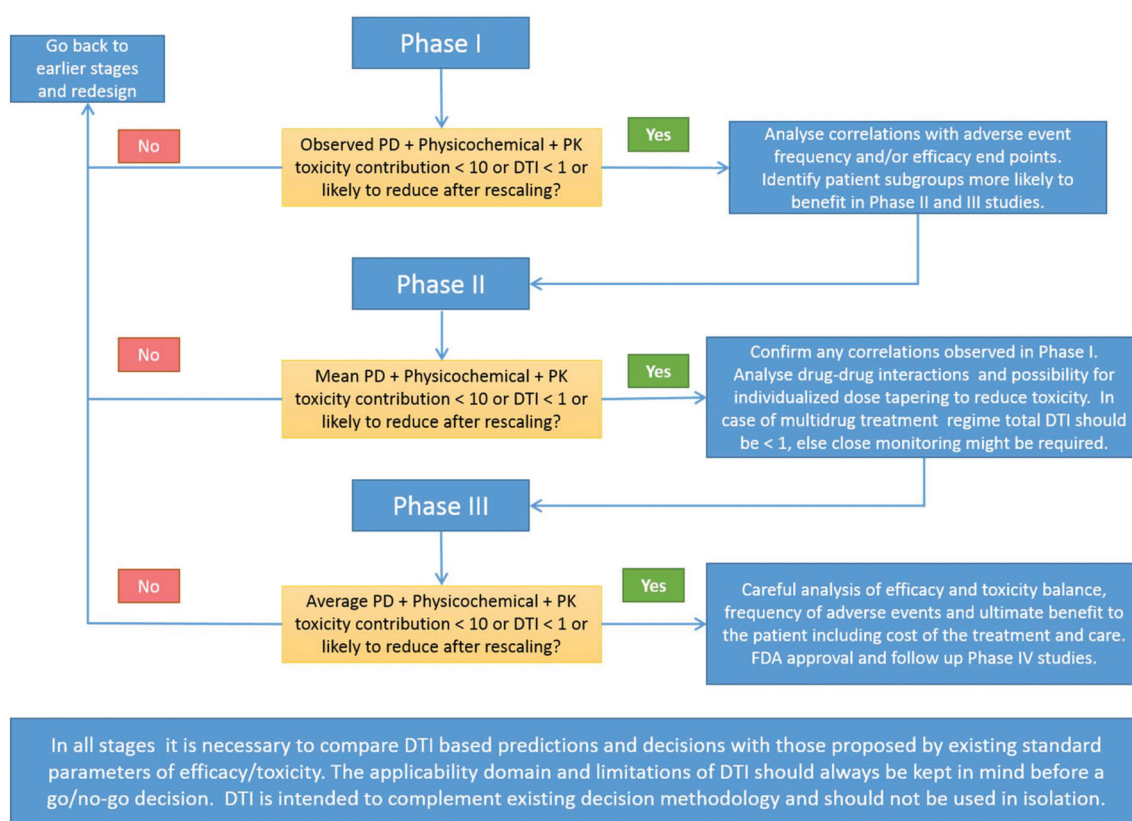


Fig. 7 Physicochemical and PD toxicity contributions for candidates entering Phase I can be altered only by modification of dose and/or oral bioavailability. PK contributions can be estimated from the Phase I C_{max} and estimates of oral bioavailability. DTI and different contributions can be interpreted with confidence in therapeutic areas with good correlation of efficacy/toxicity endpoints with DTI demonstrated in this study. Generally candidates with $\text{DTI} < 0.5$ can be considered for the next phase, while those with $\text{DTI} \geq 1 \pm 0.5$ should be reassessed carefully. Similar thinking can be applied during Phase II and Phase III to gain insights into the mechanisms of toxicity and propose measures to monitor and reduce them. Applications of DTI in combination therapies, personalized dosing/therapies and drug monitoring are principally possible but need to be established with a careful study design and data analysis. Blue filled boxes on the right indicate the data analysis required while moving to the next stage of clinical trials. Light orange color filled boxes indicate decision points based on toxicity contributions (<10) or DTI values (<1).

The DTI concept is not directly applicable to drugs with purely physical (*e.g.* sorbitol and deoxycholic acid), chemical (*e.g.* Clofazimine, Cyclophosphamide and Methoxsalen) or unknown mechanisms of action (*e.g.* Quinacrine, Omega-3 acid ethyl esters and Stiripentol). For these drugs biological targets are seldom well defined against which IC_{50} or K_d values can be evaluated. Estimation of off target PD, PK and physico-chemical contributions might still prove useful for understanding toxicity profiles.

Since artificial intelligence (AI) or machine learning methods present unique difficulties (*e.g.* requirement of larger datasets, computational time, hardware, well-defined and consistent bioactivity end points) in the implementation and interpretation of complex drug discovery data, these methods were not applied in this study. Very few cases of AI based toxicity predictions, directly relevant here, exist.⁵⁸ Nevertheless, the application of AI and deep learning may improve predictions; this will be taken up in our future attempts of toxicity studies.

4. Conclusions

The DTI, an improved mathematical model, has been presented to redefine drug toxicity as biphasic and exponential functions of typical PD, PK and physicochemical parameters. It has been found to perform better than some of the widely used ligand efficiency indices, drug-likeness rules, and exposure based TIs. DTI is useful in differentiating withdrawn drugs from nine other WHO ATC drug categories and correctly identifies higher toxicity potential for antiviral and antibiotic drugs. DTI correctly identifies the DILI concern with excellent accuracy for different WHO ATC drug categories. It does particularly well for CNS, Cardiovascular, Musculoskeletal and Antiviral categories. DTI was also found to be useful in estimating relative drug efficacy in eight therapeutic areas (Antibiotics, CNS: antipsychotics and antidepressants, Metabolic and GIT, Gentic-Urinary and Sex, Antiviral, Blood and blood forming agents, and Musculoskeletal) with good correlations with clinical outcome parameters. Predictions for Antiparasitic, Systemic & Hormonal, and Blood and blood forming agents should be interpreted cautiously due to the lower number of drugs analyzed in these categories. Contributions of PD, PK and physicochemical parameters to DTI give insights for understanding drug toxicity mechanisms and are useful for identifying potentially toxic drugs and doses. DTI is intended to complement the LEI, LLE, classical drug-likeness rules, lead optimization strategies and TI heatmap grids generated during drug development. Results for the recently approved drugs generate hope for applications in drug discovery and therapeutic drug monitoring. Flowchart like decision trees for the application of DTI throughout the drug discovery process has been proposed. Limitations of the current DTI approach must be considered during such applications. Additional studies are required to further explore the utility of the DTI concept in other therapeutic areas and settings.

Funding information

No funding was received for performing this work.

Abbreviations

TI	Therapeutic index
LEI	Ligand efficiency index
PK	Pharmacokinetic
PD	Pharmacodynamic
DTI	Drug toxicity index
C_{max}	Maximum plasma drug concentration
IC_{50}	Half maximal inhibitory concentration
LTKB	Liver toxicity knowledge base
GPCR	G-protein coupled receptor
AChE	Acetylcholine esterase
COX	Cyclo-oxygenase
MAO	Monoamine-oxidase
RO5	Rule of five
RO2	Rule of two
RO2-RM	Rule of two-reactive metabolite
nM	Nanomolar
μ M	Micromolar
C_{min}	Minimum plasma drug concentration
C_{eff}	Pharmacologically effective concentration
PPB	Plasma protein binding
free C_{max}	Unbound maximum plasma drug concentration
LE	Ligand efficiency
DEI	Drug efficiency index
ADME	Absorption, distribution, metabolism and elimination
ADMET	Absorption, distribution, metabolism, elimination and toxicity
Vd	Volume of distribution
Clog P	Calculated log P
LLE	Lipophilic ligand efficiency
ANOVA	Analysis of variance
GIT	Gastrointestinal tract
DILI	Drug induced liver injury
WHO	World Health Organization anatomical therapeutic chemical classification
ATC	
PDE5	Phosphodiesterase 5
PPAR γ	Peroxisome proliferator activated receptor gamma
K_d	Dissociation rate constant
AI	Artificial intelligence

Author contributions

VAD conceived the idea, conducted the literature search, collected data, performed analysis and wrote the manuscript.

Conflicts of interest

There are no conflicts of interest to declare.

References

- 1 P. Y. Muller and M. N. Milton, *Nat. Rev. Drug Discovery*, 2012, **11**, 751.
- 2 L. B. Laurence, A. C. Bruce and C. K. Bjorn, *Goodman and Gilman's the pharmacological basis of therapeutics*, The McGraw-Hill Companies, Twelfth Ed., 2011.
- 3 D. Cook, D. Brown, R. Alexander, R. March, P. Morgan, G. Satterthwaite and M. N. Pangalos, *Nat. Rev. Drug Discovery*, 2014, **13**, 419.
- 4 M. J. Waring, J. Arrowsmith, A. R. Leach, P. D. Leeson, S. Mandrell, R. M. Owen, G. Pairaudeau, W. D. Pennie, S. D. Pickett, J. Wang, O. Wallace and A. Weir, *Nat. Rev. Drug Discovery*, 2015, **14**, 475.
- 5 C. B. Eap, T. Buclin and P. Baumann, *Clin. Pharmacokinet.*, 2002, **41**, 1153–1193.
- 6 D. Wang, V. V. Patel, E. Ricciotti, R. Zhou, M. D. Levin, E. Gao, Z. Yu, V. A. Ferrari, M. M. Lu, J. Xu, H. Zhang, Y. Hui, Y. Cheng, N. Petrenko, Y. Yu and G. A. FitzGerald, *Proc. Natl. Acad. Sci. U. S. A.*, 2009, **106**, 7548.
- 7 J. Bowes, A. J. Brown, J. Hamon, W. Jarolimek, A. Sridhar, G. Waldron and S. Whitebread, *Nat. Rev. Drug Discovery*, 2012, **11**, 909.
- 8 M. P. Gleeson, *J. Med. Chem.*, 2008, **51**, 817–834.
- 9 J. D. Hughes, J. Blagg, D. A. Price, S. Bailey, G. A. DeCrescenzo, R. V. Devraj, E. Ellsworth, Y. M. Fobian, M. E. Gibbs, R. W. Gilles, N. Greene, E. Huang, T. Krieger-Burke, J. Loesel, T. Wager, L. Whiteley and Y. Zhang, *Bioorg. Med. Chem. Lett.*, 2008, **18**, 4872–4875.
- 10 C. A. Lipinski, *Drug Discovery Today: Technol.*, 2004, **1**, 337–341.
- 11 H. Jhoti, G. Williams, D. C. Rees and C. W. Murray, *Nat. Rev. Drug Discovery*, 2013, **12**, 644.
- 12 M. Chen, J. Borlak and W. Tong, *Hepatology*, 2012, **58**, 388–396.
- 13 M. Chen, J. Borlak and W. Tong, *Hepatology*, 2016, **64**, 931–940.
- 14 J. P. Hughes, S. Rees, S. B. Kalindjian and K. L. Philpott, *Br. J. Pharmacol.*, 2010, **162**, 1239–1249.
- 15 M. C. Hutter, *Future Med. Chem.*, 2018, **10**, 1623–1635.
- 16 D. J. Cummins and M. A. Bell, *J. Med. Chem.*, 2016, **59**, 6999–7010.
- 17 M. C. Wenlock, *MedChemComm*, 2016, **7**, 706–719.
- 18 A. L. Hopkins, G. M. Keserü, P. D. Leeson, D. C. Rees and C. H. Reynolds, *Nat. Rev. Drug Discovery*, 2014, **13**, 105.
- 19 M. M. Trachez, G. Zapata-Sudo, O. R. Moreira, N. G. B. Chedid, V. F. T. Russo, E. M. S. Russo and R. T. Sudo, *Acta Anaesthesiol. Scand.*, 2005, **49**, 66–71.
- 20 S. Braggio, D. Montanari, T. Rossi and E. Ratti, *Expert Opin. Drug Discovery*, 2010, **5**, 609–618.
- 21 D. Montanari, E. Chiarparin, M. P. Gleeson, S. Braggio, R. Longhi, K. Valko and T. Rossi, *Expert Opin. Drug Discovery*, 2011, **6**, 913–920.
- 22 P. L. M. van Giersbergen, A. Halabi and J. Dingemans, *Br. J. Clin. Pharmacol.*, 2002, **53**, 589–595.
- 23 D. Bonchev, *J. Mol. Graphics Modell.*, 2001, **20**, 65–75.
- 24 T. Kron, A. Elliott and P. Metcalfe, *Med. Phys.*, 1993, **20**, 1429–1438.
- 25 M. J. Waring, *Expert Opin. Drug Discovery*, 2010, **5**, 235–248.
- 26 W. O. Foye, T. L. Lemke and D. A. Williams, *Foye's principles of medicinal chemistry*, Wolters Kluwer Health/Lippincott Williams & Wilkins, 2013.
- 27 P. Duan, S. Li, N. Ai, L. Hu, W. J. Welsh and G. You, *Mol. Pharm.*, 2012, **9**, 3340–3346.
- 28 ToxTutor, <https://toxmentor.nlm.nih.gov/03-002.html>, (accessed 15 September 2018).
- 29 H. Padma-Nathan, I. Goldstein, I. Klimberg, C. Coogan, S. Auerbach, P. Lammers and V. S. Group, *Int. J. Impotence Res.*, 2002, **14**, 266.
- 30 C. C. Ogu and J. L. Maxa, *Proc. (Bayl. Univ. Med. Cent.)*, 2000, **13**, 421–423.
- 31 E. S. Björnsson and J. H. Hoofnagle, *Hepatology*, 2015, **63**, 590–603.
- 32 J. Chen and J. Ahmad, *J. Gen. Intern. Med.*, 2008, **23**, 1914–1916.
- 33 A. W. Hayes, *Principles and methods of toxicology*, CRC Press, 2007.
- 34 D. A. Smith, L. Di and E. H. Kerns, *Nat. Rev. Drug Discovery*, 2010, **9**, 929.
- 35 H.-G. Eichler, B. Bloechl-Daum, E. Abadie, D. Barnett, F. König and S. Pearson, *Nat. Rev. Drug Discovery*, 2010, **9**, 277.
- 36 H. Kim, L. Gurrin, Z. Ademi and D. Liew, *Br. J. Clin. Pharmacol.*, 2013, **77**, 116–121.
- 37 D. Henry, L. L.-Y. Lim, L. A. G. Rodriguez, S. P. Gutthann, J. L. Carson, M. Griffin, R. Savage, R. Logan, Y. Moride, C. Hawkey, S. Hill and J. T. Fries, *Br. Med. J.*, 1996, **312**, 1563.
- 38 B. J. Knottnerus, L. Grigoryan, S. E. Geerlings, E. P. Moll van Charante, T. J. M. Verheij, A. G. H. Kessels and G. ter Riet, *Fam. Pract.*, 2012, **29**, 659–670.
- 39 A. Zimmer, I. Katzir, E. Dekel, A. E. Mayo and U. Alon, *Proc. Natl. Acad. Sci. U. S. A.*, 2016, **113**, 10442.
- 40 S. Leucht, A. Cipriani, L. Spineli, D. Mavridis, D. Örey, F. Richter, M. Samara, C. Barbui, R. R. Engel, J. R. Geddes, W. Kissling, M. P. Stapf, B. Lässig, G. Salanti and J. M. Davis, *Lancet*, 2013, **382**, 951–962.
- 41 M. Curto, N. Girardi, L. Lionetto, G. M. Ciavarella, S. Ferracuti and R. J. Baldessarini, *Curr. Psychiatry Rep.*, 2016, **18**, 68.
- 42 A. L. Khoo, H. J. Zhou, M. Teng, L. Lin, Y. J. Zhao, L. B. Soh, Y. M. Mok, B. P. Lim and K. P. Gwee, *CNS Drugs*, 2015, **29**, 695–712.
- 43 E. S. Mearns, D. M. Sobieraj, C. M. White, W. J. Saulsberry, C. G. Kohn, Y. Doleh, E. Zaccaro and C. I. Coleman, *PLoS One*, 2015, **10**, e0125879–.
- 44 L. Chen, S. E. L. Staubli, M. P. Schneider, A. G. Kessels, S. Ivic, L. M. Bachmann and T. M. Kessler, *Eur. Urol.*, 2015, **68**, 674–680.
- 45 J. Yuan, R. Zhang, Z. Yang, J. Lee, Y. Liu, J. Tian, X. Qin, Z. Ren, H. Ding, Q. Chen, C. Mao and J. Tang, *Eur. Urol.*, 2013, **63**, 902–912.

- 46 D. A. Patel, S. J. Snedecor, W. Y. Tang, L. Sudharshan, J. W. Lim, R. Cuffe, S. Pulgar, K. A. Gilchrist, R. R. Camejo, J. Stephens and G. Nichols, *PLoS One*, 2014, **9**, e105653–.
- 47 A. Tawfik, J. Bielecki, M. Krahn, P. Dorian, J. Hoch, H. Boon, D. Husereau and P. Pechlivanoglou, *Clin. Pharmacol.: Adv. Appl.*, 2016, **8**, 93–107.
- 48 A. Fretheim, J. Odgaard-Jensen, O. Brørs, S. Madsen, I. Njølstad, O. F. Norheim, A. Svilaas, I. S. Kristiansen, H. Thürmer and S. Flottorp, *BMC Med.*, 2012, **10**, 33.
- 49 K. Jordan, H. J. Schmoll and M. S. Aapro, *Crit. Rev. Oncol. Hematol.*, 2007, **61**, 162–175.
- 50 R. K. Harrison, *Nat. Rev. Drug Discovery*, 2016, **15**, 817.
- 51 V. A. Dixit and P. V. Bharatam, *Chem. Res. Toxicol.*, 2011, **24**, 1113–1122.
- 52 C.-M. Loi, C. W. Alvey, A. B. Vassos, E. J. Randinitis, A. J. Sedman and J. R. Koup, *J. Clin. Pharmacol.*, 2013, **39**, 920–926.
- 53 Z. V. Varga, P. Ferdinandy, L. Liaudet and P. Pacher, *Am. J. Physiol.: Heart Circ. Physiol.*, 2015, **309**, H1453–H1467.
- 54 S. G. Meeves and S. Appajosyula, *J. Allergy Clin. Immunol.*, 2003, **112**, S69–S77.
- 55 H. Vidot, S. Carey, M. Allman-Farinelli and N. Shackel, *Aliment. Pharmacol. Ther.*, 2014, **40**, 221–232.
- 56 E. M. Stein, C. D. DiNardo, D. A. Pollyea, A. T. Fathi, G. J. Roboz, J. K. Altman, R. M. Stone, D. J. DeAngelo, R. L. Levine, I. W. Flinn, H. M. Kantarjian, R. Collins, M. R. Patel, A. E. Frankel, A. Stein, M. A. Sekeres, R. T. Swords, B. C. Medeiros, C. Willekens, P. Vyas, A. Tosolini, Q. Xu, R. D. Knight, K. E. Yen, S. Agresta, S. de Botton and M. S. Tallman, *Blood*, 2017, **130**, 722.
- 57 V. A. Dixit, L. A. Lal and S. R. Agrawal, *WIREs Comput. Mol. Sci.*, 2017, **7**, e1323.
- 58 A. Mayr, G. Klambauer, T. Unterthiner and S. Hochreiter, *Front. Environ. Sci.*, 2016, **3**, 80.