New Associations Between Drug-Induced Adverse Events in Animal Models and Humans Reveal Novel Candidate Safety Targets

Journal:	*Chemical Research in Toxicology*
Manuscript ID	tx-2020-00311s.R2
Manuscript Type:	Article
Date Submitted by the Author:	04-Dec-2020
Complete List of Authors:	Giblin, Kathryn; AstraZeneca PLC, Oncology Medicinal Chemistry Basili, Danilo; University of Cambridge, Department of Chemistry Afzal, Avid; AstraZeneca PLC, Data Sciences & Quantitative Biology, Discovery Sciences Rosenbrier-Ribeiro, Lyn; Medicines Discovery Catapult Greene, Nigel; Astrazeneca, Drug Safety & Metabolism Barrett, Ian; Discovery Sciences Hughes, Samantha; AstraZeneca PLC, Oncology Medicinal Chemistry Bender, Andreas; University of Cambridge, Chemistry Department

New Associations Between Drug-Induced Adverse Events in Animal Models and Humans Reveal Novel Candidate Safety Targets

Kathryn A. Giblin,[1,2*] Danilo Basili,[1] Avid M. Afzal,[3] Lyn Rosenbrier-Ribeiro,[4†] Nigel Greene,[5] Ian Barrett,[3] Samantha J. Hughes,[2] Andreas Bender[1*%]

1. Centre for Molecular Informatics, Department of Chemistry, University of Cambridge, Lensfield Road, Cambridge, CB2 1EW, UK

2. Medicinal Chemistry, Research and Early Development, Oncology R&D, AstraZeneca, Cambridge, UK.

3. Data Sciences & Quantitative Biology, Discovery Sciences, R&D, AstraZeneca, Cambridge, UK

4. Safety Platforms, Clinical Pharmacology and Safety Sciences, R&D, AstraZeneca, Cambridge, UK.

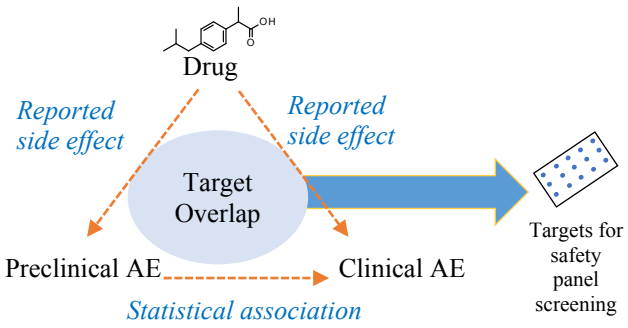
5. Data Science and AI, Clinical Pharmacology and Safety Sciences, R&D, AstraZeneca, Boston, US.

*kathryn.giblin@astrazeneca.com, ab454@cam.ac.uk

†present address: Virtual R&D, Medicines Discovery Catapult, Block 35, Alderley Park, Cheshire, SK10 4TG, UK.

%present address: Data Science and AI, Clinical Pharmacology and Safety Sciences, R&D, AstraZeneca, Cambridge, UK

For table of contents only



Abstract

To improve our ability to extrapolate preclinical toxicity to humans, there is a need to understand and quantify the concordance of adverse events (AEs) between animal models and clinical studies. In the present work, we discovered 3,011 statistically significant associations between preclinical and clinical AEs caused by drugs reported in the PharmaPendium database of which 2,952 were new associations between toxicities encoded by different MedDRA terms across species. To find plausible and testable candidate off-target drug activities for the derived associations, we investigated the genetic overlap between the genes linked to both a preclinical and a clinical AE and the protein targets found to interact with one or more drugs causing both AEs. We discuss three associations from the analysis in more detail, for which novel candidate off-target drug activities could be identified; namely the association of preclinical mutagenicity readouts with clinical TERATOSPERMIA and OVARIAN FAILURE, the association of preclinical REFLEXES ABNORMAL with clinical POOR-QUALITY SLEEP and the association of preclinical PSYCHOMOTOR HYPERACTIVITY with clinical DRUG WITHDRAWAL SYNDROME. Our analysis successfully identified a total of 77 % known safety targets currently tested in *in vitro* screening panels plus an additional 431 genes which were proposed for investigation as future safety targets for different clinical toxicities. This work provides new translational toxicity relationships beyond adverse event term-matching, the results of which can be used for risk-profiling of future new chemical entities for clinical studies and for the development of future *in vitro* safety panels.

Keywords

Toxicity Translation, Safety Targets, Statistical Association, Risk Profiling, Adverse Events

Introduction

Testing new chemical entities (NCEs) in animal models is a regulatory requirement for toxicity assessment before administration of a drug to humans in clinical trials. However, drug-induced animal toxicity and drug-induced human toxicity do not always correlate, which can lead to either unwarranted attrition [1,2] or increased costs and risk to human health [3]. For example, it is well-known that cardiovascular endpoints are well predicted from animal models [4], however, cutaneous and neurological toxicities translate less well across species [5]. From an analysis of the attrition of drug candidates from four pharmaceutical companies, it was found that 40 % of drug candidates were terminated due to non-clinical toxicology findings and 11 % due to clinical safety findings [6]. Safety was found to be the highest contributor to attrition in both preclinical and phase I studies [6]. In an AstraZeneca study conducted on drug projects between 2005–2010, it was found that of those drug projects terminated preclinically, 82 % were due to toxicity [7]. Furthermore, of those drug projects which passed into clinical trials, 62 %, 35 % and 12 % of them were terminated due to toxicity in phase I, phase II and phase III, respectively [7], showing that also in the clinical phases (and even in clinical phases II and III, which have efficacy as their main objective), toxicity is a major cause of attrition. More recently, there has been a renewed focus on the "right safety" of compounds within AstraZeneca [8], leading to a reduction in project closures due to lack of safety between 2012-2016, down to 38 % for phase I and 8 % for phase II. This result shows the vital contribution that toxicity models can make to project success, and that continued improvement in this area is likely to further reduce drug attrition. Since clinical toxicity is still a significant source of drug attrition, it is clear that the predictivity of animal models is currently not sufficient to anticipate compound safety and toxicity of later tests in man [9]. Increasingly, there is a drive towards the reduction of animal usage in support of the 3R's initiative [10] by prioritising the use of only the most toxicologically

relevant animal models, as well as implementing and developing alternative *in vitro* models. To this extent, earlier stage assessment based on secondary pharmacology screening may be successful in driving the reduction of animal testing. There are different screening panels of targets which are used to screen molecules for off-target liabilities, including the safety panel published in Lounkine *et al.,* 2012 [11], based on the Novartis safety panel,[12] and the Cross Pharma Consensus panel known as panel-44, which uses the targets published by Bowes *et al.*, 2012 [13]. The study by Lynch *et al.,* 2017 [14] summarizes the evidence for adverse events induced by agonism/activation and antagonism/inhibition of each of 70 off-targets included in AbbVie's *in vitro* screening panel[15]. Other studies have proposed targets for secondary pharmacology panels using data driven analysis, including the recent study by Deaton and collaborators, which provides a suggested off-target screening panel based on deriving links between off-target phenotypes and the side effect profile of a drug using both enrichment analysis and multivariate logistic regression [16], and the machine learning study by Letswaart and collaborators, which related off-targets to adverse drug reactions[17]. For the reasons mentioned above, it is important to understand which toxicity endpoints translate from preclinical to clinical studies and which ones do not. Such efforts to understand and quantify the concordance of adverse events (AEs) between animal models and clinical studies have been conducted previously using retrospective statistical analyses. Studies of this type are not trivial to conduct, due to small data set sizes, biological variability and species exposure differences [18], as well as biased data such as due to 'survivor bias', since data for drugs which are terminated before clinical trials due to safety (or other) reasons will not be present in the analysis.

It is important that analyses such as the above are conducted using high quality safety datasets. The PharmaPendium database, used in previous concordance studies as well as in our study, contains curated safety data from a range of high quality sources. Regulatory healthcare

agencies, including the European Medicines Agency (EMA) and Food and Drug Administration (FDA), have requirements for the reporting of AEs to pharmacovigilance databases, which leads to a source of clinical data which can be incorporated into databases such as PharmaPendium. The current version of PharmaPendium contains over 2 million adverse event reports for 4704 drugs and includes information reported from over 70 years of past drug approvals. Sources of AE data include FDA approval packages, EMA approval documents, Meyler's side effects of drugs,[19]Mosby's Drug Consult [20] and toxicity data curated by PharmaPendium from publications[21]. Pharmacovigilance databases including PharmaPendium are often coded to a medical dictionary of terms to aid analysis; the main dictionary used in the UK, US and Japan is the Medical Dictionary for Regulatory Activities (MedDRA)[22], which is also used in PharmaPendium[21]. (MedDRA) provides a 5-level hierarchy of AE terms, where the highest level of this hierarchy is the System Organ Class (SOC) and the lowest level terms are more likely to be the original terms described in a document.[22] It is important to note that since MedDRA terminology was designed for the description of human AEs, it doesn't cover pathological terms very well, but focuses more on physiological or functional impairment effects. Therefore, there will be limitations in the ability of the terminology to accurately map to all text descriptions of preclinical observations. However, despite these limitations the MedDRA dictionary provides a highly governed terminology with a hierarchical dictionary to map to descriptions of AEs at different levels of granularity[22].

The main findings from previous studies included that haematological, gastrointestinal, injection site and some specific cardiovascular AEs displayed a high concordance (demonstrated by LR^+ values of greater than 10 or sensitivities greater than 70 %, Table S1)[4,18,23,24]; however, neurological toxicity and cutaneous toxicities have a poor concordance (demonstrated by a sensitivity of less than 35 %) [4,24]. It was also observed that concordance was higher for small molecules than it was for large molecules (e.g. antibodies) [24,25] and that

there were differences between concordance with humans for different preclinical species [4,25,26]. For a comprehensive review of previous concordance literature the reader is referred to Monticello *et al*. [27] and to Table S1.

What all previous studies have in common, and the major difference to the current work, is that they measured the association between the same preclinical toxicity, or a preclinical toxicity related to the same system organ class (SOC) as the clinical AE. Some adverse effects in humans are not predicted by the same AE in animals, due to the differences in anatomy, physiology and biology between species. An example of this is the lack of a vomiting response in rats [28], which excludes this from being used as a model for vomiting in humans. Instead, taste aversion/food avoidance responses in rodents or ferret, or dog emesis models are used [28,29]. As another example, species differences exist for some teratogenic based toxicities, for example corticosteroids are teratogenic in animal models but not in humans [30] and conversely thalidomide is a teratogen in humans but not in many animal species, which has been attributed in part to the differences in metabolism across species [31]. In addition, the lack of correlation of drug bioavailability between species [5], linked to poor dose extrapolation between species, can lead to differences in toxicity observations [32]. These and other reasons demonstrate how concordance between the same AE across species may limit our ability to translate toxicity findings from animals to humans.

Aiming to account for species differences when observing animal-based safety signals and at the same time attempting to widen the predictivity of available animal endpoints, in this study we implemented a data-driven approach to identify relationships between animal and human toxicities that are not limited by considering the same system organ class (SOC) as the clinical AE. More specifically, we extended the previous concordance approaches to find associations between AEs encoded by different MedDRA terms. To provide biological support for our newly discovered associations and to propose biological targets for secondary pharmacology

screens, we have next set out to explore the genetic information shared between the drugs and both the preclinical and clinical AEs, since drugs with genetically supported targets have been recently demonstrated more likely to be successful in Phase II and Phase III trials [33]. More specifically, for each association we retrieved the genes linked to the ontology terms representative of the preclinical and clinical AEs as well as the genes whose protein targets were found to interact with one of more drugs causing both AEs. We then determined the overlap of genetic evidence among the three domains identifying a set of plausible off-target activities. We were hence able to propose targets as candidates for incorporation into *in vitro* secondary pharmacology screening panels, pending further experimental validation.

Results

Identifying concordance between preclinical and clinical AEs within the same MedDRA term

In order to identify and quantify the concordance between preclinical and clinical AEs, we computed the Mutual Information (MI) and Likelihood Ratio (LR) values and assessed the significance for those AEs which were recorded by the same MedDRA term[18], similar to previous studies (Table S1). MedDRA terms are represented throughout the text in small capitals to provide clarity. As mentioned in the introduction, MedDRA is a hierarchical dictionary of terms, however, here we analyse the associations between the preferred term (PTs) level of the hierarchy, since grouping a collection of terms with different underlying modes of action into a single class makes the gene to phenotype associations that we utilise in this study more difficult to identify. Overall, 188 of the possible 473 (40%) matching preclinical and clinical MedDRA terms were identified as significantly associated

(Supplementary Data File 1). The 10 highest associations for the same adverse event reported in preclinical and clinical studies are shown in Table S2. A total of 6 out of these 10 highest associations, including DRUG SPECIFIC ANTIBODY PRESENT (LR^+ 118), BLOOD PROLACTIN INCREASE (LR^+ 23), ELECTROCARDIOGRAM QT CORRECTED INTERVAL PROLONGED (LR^+ 11), ALANINE AMINOTRANSFERASE INCREASED (LR^+ 3), DIARRHOEA (LR^+ 4) and INJECTION SITE ERYTHEMA (LR^+ 10), were also reported in previous studies with comparable LR^+ values [18,23]. In addition to these already discovered associations, we also find that RENAL PAPILLARY NECROSIS, CARDIOTOXICITY, SEDATION and PLATELET COUNT INCREASED were concordant between animal species and humans with LR^+ values of 33, 25, 5 and 4, respectively, highlighting the relevance of these readouts in animal models to the prediction of the clinical AE. Since only 30 % (473 out of 1,740) of MedDRA terms for clinical AEs were matched by a preclinical AE with the same term and we know that biological relationships exist between different cell types, tissues and organs, this approach may lack the ability to identify additional toxicity relationships between non-matching terms. Additionally, among these 473 matched terms only 40 % were statistically associated by our analysis, which poses the question as to whether other preclinical AEs in the dataset might be more relevant than the exact MedDRA term match in predicting certain clinical AEs and suggests that novel targets of toxicity could be identified.

Linking Preclinical and Clinical AEs across MedDRA terms

We hence next investigated which drug-induced animal toxicities were predictive for human AEs, without the requirement of the AE to be encoded by the same term in both species. In total, 3,011 statistically significant associations (p-value < 0.01 assessed by normalised mutual information followed by Bonferroni post-hoc correction; see methods for details) were found between preclinical and clinical AEs (Supplementary Data File 2), with Mutual Information (MI) values ranging between 0.02 - 0.33, positive Likelihood Ratio (LR^+) values ranging

between 1.5 - 118.5 and negative Likelihood Ratio (LR^-) values ranging between 0.10-0.97. Consistent with a previous study [18], we find that whilst the positive Likelihood Ratios show a significant change in risk of the clinical AE given the preclinical AE (94 % of significant associations had LR^+ greater than 2), the negative Likelihood Ratios show that the absence of a preclinical AE does not alter the probability of clinical safety by an important degree (94 % of the LR^- values were between 0.5-1). The 20 associations with the highest Mutual Information values are presented in Table 1, along with their Likelihood Ratios to quantify the conditional risk. The association with the highest MI score of 0.33 as well as the highest positive Likelihood Ratio (LR^+ 118) was the preclinical finding of DRUG SPECIFIC ANTIBODY PRESENT associated with the same finding clinically. The same AE preclinically is also associated with clinical INFUSION RELATED REACTION (MI= 0.24 and LR^+ 44), most likely because monoclonal antibodies are administered by infusion [25]. The other members of the 20 highest MI associations contain preclinical findings including WAXY FLEXIBILITY, BLOOD PROLACTIN INCREASED, INTESTINAL ULCER, ADRENAL CORTEX ATROPHY and PERITONITIS. Preclinical WAXY FLEXIBILITY and BLOOD PROLACTIN INCREASED were found to be associated with the most clinical toxicities for the top 20 associations (11 and 5, respectively). These preclinical findings were predictive of clinical MEIGE'S SYNDROME, AUTONOMIC NERVOUS SYSTEM IMBALANCE, SCHIZOPHRENIA, FACIAL SPASM and TARDIVE DYSKINESIA, a range of central nervous system (CNS) side effects which have been recognised as a toxicity area which needs to be better predicted in preclinical models [34,35]. From our analysis we find that behavioural symptoms of catatonia in animal species (e.g. WAXY FLEXIBILITY) may serve as good indicators of a variety of disorders, including dyskinesias (MI=0.2 and LR^+ 43), spasms (MI=0.2 and LR^+ 43) and male fertility problems (MI=0.18-019 and LR^+ 39-42) (Table 1). While the association between catatonia and neurological disorders has now been elucidated [36], our approach identified potential new relationships with male fertility. In addition, our study highlights that the behavioural

observation of WAXY FLEXIBILITY when present in preclinical studies elevates the risk of the aforementioned adverse effects in the clinic by up to 44 times. We also identify that the preclinical BLOOD PROLACTIN INCREASED elevates the risk of some neurological toxicities by over 50 times. Given that most of the drugs responsible for these associations in this study are antipsychotics, antianxiolitics and antidepressants, this link is most likely correlative rather than causative, since these drug classes modulate serotonin and dopamine levels, neurotransmitters known to affect circulating prolactin levels[37,38] Nevertheless, a wide range of liquid biomarkers have been previously identified as being important for neurological toxicity prediction [35].

Table 1: The 20 statistical associations with the highest MI values between preclinical and clinical adverse events. Reported are the normalised mutual information, Bonferroni corrected p-value from Fisher's exact test, positive and negative likelihood ratios and the number of intersecting drugs between the preclinical and clinical AE terms. Grouped by preclinical AE

Preclinical AE	Clinical AE	Normalised Mutual Information (MI)	Bonferroni corrected Fishers' exact test p-value	Positive Likelihood Ratio (LR^+)	Negative Likelihood Ratio (LR^-)	Number of intersecting drugs between preclinical and clinical AE
DRUG SPECIFIC ANTIBODY PRESENT	DRUG SPECIFIC ANTIBODY PRESENT	0.33	1.37×10^{-39}	118	0.67	35
	INFUSION RELATED REACTION	0.24	4.40×10^{-24}	44	0.66	24
WAXY FLEXIBILITY	DYSKINESIA OESOPHAGEAL	0.20	3.98×10^{-11}	43	0.60	11
	FACIAL SPASM	0.20	3.98×10^{-11}	43	0.60	11
	MEIGE'S SYNDROME	0.20	3.65×10^{-10}	44	0.57	10
	INTESTINAL DILATATION	0.19	6.20×10^{-10}	42	0.59	10
	SPERM COUNT INCREASED	0.19	6.20×10^{-10}	42	0.59	10
	ASPHYXIA	0.19	1.92×10^{-12}	39	0.67	13
	RETROGRADE EJACULATION	0.18	1.63×10^{-10}	39	0.64	11
	AUTONOMIC NERVOUS SYSTEM IMBALANCE	0.18	9.39×10^{-14}	37	0.72	15

	ASPERMIA	0.18	1.65×10^{-9}	39	0.62	10
	BLOOD ANTIDIURETIC HORMONE INCREASED	0.18	1.65×10^{-9}	39	0.62	10
BLOOD PROLACTIN INCREASED	COGWHEEL RIGIDITY	0.26	4.35×10^{-15}	53	0.46	13
	DROOLING	0.20	4.71×10^{-12}	33	0.58	12
	FACIAL SPASM	0.18	1.81×10^{-10}	36	0.60	11
	SCHIZOPHRENIA	0.18	3.36×10^{-11}	35	0.63	12
	TARDIVE DYSKINESIA	0.18	1.58×10^{-16}	33	0.75	19
INTESTINAL ULCER	RENAL PAPILLARY NECROSIS	0.25	8.07×10^{-19}	60	0.65	17
ADRENAL CORTEX ATROPHY	ADRENAL SUPPRESSION	0.20	7.97×10^{-20}	25	0.54	22
PERITONITIS	RENAL PAPILLARY NECROSIS	0.19	4.08×10^{-18}	27	0.59	20

Preclinical INTESTINAL ULCER and preclinical PERITONITIS were both associated with RENAL PAPILLARY NECROSIS within the 20 most significant associations and INTESTINAL ULCER was found be more predictive for the presence of RENAL PAPILLARY NECROSIS ($LR^+ = 60$) than the preclinical observation of RENAL PAPILLARY NECROSIS itself ($LR^+ = 33$). INTESTINAL ULCER, PERITONITIS and RENAL PAPILLARY NECROSIS are side effects observed in nonsteroidal anti-inflammatory drugs (NSAIDs), and all drugs supporting this association were indeed NSAID drugs. This is an example which illustrates the limits of co-occurrence for finding useful human AE predictors from preclinical AEs when the association is supported by a single drug class. It is likely that these AEs are species-specific, although there is some evidence to suggest that the two events are connected since Cox-1 inhibition is known to lead to decreased prostaglandin production which is necessary for both mucosal integrity in the GI tract and maintenance of blood flow to the renal papillae [39]. Our analysis also confirmed the potential of using preclinical ADRENAL CORTEX ATROPHY to predict ADRENAL SUPPRESSION clinically ($LR^+ = 25$). ADRENAL SUPPRESSION is one of the main side effects associated with long term corticosteroid usage [40]. Since the adrenal glands are suppressed from corticosteroid usage, less cortisol is released over

time. In animals this effect is often measured by histopathological examination of the adrenal glands [41]. Taken all together, these highest-ranking associations can be backed up by strong literature evidence hence supporting the effectiveness of our approach in identifying meaningful associations. Notably, all clinical AEs in the 20 most significant associations, with the exception of DRUG SPECIFIC ANTIBODY PRESENT, had a higher MI value of association with a preclinical AE which was *not identical* in clinical and preclinical space and the identical preclinical AE was found to have one of the highest three MI values for only 59 clinical AEs. We show that this method was able to discover 2,952 statistically significant relationships between preclinical and clinical AEs that were beyond the straightforward term matching, as well as 59 concordant associations, to make up 3,011 associations overall.

Identifying plausible off-target drug activities

We next set out to explore potential off-target activities linked to our associations, for full details see methods. Briefly, we first identified targets modulated by drugs inducing both AEs on the animal as well as the human level in the association, by mining multiple drug-target interaction databases. Then, we retrieved genes associated with preclinical and clinical AEs by converting AEs to diseases or phenotype IDs and subsequently mapping the resultant IDs to phenotype-gene and disease-gene association databases. Finally, we mapped all the identified genes to their human orthologs and we then conducted an overlap analysis to identify genes shared between the preclinical and clinical AEs and one or more of the drugs which induced AEs in both domains. We conducted our overlap analysis on a subset of 248 highly scoring associations as measured by the MI (see methods for details) (Supplementary Data File 2) for which we found overlapping evidence on the gene level for 59 associations (see Supplementary Data File 3). A total of 487 unique genes were found across all associations (see Supplementary Data File 4). The strength of association linking preclinical and clinical AEs induced by drug treatment, measured as the number of shared off-target drug activities, is shown in Fig. 1. The

association with the highest number of unique genes was the association of preclinical FOETAL MALFORMATION with clinical NERVOUS SYSTEM DISORDERS supported by 310 different genes, with the high number of genes involved likely to be due to the broad nature of the term NERVOUS SYSTEM DISORDERS, as well as overlap between pathways in developmental and CNS disorders [42]. Conversely, for some associations there were very few genes found between preclinical and clinical AEs, for example, the associations of preclinical BLOOD PROLACTIN INCREASED with BLOOD GROWTH HORMONE INCREASED was linked by only one gene for four drugs (chlorpromazine hydrochloride, clozapine, haloperidol and quetiapine fumarate), Peroxisome proliferator-activated receptor gamma (PPARG), known to play a key role in lipid metabolism [43], a biological process indeed modulated by both prolactin and growth hormone. Among the four drugs which had activity at PPARG and supported this association, quetiapine fumarate, clozapine and haloperidol have reported AEs associated with thyroid, endocrine and lipid disorders[21] In addition, drugs in PharmaPendium with on-target PPARG activity included ciglitazone, darglitazone, troglitazone, englitazone, pioglitazone hydrochloride and rosiglitazone maleate, the latter three of which have been found to cause lipid metabolism related diseases and hypothalamo-pituitary disorders in humans and rat[21]. This suggests that the association found could be either a direct link to lipid metabolism or to the hypothalamic-pituitary-thyroid (HPT) axis. The number of drugs driving each association as well as the proportion of drugs for which shared off-target drug activities were identified, is reported in Fig. S1. We also observe links between other AEs for which off-target drug activities were of particular relevance, including the association of preclinical ABNORMAL REFLEXES with clinical POOR QUALITY SLEEP, the association of PRECLINICAL PSYCHOMOTOR HYPERACTIVITY AND CLINICAL DRUG WITHDRAWAL SYNDROME and the associations of preclinical MUTAGENIC tests with clinical TERATOSPERMIA and OVARIAN FAILURE which are covered in the discussion.

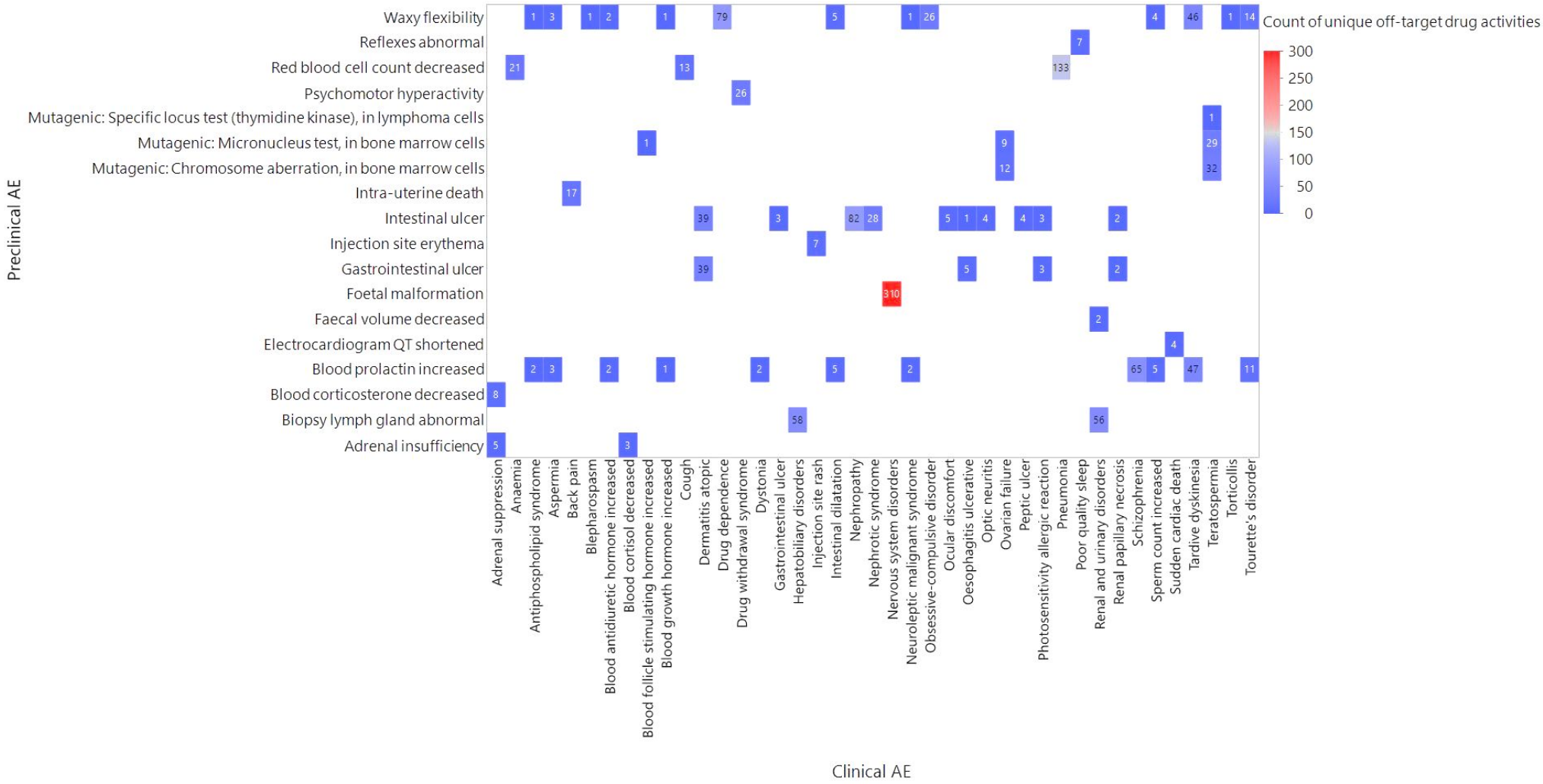


Fig. 1: Association strength between preclinical and clinical AEs measured by the number of unique shared off-target drug activities. The plot is coloured by the number of unique genes overlapping between drug targets and genes involved in phenotypes in both preclinical and clinical AE space. The minimum number of genes was 1 and the maximum number of genes was 310.

We were able to identify off-target drug activities for only about 25% of the studied links between preclinical and clinical AEs. However, those links are now in turn more likely to be biologically meaningful, and a potential biomarker/safety target hypothesis for the respective AE can be derived.

We found that the number of associations with overlap between preclinical, clinical and drug space amounted to only about 25 % of our initial associations, which whilst low offers increased confidence into the connections identified. This information can now be used to identify preclinical readouts whose link to a clinical readout of interest is biologically supported, to identify preclinical safety models that go beyond simple term and observation matching between the preclinical and clinical domain.

Identifying novel candidate safety targets for clinical toxicities

We next assessed the correspondence between these genes and the published safety target panel associated genes to determine the ability of our analysis to find known safety targets. In total, our analysis identified 56 out of 73 (77 %) of known safety target genes (Fig. 2), as defined from the combination of the panels published in Lounkine *et al.,* 2012 [11] and Bowes *et al.*, 2012 [13], which was high compared to a mean of 16 targets out of 73 (21 %) retrieved from repeated random samples of 487 targets with an active data point in ChEMBL (Fig. S2). Some of these safety endpoints are probably not target-related (or only partially) hence we do not expect links for all of them [44]. Certain genes can be associated with groups of (as opposed to individual) clinical AEs. For example, considering nervous system and psychiatric disorders, the terms DRUG DEPENDENCE, OBSESSIVE COMPULSIVE DISORDER, SCHIZOPHRENIA, TARDIVE DYSKINESIA and TOURETTE'S DISORDER were associated with many genes, including the Dopamine Receptor subtypes 1, 2 and 3, which are known to be associated with CNS adverse events [14], however, the precise link between these receptors and OBSESSIVE COMPULSIVE DISORDER, SCHIZOPHRENIA and TOURETTE'S DISORDER was not reported [14]. The AR and ESR1 genes, which encode the Androgen Receptor and Estrogen Receptor 1 respectively, are promiscuous across clinical AEs, due to their roles in a wide range of biological processes, specifically linking to endocrine and CNS adverse events [14]. Similarly, the GCR gene which encodes the Glucocorticoid Receptor was also found to be associated with many of the clinical AEs, in

particular a range of CNS adverse events, which have not been previously identified [14]. More specific gene-clinical AE associations from our analysis included the Acetylcholinesterase (ACES) and Tyrosine-Protein Kinase LCK (LCK) genes associated with PNEUMONIA. Importantly, ACES has been previously associated with respiratory adverse events [14]. In addition, the Neuronal Acetylcholine Receptor Subunit Alpha-4 (ACHA4) gene was associated with POOR-QUALITY SLEEP, the Neuronal Acetylcholine Receptor Subunit Alpha-7 (ACHA7), the Neuronal Acetylcholine Receptor Subunit Beta-4 (ACHB4), the Substance-P Receptor (NK1R) and the Glutamate Receptor Ionotropic NMDA 2A (NMDE1) genes were associated with DRUG WITHDRAWAL SYNDROME and the Steroid Hormone Receptor ERR2 (ERR2) and Gamma-aminobutyric acid Receptor Subunit Alpha-5 (GBRA5) were associated with NERVOUS SYSTEM DISORDERS, none of which are summarized in the previous study [14], representing novel associations between previously identified safety target encoding genes and adverse events. Hence, these findings support the ability of the approach we developed to identify meaningful associations. The remaining 431 genes which are linked by preclinical and clinical AE observations as well as a genetic basis are not currently used in secondary pharmacology panels, and they are provided to the reader as a resource (Supplementary Data File 5) for further evaluation.

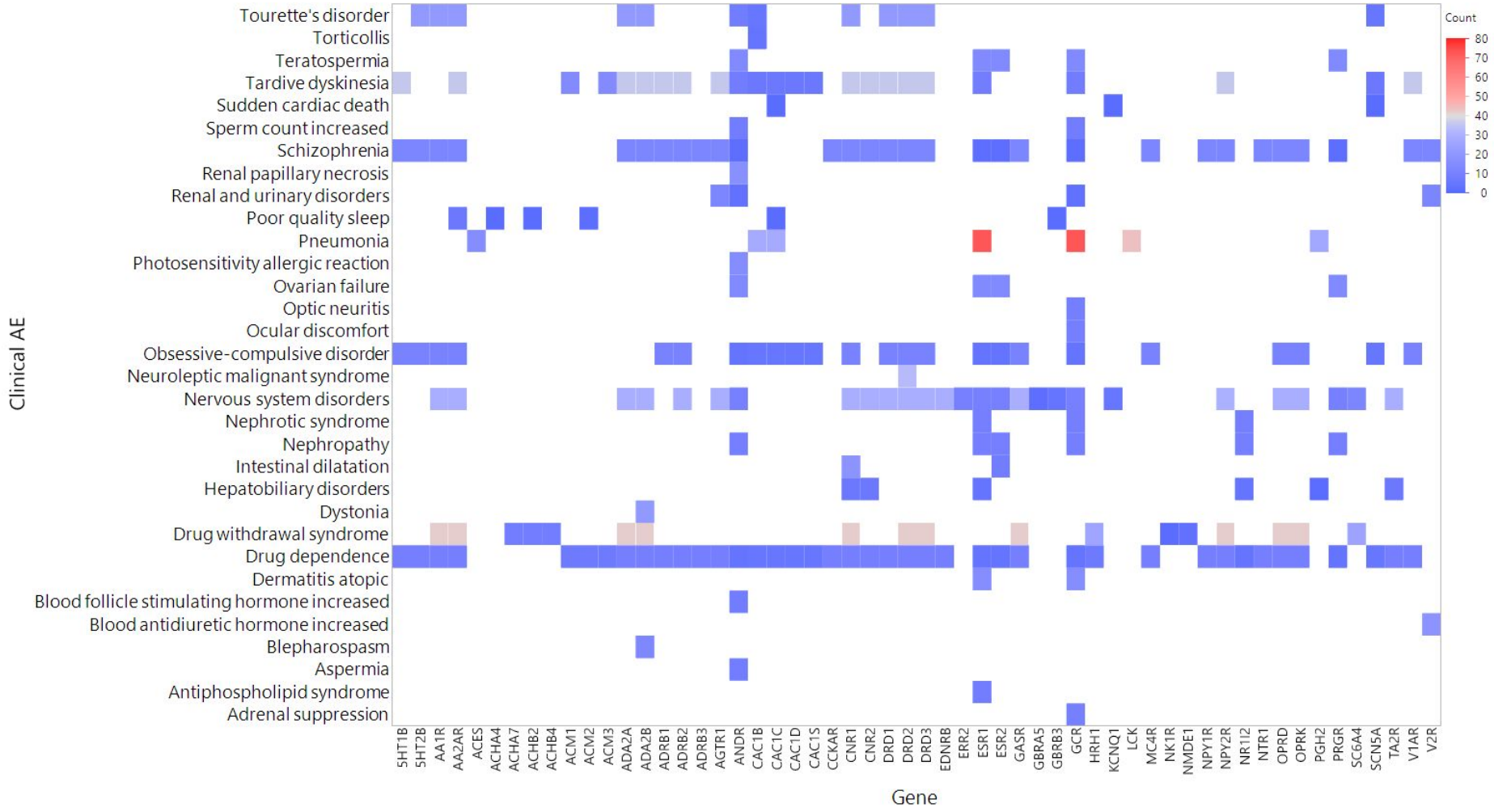


Fig. 2: Genes whose resulting protein products are targeted by drugs, are associated with toxicities on the preclinical and clinical level and overlap with known safety targets from the published studies by Lounkine et al., 2012 [11] and Bowes et al., 2012 [13]. Matrix elements are coloured by the number of times the gene was associated with each clinical AE. The most promiscuous genes across AEs, included the Androgen Receptor (ANDR), the Estrogen Receptor 1 (*ESR1*) and the Glucocorticoid Receptor (*GCR)*. Importantly, some of the safety targets are not target-related so we do not expect clear links for all of them [44].

Discussion

Harm-benefit analysis of the use of animal models for drug toxicity testing has long been discussed in the field of pharmaceutical research with the conclusion that in some cases, such as QT prolongation and arrhythmias, predictivity seems rather high [18,23], while in others, such as neurological toxicities and cutaneous toxicities, this does not appear to be the case [24]. This assessment of the utility of animal studies is guided by the principles of the 3Rs to replace, reduce or refine the use of animals in research to protect animal welfare [10] whilst still using animals in research, where the toxicity endpoint is predictive of human toxicity to the extent that it can be used to inform decisions for drug progression to the clinic. Previous concordance studies have led to data-driven conclusions for the risk of experiencing clinical AEs given *similar* preclinical AEs. Despite the recognition that AEs manifest differently between animals and humans, for reasons including differences in anatomy and physiology between species, as well as differences on the genetic and cellular level, no studies have yet taken advantage of the opportunity to associate preclinical AEs with substantially different clinical AEs. The method presented here has now extended previous approaches in such a way and identified 2,952 new associations (Supplementary Data File 2) between AEs which were encoded by *different* terms preclinically and clinically. Firstly, this highlights the scale on which we were able to find novel predictive links between preclinical AEs and clinical AEs. Secondly, we showed that there were a large number of clinical AEs *which are better predicted* by different preclinical AEs than *via* their corresponding term preclinically. We observed that 57 % of the generated associations have a positive Likelihood Ratio of greater than 2, which indicates a significant change in risk of experiencing the clinical AE given the preclinical AE. This ratio can be used to assess the diagnostic ability of preclinical AEs for predicting a clinical outcome, which can be used in safety assessment cases. Acknowledging that statistical associations can be spurious,

we next explored off-target drug activities shared by concordant AEs, with the aim to both focus on biologically meaningful relationships between preclinical and clinical AEs and also to identify a way of supporting the choice of novel safety targets in this way. Interestingly, we identified a total of 487 off-target drug activities including 56 of the 73 (77%) known safety targets currently used in in vitro safety panels [11,13] supporting the ability of the approach we developed to identify meaningful off-target activities. In addition, we identified 431 genes which we propose could be investigated as future safety target encoding genes for different clinical toxicities. Among the associations between preclinical and clinical AEs along with the associated off-target activities identified, we identified a few examples to be further discussed for the relevance of the clinical adverse effect, which include the association between preclinical ABNORMAL REFLEXES and clinical POOR QUALITY SLEEP, preclinical PSYCHOMOTOR HYPERACTIVITY and clinical DRUG WITHDRAWAL SYNDROME and a range of preclinical mutagenicity readouts with both clinical TERATOSPERMIA and OVARIAN FAILURE (Fig. 3).

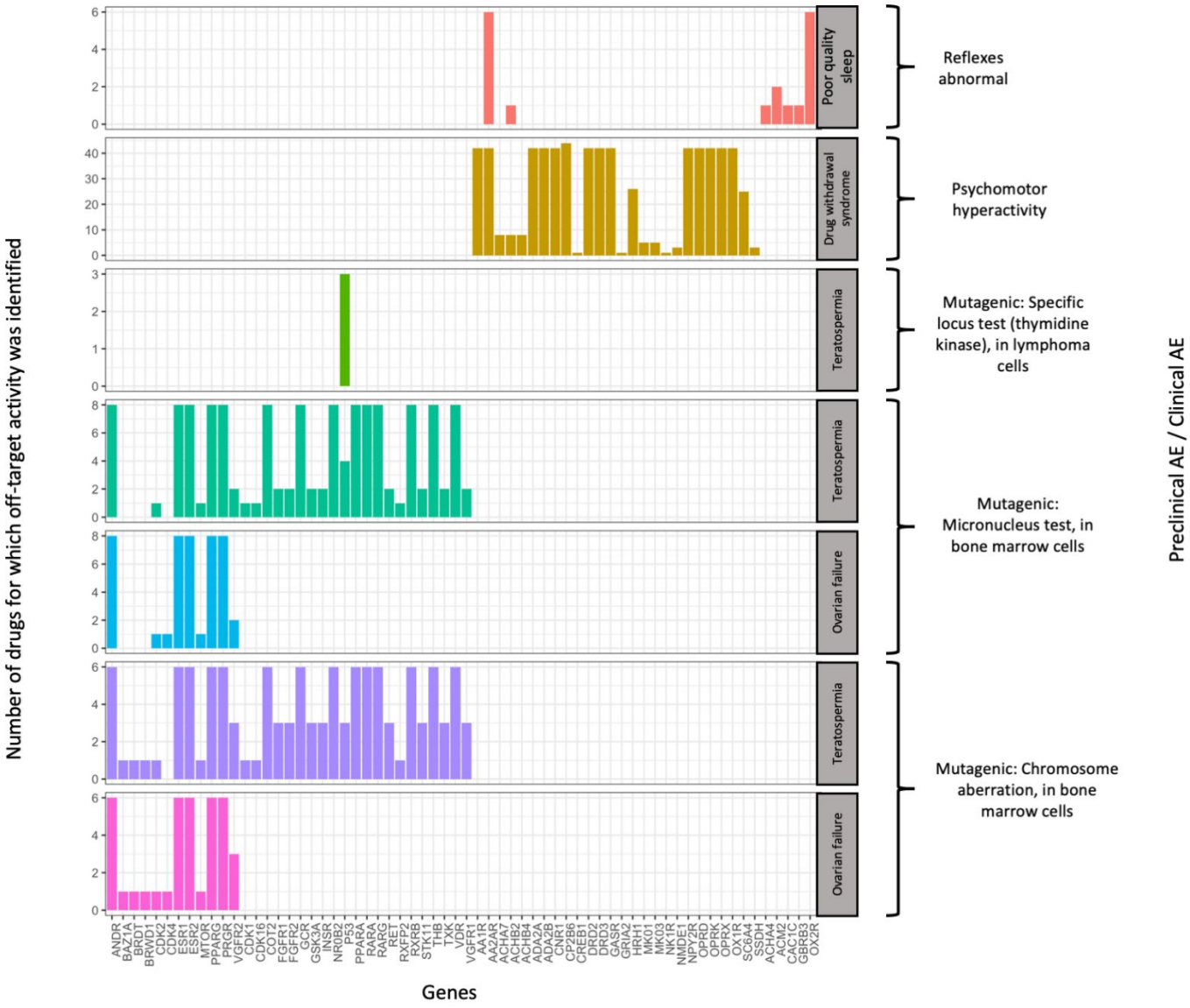


Fig. 3: Targets for the associations between mutagenic preclinical adverse events and clinical TERATOSPERMIA and OVARIAN FAILURE, between REFLEXES ABNORMAL and POOR QUALITY SLEEP and between preclinical PSYCHOMOTOR HYPERACTIVITY and clinical DRUG WITHDRAWAL SYNDROME. Histograms shows the number of drugs for which off-target activity against the listed genes was identified. As expected, many targets overlap when considering preclinical mutagenic tests, especially the two estrogen receptors (ESR1 and ESR2), the androgen receptor (ANDR), the progesterone receptor (PRGR) and the peroxisome proliferator-activated receptor gamma (PPARG).

Preclinical REFLEXES ABNORMAL are associated with clinical POOR-QUALITY SLEEP via multiple genes/protein targets

Many pharmaceutical drugs have been shown to induce sleep disorders including insomnia, daytime sleepiness, nightmares and changes in the sleep architecture, ultimately affecting the

quality of life [45]. Animal models exist for assessing the risk of poor-quality of sleep in patients; however, many differences in sleep patterns exist between animals and humans [46]. In our study we identified the association between the preclinical AE REFLEXES ABNORMAL with clinical AE POOR-QUALITY SLEEP with a LR^+ of 17.3 (Fig. 3 and Supplementary Data File 2). This association is driven by the observation of BOTH side effects for 10 individual drugs in the dataset. Of these 10 drugs, 7 were found to share at least one gene target with the two phenotypes (Supplementary Data File 3). The most common targets found for this association were the Adenosine A2a receptor (AA2AR), annotated as a target of six drugs and Orexin 2 receptor (OX2R), also annotated as a target of six drugs. Studies of A2aR-knockout mice show that AA2AR is linked with a decrease in sensitivity to painful stimuli, which can lead to reduction in the reflex response [47]. AA2AR also plays a role in sleep directly, and mice lacking the functional Adenosine A2a receptor no longer show increased wakefulness in response to caffeine [48]. In humans, polymorphisms in the A2a receptor are associated with impaired sleep, showing a role for A2a receptors in sleep quality directly [49]. The orexin receptors are well-known for their link with sleep and wakefulness cycles [50–52]. in particular the conditions of insomnia and narcolepsy [53]. More recently, orexin has also been linked to antinociceptive effects, showing a role in neuropathic pain modulation, which may have a role in the reflex response to pain [54]. While the A2a receptor is part of previously published safety panels, OX2R is currently not found in those [55]. However, in particular where sleep modulation is a concern during drug development, including OX2R in screening panels would be warranted according to the current analysis.

Preclinical PSYCHOMOTOR HYPERACTIVITY is associated with clinical DRUG WITHDRAWAL SYNDROME via multiple genes/protein targets

Drug withdrawal syndrome presents a real challenge to patients and in some cases can be life-threatening.[56] From our analysis, PSYCHOMOTOR HYPERACTIVITY in the preclinical setting is significantly associated with clinical DRUG WITHDRAWAL SYNDROME, with a LR^+ of 6.2 ((Fig. 3 and Supplementary Data File 2). In total, 53 drugs were found which induced both phenotypes, of which 49 had one or more gene overlapping between itself and the two AEs in the association. The genes whose protein targets were modulated by the highest number of drugs for this association included CP2B6, AA1R, AA2AR, ADA2A, ADA2B, CNR1, DRD2, DRD3, GASR, NPY2R, OPRD, OPRK, OPRX, OX1R, which were targets for at least 42 of the drugs supporting the association. Many of the links identified from the data were supported by biological knowledge. For example, the Cytochrome P450 family 2 subfamily B member 6 (CP2B6) polymorphisms, in combination with attention deficit hyperactivity disorder (ADHD) symptoms, are found to be linked to nicotine addiction.[57] Antagonism of the Adenosine receptors (AA1R and AA2AR) leads to psychomotor phenotypes from the reduction in adenosine and regulation of genes in the striatum signalling pathway.[58] AA1R and AA2AR agonists have been linked to withdrawal symptoms for benzodiazepine drugs,[59] and agonism of these receptors antagonistically modulates dopaminergic neurotransmission and therefore reward systems.[60] Additionally, the Cannabinoid receptor 1 (CNR1) and Dopamine receptors (DRD2, DRD3) have well-known links to ADHD, neuropsychiatric disorders and substance withdrawal, as do the Opioid receptors (OPRD, OPRK, OPRX).[61–64] Cholecystokinin B receptor (GASR) mutations are related to behavioural changes in animals[65] and also affect morphine induced hyperactivity and withdrawal symptoms.[66,67] Finally, the Orexin 1 (OXR1) receptor is involved in ADHD as well as naloxone-precipitated morphine withdrawal.[68–70]

Our analysis found a link between preclinical PSYCHOMOTOR HYPERACTIVITY and clinical DRUG WITHDRAWAL SYNDROME, supported by multiple genetic links. Of the targets identified, the majority are present in existing safety panels, however, we also identified three targets, namely CP2B6, OPRX and OX1R, which are not routinely screened according to published in vitro safety panels, (although CP2B6 is screened in drug metabolism and pharmacokinetic (DMPK) assays for potential drug-drug interactions[71]). According to the links found here, these three new targets could be further investigated for future safety panel inclusion.

Preclinical mutagenicity tests are associated with clinical TERATOSPERMIA and OVARIAN FAILURE, supported via by multiple genes/protein targets

From our analysis, observations of preclinical mutagenicity using the MUTAGENIC: MICRONUCLEUS TEST, IN BONE MARROW CELLS test, MUTAGENIC: CHROMOSOME ABERRATION, IN BONE MARROW CELLS and MUTAGENIC: SPECIFIC LOCUS TEST (THYMIDINE KINASE), IN LYMPHOMA CELLS were associated with clinical TERATOSPERMIA and OVARIAN FAILURE with LR+ values of 20.7, 23.1, 28.9 and 17.2, 19.2, 24, respectively (Fig. 3 and Supplementary Data File 2). This showed a very high conditional likelihood for experiencing TERATOSPERMIA and OVARIAN FAILURE in clinical trials, given a positive readout in either of these preclinical tests. Interestingly, the main gene present for all associations was the human ANDR gene. This gene (the Androgen receptor) is well-known to be linked to TERATOSPERMIA, as it has a vital role in spermatogenesis and mutations in the gene lead to male infertility [72–74]. OVARIAN FAILURE is linked to the absence of the androgen receptor in mice and serum androgen levels are elevated in women with ovarian failure [75,76]. Many of the other 32 genes for these associations are related to DNA damage or repair, apoptosis, cellular proliferation, angiogenesis, methylation effects and transcriptional regulation and are not shared by all the associations. Overall, we present the genes which were found to support our associations between mutagenic preclinical AEs and the clinical AEs of TERATOSPERMIA and OVARIAN FAILURE, providing protein targets which

could be investigated further for their roles in these serious reproductive toxicities of drugs. Despite the association between preclinical mutagenicity tests and clinical TERATOSPERMIA and OVARIAN FAILURE being known, only five out of 33 identified genes are currently routinely screened as *in vitro* safety panels, namely ANDR, ESR1, ESR2, GCR and PRGR. The remaining 28 genes/their protein products, namely BAZ1A, BRDT, BRWD1, CDK1, CDK16, CDK2, CDK4, COT2, FGFR1, FGFR2, GSK3A, INSR, MTOR, NR0B2, P53, PPARA, PPARG, RARA, RARG, RET, RXFP2, RXRB, STK11, THB, TXK, VDR, VGFR1 and VGFR2, are not currently routinely used within published safety screening panels but their future incorporation could be investigated further..

Conclusion

In this work we developed an approach to discover statistically significant associations between preclinical and clinical AEs caused by drugs reported in the PharmaPendium database across different MedDRA terms. In addition, by investigating the genetic information shared between the drugs and both the preclinical and clinical AEs we identified plausible and testable candidate off-targets drug activities which are not currently tested in *in vitro* screening and that we propose for investigation as future safety targets for different clinical toxicities. The relevance of the proposed targets is supported by the meaningful associations identified between mutagenic preclinical AEs and the clinical AEs of TERATOSPERMIA and OVARIAN FAILURE, between preclinical REFLEXES ABNORMAL and clinical POOR-QUALITY SLEEP and between preclinical PSYCHOMOTOR HYPERACTIVITY and clinical DRUG WITHDRAWAL SYNDROME. Notably, we were not able to identify off-target drug activities for up to 75% of the associations identified and this could be due to multiple reasons. First and most importantly, many toxicity events are simply not linked to proteins directly [77]. Secondly, the individual AEs

could be caused by separate and distinct mechanisms of the same drugs (i.e., two AEs are produced by different off-targets of the same drug) and therefore we find statistical, but no causal associations between (mechanistically distinct) AEs. Last, in some cases we were unable to map the AE terms to a representative disease or phenotype due to a lack of knowledge about the involvement of genes in the linked diseases or phenotypes, as annotated in the databases used here. When interpreting the results of this study there are certain limitations that should be noted related to the dataset and the methods used. Firstly, this analysis did not consider the dosing of drugs administered for both the preclinical and clinical studies due to the lack of this information in the PharmaPendium database. Therefore, the dosages in preclinical studies were assumed to be relevant to the human dose in clinical trials and plasma concentration assumed to be similar across species. Despite our analysis not considering dose, our approach led to the discovery of novel targets of toxicity which dose-response relationship can be investigated *in vitro*. Furthermore, this analysis does not consider the frequency or the number of animals or humans which experienced the toxicity, instead encoding the presence or absence of an AE across all subjects and studies conducted for each drug. This is because this quantitative information was not captured within the PharmaPendium database. With regards to the off-target drug activities analysis, we did not discriminate on the level of association required for a gene to be associated with a phenotype or disease because we wanted to identify all plausible links which could be used for hypothesis generation. Despite these limitations, this study has been able to identify new associations between preclinical and clinical AEs along with a set of novel targets of toxicity that can be further explored before being considered for *in vitro* secondary pharmacology screening panels. In addition, the approach developed in this work for the identification of associations between preclinical and clinical toxicities, can be applied for similar datasets to determine the interrelationship between toxicological data and to evaluate which preclinical models are useful to assess the risk of experiencing clinical

toxicities. We furthermore see the utility in hypothesis generation methods such as this for finding targets of toxicity for inclusion in toxicity panel screens. This analysis mainly reports one-to-one associations however, as demonstrated for reproductive toxicity, drug-induced toxicities are interrelated in the preclinical and clinical spaces. Future work can explore the multi-item nature of associations in more detail. Finally, this type of analysis could provide semantic reasoning between AEs, such as the mechanistic relationships between MedDRA terms, as well provide a basis for machine learning models of clinical toxicity.

Materials and Methods

Dataset

Preclinical and clinical adverse event (AE) data encoded by the Medical Dictionary for Regulatory Activities (MedDRA) [78] preferred term (PT) were manually extracted from PharmaPendium (2017-04) [21] for all drugs in the database. These drugs were filtered to only retain those drugs which had at least one reported AE in a preclinical and a clinical study, resulting in 2,259 drugs. Duplicate entries of the same drug and AE combination were removed, retaining only one instance of each pair. The dataset was converted into a binary matrix of AEs against drugs where presence of an AE for a drug was encoded by 1 and absence encoded by 0. In total, 4,585 preclinical AE variables and 7675 clinical AE variables were extracted. The breakdown of reported preclinical AEs by species is shown in Fig. S3, which shows that rat, dog, mouse, rabbit and monkey dominate animal models of toxicity in this database.

Feature filtering

To avoid the biases of calculating associations between variables with small classes, near-zero variance AE features were removed using the VarianceThreshold function from

sklearn.feature_selection in Python [79], using the variance for Bernoulli random variables = p(1-p), where p=0.99. We experimented with different thresholds, however, this value was chosen as any lower probability led to a large drop off in the number of variables retained (Fig. S4). 751 preclinical AEs and 1,740 clinical AEs remained after filtering, with a minimum and maximum frequency of 23 and 1,862 times respectively a given AE was present across drugs.

Mutual Information Associations

Concordance analysis

The normalized Mutual Information (MI) [80] between each preclinical to clinical AE encoded by the same MedDRA term was calculated using normalized_mutual_info_score from sklearn.metrics in Python [79]. Here we use the MI as a measure of the dependence between the preclinical and clinical variables, as opposed to other methods of correlation, because it can measure the general dependence rather than the linear dependence between two variables and doesn't depend on the exact values but the probability distribution of the variables [81].

The two-tailed Fisher's exact test [82] implemented using *scipy.stats.fisher_exact* in Python [83] was used to assess the significance of the associations with sample size of 2,259 drugs and using a cut-off of 0.01 for the Bonferroni corrected p-value (number of tests=473), to reduce the type 1 error.

Statistical Associations between all preclinical and clinical AEs

The same method as above was used to calculate the values for the MI between all preclinical AEs and all clinical AEs, regardless of the MedDRA term. For each clinical AE, the largest three preclinical MI scores were retained.

Assessing significance

The two-tailed Fisher's exact test [82] implemented using *scipy.stats.fisher_exact* in Python [83] was again used to assess the significance of these associations sample size of 2,259 drugs and using a cut-off of 0.01 for the Bonferroni corrected p-value (number of tests=5220). Furthermore, for each individual preclinical and clinical AE the binary labels were randomised using *rv_discrete* from SciPy, preserving the computed probability of observing the AE for each of the vectors and generating randomised vectors of same length as the real vectors (comprising 2,259 drugs). Then the normalized mutual information was calculated as before for the random set of vectors for all preclinical to clinical AEs, repeated 10 times. We took the 99th percentile of the distribution of randomised MI values which was 0.011 as a cut-off for significant associations between the real data; any associations below this value were removed from the set. We use stringent cut-offs in both approaches to reduce the false positive rate of our newly discovered associations, hoping to avoid spurious associations, which are likely with large datasets such as these. In total, after applying the Fisher's exact test and the randomisation derived cut-off, 3,011 significant associations were identified.

Quantifying risk

In order to quantify the risk determined by these associations we employ the Likelihood Ratios. The positive and negative Likelihood Ratios [84] (LR^+, LR^-) were calculated to assess the risk of experiencing a clinical AE given the presence of a preclinical AE and the likelihood of absence of a clinical AE given the absence of a preclinical AE. The following formulas were applied:

$$LR^+ = \frac{sensitivity}{1 - specificity}$$

$$LR^- = \frac{1 - sensitivity}{specificity}$$

The Likelihood Ratio has been argued to be a better indicator of concordance for this type of analysis than the traditional sensitivity metric, as it takes into account the false negatives and

the false positive values as well as providing extra interpretation [85,86]. A positive or negative Likelihood Ratio of 1 means that there is no useful information in the probability of the preclinical finding to predict the clinical outcome, whereas a positive Likelihood Ratio (LR+) of greater than 10 shows a large and conclusive change in probability of a clinical AE given the probability of the preclinical AE. Conversely, for the negative Likelihood Ratio (LR-), values closer to 0 are more useful to determine the likelihood of the absence of a preclinical finding predicting the absence of the clinical AE [18,87]. We therefore use this as an additional measure of directionality of our associations.

Targets of Toxicity Identification Between Preclinical and Clinical AEs

Associations

Associations for Interpretation

For the purposes of interpretation, we limited our associations to only the most significant, chosen by applying a cut-off for the mutual information (MI) score of greater than 0.095, which was the highest value for the randomised distribution. This left 248 highly significant associations which were next subject to the workflow shown in Figure S6 in order to investigate for evidence of biological drivers.

Extracting targets for drugs

First, we set to retrieve toxicity targets for each of the drug contained in the PharmaPendium [21] database (Fig. S5, Step 1). All 2,259 drugs were standardised from their SMILES strings using the StandardiseMolecules function camb package in R [88] which removes salts. The standardised structures were then converted to InChIKeys using KNIME [89], to map the drugs to their target activities in other databases. Targets for the drugs in PharmaPendium were extracted from three main sources, namely the AstraZeneca ChemistryConnect database

(which contained Bioprint (2007 snapshot) [90], ChEMBL-23 [91] and GOSTAR (GVK Bio) data), DrugTargetCommons [92] and SuperDrug2 [93]. Data was extracted from ChemistryConnect by querying the database using a synonym search from the drugs in the PharmaPendium database. Compounds extracted were standardised using the same process as for the PharmaPendium drugs, and then matched by InChIKey [94] to the PharmaPendium drugs. It was necessary to carry out a second match on exact drug names between the databases, as some of the drugs were non-small molecule drugs for which SMILES were not available. Targets in ChemistryConnect were encoded by EntrezGeneID [95] which was mapped to their UniProtKB Accession IDs using the Uniprot Identifier exchange service [96]. The data in ChemistryConnect is already categorized into the classes of active and inactive based on the cut-off of 10 μM for endpoints including *Ki*, *Kd*, IC50 or % inhibition at 10 μM. We retained the active entries for our analysis. Whilst a relatively high cut-off, we chose this activity cut-off as we wanted to include as much off-target information for drugs as possible and many of the panel screens were conducted at a concentration of 10 μM. Data from the SuperDrug2 [93] database was provided by the curators. Compounds extracted were standardised using the same process as for the PharmaPendium drugs, and then matched by InChIKey and exact name matching to the PharmaPendium drugs. The targets in this database were already encoded in UniProtKB Accession IDs. DrugTargetCommons [92] data was downloaded from the web platform (https://drugtargetcommons.fimm.fi/). Since the data did not have SMILES, but did include ChEMBL IDs, we matched the database to ChEMBL-23 using these IDs to obtain SMILES. For this purpose, the ChEMBL IDs and SMILES for all ChEMBL-23 compounds were extracted using Toad for MySQL [97] and then merged with the ChEMBL IDs in Drug-Target Commons. The resulting compounds were standardised using the same process as for the PharmaPendium drugs, and then matched by InChIKey and exact name matching to the PharmaPendium drugs. For the data in DrugTargetCommons, those with an "Active" flag in

the activity comment column were retained, where a cut-off of 10 μM was used to define activity and only those records for which there was *Ki*, *Kd*, IC50 or % inhibition at 10 μM were retained. The targets in this database were already encoded in UniProtKB Accession IDs. Active drug-target interaction data from all sources was combined into one database, retaining PharmaPendium drug names, UniProtKB Accession IDs, gene names, gene symbols where possible and an external reference key to the original database. Duplicates of PharmaPendium drug name and UniProtKB Accession ID were removed. Overall 82,459 positive drug-target interactions were extracted for 1604 out of 2,259 drugs in the PharmaPendium dataset, which included 7,533 unique targets. The breakdown across databases was as follows: 1,397 drugs mapped to ChemistryConnect, 1,511 drugs mapped to DrugTargetCommons and 1,253 drugs mapped to SuperDrug2. A total of 4,863 out of 7,533 targets were found to interact with multiple drugs and 1,549 out of 1,604 drugs had more than one active target, and hence this combined dataset was found to be encoding more than just the primary target activity.

Mapping preclinical and clinical AEs to ontology terms

For each preclinical or clinical AE which was found as part of the 248 statistically significant associations generated from the analysis, the MedDRA encoded AE terms were mapped to ontology IDs using Zooma [98], identifying phenotypes and diseases that describe the AE (Fig. S5, Step 2). The ontologies searched in Zooma included the Human Disease Ontology (DOID) [99], Mammalian Phenotpye Ontology (MP) [100], Human Phenotype Ontology (HP) [101], Experimental Factor Ontology (EFO) [102], Orphanet Rare Disease Ontology (ORDO) [103], National Cancer Institute Theasurus (NCIT) [104], Ontology of MIRNA Target (OMIT) [105], Ontology of Adverse Events (OAE) [106], Monarch Merged Disease Ontology (MONDO) [107], Symptom Ontology (SYMP) [108], Mental Disease Ontology (MFOMD) [109], Mouse Pathology Ontology (MPATH) [110], Ontology of Biological Attributes (OBA) [111] and BioAssay Ontology (BAO) [112]. The mapping results were manually filtered to leave only terms which have the

same meaning as the MedDRA encoded AE term. This resulted in 183 Zooma mappings for preclinical AEs and 405 mappings for clinical AEs. Not all AEs were mapped to ontology identifiers. The list of Zooma Ontology terms mapped from PharmaPendium MedDRA terms are in Supplementary Data Files 6 and 7.

Extracting genes for preclinical and clinical AEs

From these mappings, the OpenTargets (version 3.4) [113] database was queried using the OpenTargets Python client [114] to extract the genes associated with diseases encoded by the EFO and DOID ontologies (Fig. S5, Step 3). The genes found were encoded by Ensembl gene (ENSG) IDs [115] from OpenTargets. In total 22,056 genes were mapped to preclinical AEs and 44,718 genes were mapped to clinical AEs. For the phenotype ontologies including HPOs, MPOs, NCITs, Orphanet, as well as others, genes were extracted from the Monarch Initiative [116] using the requests library in Python [117] to import the data from the URL for the matched ontology ID. In total 26,216 genes were mapped to preclinical AEs and 50,415 genes were mapped to clinical AEs. The output gene IDs were HUGO Gene Nomenclature Committee (HGNC) symbols [118] for human genes and the relevant non-human gene IDs for other organisms including the Mouse Genome Informatics (MGI) [119] IDs.

For disease ontologies including DOIDs and MONDO IDs, genes were extracted from the Monarch Initiative [116], using the requests library in Python [117] to import the tsv file from the url for the matched ontology ID. In total 9,252 genes were mapped to preclinical AEs and 35,239 genes were mapped to clinical AEs. The output gene IDs were as for the previous step. Gene IDs from all sources were then mapped to Uniprot KB identifiers using the Uniprot Identifier exchange service. The genes were mapped back to their original AE term. Duplicate UniprotKB Accession IDs for each AE were removed. In total, 242,997 gene-preclinical AE pairings were found across all preclinical AEs and 546,902 gene-clinical AE pairings were found for clinical AEs.

Gene filtering and mapping animal genes to human orthologs

Next, we set out to map the UniprotKB gene IDs associated with drugs, preclinical AEs and clinical AEs to human ortholog (Fig. S5, Step 4). For preclinical AE genes, we retained those which were from non-human species while for clinical AE genes we retained only human genes. For the drug-gene associations the non-human genes were separated from the human genes for the next step. Next, for the preclinical AE-associated non-human genes, we mapped their gene identifiers to their orthologs in humans, utilising the Uniprot Identifier exchange service to map to Eggnog [120] identifiers and then back to human UniprotKB genes *via* the REST API [96]. We chose the Eggnog database due to the higher level of overlap with UniprotKB identifiers than other ortholog mapping databases. The non-human genes for drugs were also mapped using the same method to their human orthologs, meaning all genes were mapped to human UniprotKB identifiers for the overlap analysis.

Overlap analysis of genes from preclinical AE, clinical AE and association driving drugs

Finally, we set to explore the genetic overlap between all the drug associated genes, clinical AE associated genes and the preclinical AE associated gene orthologs (Fig. S5, Step 5). The following analysis was conducted in Python utilising the pandas [121] library. The overlap analysis was designed to check whether, for each association, the drugs which displayed both the preclinical and clinical endpoints possess protein targets which are encoded by genes associated with both the preclinical and clinical endpoint, and which could hence mechanistically be associated with AEs in both animals and humans. All matches of a drug with protein targets which are present in the genes associated with both preclinical and clinical AEs were retained for subsequent analysis. For each association, the intersection between genes for the preclinical and clinical AE terms was extracted as a list of genes. As mentioned

above, the overlap analysis discussed in this study was conducted by only using human genes found from databases for matching to clinical AEs and the orthologs of animal genes found from databases for matching to preclinical AEs. We also conducted the same analysis agnostic of species, by not filtering out animal genes for clinical AEs and human genes for preclinical AEs respectively, and refer the interested reader to Supplementary Data File 9, which reports these results.

Comparison of Mechanistic Targets to *In Vitro* Safety Screening Panels

In vitro target safety panels were extracted from Lounkine *et al.,* 2012 [11] and Bowes *et al.*, 2012 [13] which represent the Novartis Safety Target panel and Panel-44 respectively. The genes found from the overlap analysis were compared to the genes represented by the safety targets in the combined panels and flagged as either known safety target genes (if present in the panels) or potentially novel safety target genes (if not present in the panels). To assess how much better our method was at identifying known safety targets than a random selection of proteins associated with active compounds, we simulated the sampling of 487 targets with known actives from ChEMBL and compared these targets to the safety panel targets. Sampling was repeated 1000 times to produce a distribution of the number of matches to safety panel targets (Fig. S2), from which the mean was calculated to be 16, identifying 22 % of known safety panel targets on average.

Visualisations

Visualisations were generated in TIBCO spotfire [122] and R.

Acknowledgements

General: The authors would like to thank Ines Smit, Dezső Módos, Tobias Noeske, Carrie Larner, Ola Engkvist and Willem Nissink for their useful discussions and input into the project.

Funding:

This work was supported as part of the PhD project of K.A.G funded by the European Research Council and AstraZeneca Early Oncology TDE.

Author Contributions:

K.A.G., N.G., I.B. L.R. contributed to the overall analysis design and database selections. K.A.G. devised and conducted the analysis. K.A.G and D.B wrote the manuscript. A.B. and S.J.H. supervised the study, contributed to project direction and manuscript preparation. A.M.A advised on the statistical methods. L.R. contributed to the biological interpretation.

Competing Interests:

N.G., I.B., L.R., S.J.H., A.M.A. K.A.G and A.B are employees at AstraZeneca.

Data and Materials Availability:

All data needed to support conclusions are either available in the main text or the supplementary materials, including Supplementary Data Files of the full lists of associations and gene overlap analyses.

Supplementary Materials

Supporting Information:

Fig. S1: Number of drugs with gene overlap for each significant association

Fig. S2: Distribution of the number of matches to safety panels from 1000 random samples of 487 targets from ChEMBL

Fig. S3: Breakdown of the number of times each animal species was reported for adverse events

Fig. S4: Feature selection for preclinical and clinical adverse events

Fig. S5: Schematic to show the method for the gene overlap analysis

Table S1: Summary of the previous concordance literature relating preclinical adverse events to clinical adverse events

Table S2: The 10 most statistically significant concordant adverse events as ranked by the Mutual Information

Supplementary Data Files: Can be found in separate attached excel spreadsheet tabs; these are larger data sets of results

File 1: Concordant associations

File 2: All Significant Associations

File 3: Gene Overlap Analysis Results

File 4: Unique Genes

File 5: Genes in Safety Panels

File 6: Preclinical Ontology Mapping

File 7: Clinical Ontology Mapping

File 8: Drug ATC Codes

File 9: Alternative Overlap Analysis

References

(1) Leenaars, Cathalijn H. C.; Kouwenaar, Carien; Stafleu, Frans R.; Bleich, André; Ritskes-Hoitinga, Merel; De Vries, Rob B. M.; Meijboom, Franck L. B. Animal to Human Translation: A Systematic Scoping Review of Reported Concordance Rates. *Journal of Translational Medicine*. BioMed Central Ltd. July 2019, p 223.

(2) Gintant, Gary; Sager, Philip T.; Stockbridge, Norman. Evolution of Strategies to Improve Preclinical Cardiac Safety Testing. *Nat. Rev. Drug Discov.* **2016**, *15*, 457–471.

(3) Kim, Eunjung; Rebecca, Vito W.; Smalley, Keiran S. M.; Anderson, Alexander R. A. Phase i Trials in Melanoma: A Framework to Translate Preclinical Findings to the Clinic. *Eur. J. Cancer* **2016**, *67*, 213–222.

(4) Olson, Harry; Betton, Graham; Robinson, Denise; Thomas, Karluss; Monro, Alastair; Kolaja, Gerald; Lilly, Patrick; Sanders, James; Sipes, Glenn; Bracken, William; Dorato, Michael; Van Deun, Koen; Smith, Peter; Berger, Bruce; Heller, Allen. Concordance of

the Toxicity of Pharmaceuticals in Humans and in Animals. *Regul. Toxicol. Pharmacol.* **2000**, *32*, 56–67.

(5) Shanks, Niall; Greek, Ray; Greek, Jean. Are Animal Models Predictive for Humans? *Philos. Ethics, Humanit. Med.* **2009**, *4*, 2.

(6) Waring, Michael J.; Arrowsmith, John; Leach, Andrew R.; Leeson, Paul D.; Mandrell, Sam; Owen, Robert M.; Pairaudeau, Garry; Pennie, William D.; Pickett, Stephen D.; Wang, Jibo; Wallace, Owen; Weir, Alex. An Analysis of the Attrition of Drug Candidates from Four Major Pharmaceutical Companies. *Nat. Rev. Drug Discov.* **2015**, *14*, 475–486.

(7) Cook, David; Brown, Dearg; Alexander, Robert; March, Ruth; Morgan, Paul; Satterthwaite, Gemma; Pangalos, Menelas N. Lessons Learned from the Fate of AstraZeneca's Drug Pipeline: A Five-Dimensional Framework. *Nat. Rev. Drug Discov.* **2014**, *13*, 419–431.

(8) Morgan, Paul; Brown, Dean G.; Lennard, Simon; Anderton, Mark J.; Barrett, J. Carl; Eriksson, Ulf; Fidock, Mark; Hamrén, Bengt; Johnson, Anthony; March, Ruth E.; Matcham, James; Mettetal, Jerome; Nicholls, David J.; Platz, Stefan; Rees, Steve; Snowden, Michael A.; Pangalos, Menelas N. Impact of a Five-Dimensional Framework on R&D Productivity at AstraZeneca. *Nat. Rev. Drug Discov.* **2018**, *17*, 167–181.

(9) Van Norman, Gail A. Limitations of Animal Studies for Predicting Toxicity in Clinical Trials: Is It Time to Rethink Our Current Approach? *JACC Basic to Transl. Sci.* **2019**, *4*, 845–854.

(10) Törnqvist, Elin; Annas, Anita; Granath, Britta; Jalkesten, Elisabeth; Cotgreave, Ian; Öberg, Mattias. Strategic Focus on 3R Principles Reveals Major Reductions in the Use of Animals in Pharmaceutical Toxicity Testing. **2014**, *9*, e101638.

(11) Lounkine, Eugen; Keiser, Michael J.; Whitebread, Steven; Mikhailov, Dmitri; Hamon, Jacques; Jenkins, Jeremy L.; Lavan, Paul; Weber, Eckhard; Doak, Allison K.; Côté, Serge; Shoichet, Brian K.; Urban, Laszlo. Large-Scale Prediction and Testing of Drug Activity on Side-Effect Targets. *Nature* **2012**, *486*, 361–367.

(12) Whitebread, Steven; Hamon, Jacques; Bojanic, Dejan; Urban, Laszlo. Keynote Review: In Vitro Safety Pharmacology Profiling: An Essential Tool for Successful Drug Development. *Drug Discovery Today*. Elsevier Current Trends November 1, 2005, pp 1421–1433.

(13) Bowes, Joanne; Brown, Andrew J.; Hamon, Jacques; Jarolimek, Wolfgang; Sridhar, Arun; Waldron, Gareth; Whitebread, Steven. Reducing Safety-Related Drug Attrition: The Use of in Vitro Pharmacological Profiling. *Nat. Rev. Drug Discov.* **2012**, *11*, 909–922.

(14) Lynch, James J.; Van Vleet, Terry R.; Mittelstadt, Scott W.; Blomme, Eric A. G. Potential Functional and Pathological Side Effects Related to Off-Target Pharmacological Activity. *J. Pharmacol. Toxicol. Methods* **2017**, *87*, 108–126.

(15) Van Vleet, Terry R.; Liguori, Michael J.; Lynch, James J.; Rao, Mohan; Warder, Scott. Screening Strategies and Methods for Better Off-Target Liability Prediction and Identification of Small-Molecule Pharmaceuticals. *SLAS Discov.* **2019**, *24*, 1–24.

(16) Deaton, Aimee M.; Fan, Fan; Zhang, Wei; Nguyen, Phuong A.; Ward, Lucas D.; Nioi, Paul. Rationalizing Secondary Pharmacology Screening Using Human Genetic and Pharmacological Evidence. *Toxicol. Sci.* **2019**, *167*, 593–603.

(17) Ietswaart, Robert; Arat, Seda; Chen, Amanda X.; Farahmand, Saman; Kim, Bumjun; DuMouchel, William; Armstrong, Duncan; Fekete, Alexander; Sutherland, Jeffrey J.;

Urban, Laszlo. Machine Learning Guided Association of Adverse Drug Reactions with in Vitro Target-Based Pharmacology. *EBioMedicine* **2020**, *57*.

(18) Clark, Matthew. Prediction of Clinical Risks by Analysis of Preclinical and Clinical Adverse Events. *J. Biomed. Inform.* **2015**, *54*, 167–173.

(19) *Meyler's Side Effects of Drugs: The International Encyclopedia of Adverse Drug Reactions and Interactions*, 12th ed.; Dukes, M. N. G., Aronson, J. K., Eds.; Elsevier Science Publishers BV: Amsterdam, 2007; Vol. 44.

(20) Nissen, David. *Mosby's Drug Consult*, 13th ed.; Mosby, 2007.

(21) Reed Elsevier Properties SA. PharmaPendium https://pharmapendium.com/ (accessed Jun 4, 2018).

(22) Wood, K. L. The Medical Dictionary for Drug Regulatory Affairs (MEDDRA) Project. *Pharmacoepidemiol. Drug Saf.* **1994**, *3*, 7–13.

(23) Clark, Matthew; Steger-Hartmann, Thomas. A Big Data Approach to the Concordance of the Toxicity of Pharmaceuticals in Animals and Humans. *Regul. Toxicol. Pharmacol.* **2018**, *96*, 94–105.

(24) Nishida, Minoru; Takashima, Yoshiharu; Ogino, Yamato; Yoneta, Yasuo; Nakamura, Kazuichi; Fujiyoshi, Masato; Kodaira, Hiroshi; Hizue, Masanori; Hisada, Shigeru; Nagayama, Takashi; Hashiba, Masamichi; Ohkura, Takako; Suzuki, Kazuhiko; Yasugi, Daisaku; Tamaki, Chihiro. Potentials and Limitations of Nonclinical Safety Assessment for Predicting Clinical Adverse Drug Reactions: Correlation Analysis of 142 Approved Drugs in Japan. *J. Toxicol. Sci.* **2013**, *38*, 581–598.

(25) Bugelski, Peter J.; Martin, Pauline L. Concordance of Preclinical and Clinical Pharmacology and Toxicology of Therapeutic Monoclonal Antibodies and Fusion Proteins: Cell Surface Targets. *Br. J. Pharmacol.* **2012**, *166*, 823–846.

(26) Bailey, Jarrod; Thew, Michelle; Balls, Michael. An Analysis of the Use of Animal Models in Predicting Human Toxicology and Drug Safety. *Altern. Lab. Anim.* **2014**, *42*, 181–199.

(27) Monticello, Thomas M. T. M. Drug Development and Nonclinical to Clinical Translational Databases:Past and Current Efforts. *Toxicol. Pathol.* **2015**, *43*, 57–61.

(28) King, G. L. Animal Models in the Study of Vomiting. *Can. J. Physiol. Pharmacol.* **1990**, *68*, 260–268.

(29) Holmes, A. M.; Rudd, J. A.; Tattersall, F. D.; Aziz, Q.; Andrews, P. L. R. Opportunities for the Replacement of Animals in the Study of Nausea and Vomiting. *Br. J. Pharmacol.* **2009**, *157*, 865–880.

(30) Needs, C. J.; Brooks, P. M. Antirheumatic Medication in Pregnancy. *Br. J. Rheumatol.* **1985**, *24*, 282–290.

(31) Lepper, Erin R.; Smith, Nicola F.; Cox, Michael C.; Scripture, Charity D.; Figg, William D. Thalidomide Metabolism and Hydrolysis: Mechanisms and Implications. *Curr. Drug Metab.* **2006**, *7*, 677–685.

(32) Voisin, Emmanuelle M.; Ruthsatz, Manfred; Collins, Jerry M.; Hoyle, Peter C. Extrapolation of Animal Toxicity to Humans: Interspecies Comparisons in Drug Development. *Regul. Toxicol. Pharmacol.* **1990**, *12*, 107–116.

(33) King, Emily A.; Wade Davis, J.; Degner, Jacob F. Are Drug Targets with Genetic Support Twice as Likely to Be Approved? Revised Estimates of the Impact of Genetic Support for Drug Mechanisms on the Probability of Drug Approval. *PLoS Genet.* **2019**, *15*, e1008489.

(34) Sharma, Hari S.; Menon, Preeti; Lafuente, José Vicente; Muresanu, Dafin F.; Tian, Z.

Ryan; Patnaik, Ranjana; Sharma, Aruna. Development of *in Vivo* Drug-Induced Neurotoxicity Models. *Expert Opin. Drug Metab. Toxicol.* **2014**, *10*, 1637–1661.

(35) Roberts, Ruth A.; Aschner, Michael; Calligaro, David; Guilarte, Tomas R.; Hanig, Joseph P.; Herr, David W.; Hudzik, Thomas J.; Jeromin, Andreas; Kallman, Mary J.; Liachenko, Serguei; Lynch, James J.; Miller, Diane B.; Moser, Virginia C.; O'Callaghan, James P.; Slikker, William; Paule, Merle G. Translational Biomarkers of Neurotoxicity: A Health and Environmental Sciences Institute Perspective on the Way Forward. *Toxicol. Sci.* **2015**, *148*, 332–340.

(36) Rasmussen, Sean A.; Mazurek, Michael F.; Rosebush, Patricia I. Catatonia: Our Current Understanding of Its Diagnosis, Treatment and Pathophysiology. *World J. Psychiatry* **2016**, *6*, 391.

(37) D., La Torre; A., Falorni. Pharmacological Causes of Hyperprolactinemia. *Ther. Clin. Risk Manag.* **2007**, *3*, 929–951.

(38) Torner, Luz. Actions of Prolactin in the Brain: From Physiological Adaptations to Stress and Neurogenesis to Psychopathology. *Front. Endocrinol. (Lausanne).* **2016**, *7*.

(39) Süleyman, Halis; Demircan, Berna; Karagöz, Yalçin. Anti-Inflammatory and Side Effects of Cyclooxygenase Inhibitors. *Pharmacol. Reports* **2007**, *59*, 247–258.

(40) Nicolaides, Nicolas C.; Pavlaki, Aikaterini N.; Maria Alexandra, Maria Alexandra; Chrousos, George P. Glucocorticoid Therapy and Adrenal Suppression. In *Endotext*; Feingold, Kenneth R., Ed.; MDText.com, Inc.: South Dartmouth (MA), 2018.

(41) Inomata, Akira; Sasano, Hironobu. Practical Approaches for Evaluating Adrenal Toxicity in Nonclinical Safety Assessment. *J. Toxicol. Pathol.* **2015**, *28*, 125–132.

(42) Huang, Yue; Yu, Sui; Wu, Zhanhe; Tang, Beisha. Genetics of Hereditary Neurological Disorders in Children. *Transl. Pediatr.* **2014**, *3*, 108–119.

(43) Ahmadian, Maryam; Suh, Jae Myoung; Hah, Nasun; Liddle, Christopher; Atkins, Annette R.; Downes, Michael; Evans, Ronald M. Pparγ Signaling and Metabolism: The Good, the Bad and the Future. *Nat. Med.* **2013**, *19*, 557–566.

(44) Hu, Bingjie; Gifford, Eric; Wang, Huijun; Bailey, Wendy; Johnson, Timothy. Analysis of the ToxCast Chemical-Assay Space Using the Comparative Toxicogenomics Database. *Chem. Res. Toxicol.* **2015**, *28*, 2210–2223.

(45) Novak, Marta; Shapiro, Colin M. Drug-Induced Sleep Disturbances. Focus on Nonpsychotropic Medications. *Drug Safety*. Springer International Publishing October 1997, pp 133–149.

(46) Toth, Linda A.; Bhargava, Pavan. Animal Models of Sleep Disorders. *Comp. Med.* **2013**, *63*, 91–104.

(47) Ledent, Catherine; Vaugeois, Jean-Marie; Schiffmann, Serge N.; Pedrazzini, Thierry; Yacoubi, Malika El; Vanderhaeghen, Jean-Jacques; Costentin, Jean; Heath, John K.; Vassart, Gilbert; Parmentier, Marc. Aggressiveness, Hypoalgesia and High Blood Pressure in Mice Lacking the Adenosine A2a Receptor. *Nature* **1997**, *388*, 674–678.

(48) Yang, Amy; Palmer, Abraham A.; De Wit, Harriet. Genetics of Caffeine Consumption and Responses to Caffeine. *Psychopharmacology (Berl).* **2010**, *211*, 245–257.

(49) Rétey, J. V; Adam, M.; Khatami, R.; Luhmann, U. F. O.; Jung, H. H.; Berger, W.; Landolt, H. P. A Genetic Variation in the Adenosine A2A Receptor Gene (ADORA2A) Contributes to Individual Sensitivity to Caffeine Effects on Sleep. *Clin. Pharmacol. Ther.* **2007**, *81*, 692–698.

(50) Mochizuki, T.; Arrigoni, E.; Marcus, J. N.; Clark, E. L.; Yamamoto, M.; Honer, M.; Borroni, E.; Lowell, B. B.; Elmquist, J. K.; Scammell, T. E. Orexin Receptor 2 Expression

in the Posterior Hypothalamus Rescues Sleepiness in Narcoleptic Mice. *Proc. Natl. Acad. Sci.* **2011**, *108*, 4471–4476.

(51) Willie, Jon T.; Chemelli, Richard M.; Sinton, Christopher M.; Tokita, Shigeru; Williams, S. Clay; Kisanuki, Yaz Y.; Marcus, Jacob N.; Lee, Charlotte; Elmquist, Joel K.; Kohlmeier, Kristi A.; Leonard, Christopher S.; Richardson, James A.; Hammer, Robert E.; Yanagisawa, Masashi. Distinct Narcolepsy Syndromes in Orexin Receptor-2 and Orexin Null Mice: Molecular Genetic Dissection of Non-REM and REM Sleep Regulatory Processes. *Neuron* **2003**, *38*, 715–730.

(52) Ghanemi, Abdelaziz; Hu, Xintian. Targeting the Orexinergic System: Mainly but Not Only for Sleep-Wakefulness Therapies. *Alexandria J. Med.* **2015**, *51*, 279–286.

(53) Irukayama-Tomobe, Yoko; Ogawa, Yasuhiro; Tominaga, Hiromu; Ishikawa, Yukiko; Hosokawa, Naoto; Ambai, Shinobu; Kawabe, Yuki; Uchida, Shuntaro; Nakajima, Ryo; Saitoh, Tsuyoshi; Kanda, Takeshi; Vogt, Kaspar; Sakurai, Takeshi; Nagase, Hiroshi; Yanagisawa, Masashi. Nonpeptide Orexin Type-2 Receptor Agonist Ameliorates Narcolepsy-Cataplexy Symptoms in Mouse Models. *Proc. Natl. Acad. Sci.* **2017**, *114*, 5731–5736.

(54) Razavi, Bibi Marjan; Hosseinzadeh, Hossein. A Review of the Role of Orexin System in Pain Modulation. *Biomed. Pharmacother.* **2017**, *90*, 187–193.

(55) Whitebread, Steven; Hamon, Jacques; Bojanic, Dejan; Urban, Laszlo. Keynote Review: In Vitro Safety Pharmacology Profiling: An Essential Tool for Successful Drug Development. *Drug Discov. Today* **2005**, *10*, 1421–1433.

(56) Santos, Cynthia; Olmedo, Ruben E. Sedative-Hypnotic Drug Withdrawal Syndrome: Recognition And Treatment. *Emerg. Med. Pract.* **2017**, *19*, 1–20.

(57) Bidwell, L. Cinnamon; Garrett, Melanie E.; McClernon, F. Joseph; Fuemmeler, Bernard F.; Williams, Redford B.; Ashley-Koch, Allison E.; Kollins, Scott H. A Preliminary Analysis of Interactions between Genotype, Retrospective ADHD Symptoms, and Initial Reactions to Smoking in a Sample of Young Adults. *Nicotine Tob. Res.* **2012**, *14*, 229–233.

(58) Fisone, G.; Borgkvist, A.; Usiello, A. Caffeine as a Psychomotor Stimulant: Mechanism of Action. *Cell. Mol. Life Sci.* **2004**, *61*, 857–872.

(59) Listos, Joanna; Malec, Danuta; Fidecka, Sylwia. Influence of Adenosine Receptor Agonists on Benzodiazepine Withdrawal Signs in Mice. *Eur. J. Pharmacol.* **2005**, *523*, 71–78.

(60) Ballesteros-Yáñez, Inmaculada; Castillo, Carlos A.; Merighi, Stefania; Gessi, Stefania. The Role of Adenosine Receptors in Psychostimulant Addiction. *Front. Pharmacol.* **2018**, *8*, 985.

(61) Lu, Ake T.; Ogdie, Matthew N.; Järvelin, Marjo-Ritta; Moilanen, Irma K.; Loo, Sandra K.; McCracken, James T.; McGough, James J.; Yang, May H.; Peltonen, Leena; Nelson, Stanley F.; Cantor, Rita M.; Smalley, Susan L. Association of the Cannabinoid Receptor Gene (CNR1) with ADHD and Post-Traumatic Stress Disorder. *Am. J. Med. Genet. B. Neuropsychiatr. Genet.* **2008**, *147B*, 1488–1494.

(62) Castelli, Maura; Federici, Mauro; Rossi, Silvia; De Chiara, Valentina; Napolitano, Francesco; Studer, Valeria; Motta, Caterina; Sacchetti, Lucia; Romano, Rosaria; Musella, Alessandra; Bernardi, Giorgio; Siracusano, Alberto; Gu, Howard H.; Mercuri, Nicola B.; Usiello, Alessandro; Centonze, Diego. Loss of Striatal Cannabinoid CB1 Receptor Function in Attention-Deficit / Hyperactivity Disorder Mice with Point-Mutation of the Dopamine Transporter. *Eur. J. Neurosci.* **2011**, *34*, 1369–1377.

(63) Haughey, Heather M.; Marshall, Erin; Schacht, Joseph P.; Louis, Ashleigh; Hutchison, Kent E. Marijuana Withdrawal and Craving: Influence of the Cannabinoid Receptor 1 (*CNR1*) and Fatty Acid Amide Hydrolase (*FAAH*) Genes. *Addiction* **2008**, *103*, 1678–1686.

(64) Jones, Declan N. C.; Holtzman, Stephen G. Influence of Naloxone upon Motor Activity Induced by Psychomotor Stimulant Drugs. *Psychopharmacology (Berl).* **1994**, *114*, 215–224.

(65) Abramov, Urho; Raud, Sirli; Kõks, Sulev; Innos, Jürgen; Kurrikoff, Kaido; Matsui, Toshimitsu; Vasar, Eero. Targeted Mutation of CCK2 Receptor Gene Antagonises Behavioural Changes Induced by Social Isolation in Female, but Not in Male Mice. *Behav. Brain Res.* **2004**, *155*, 1–11.

(66) Schnur, P.; Cesar, S. S.; Foderaro, M. A.; Kulkosky, P. J. Effects of Cholecystokinin on Morphine-Elicited Hyperactivity in Hamsters. *Pharmacol. Biochem. Behav.* **1991**, *39*, 581–586.

(67) Kayser, V.; Idänpään-Hekkilä, J. J.; Christensen, D.; Guilbaud, G. The Selective CholecystokininB Receptor Antagonist L-365,260 Diminishes the Expression of Naloxone-Induced Morphine Withdrawal Symptoms in Normal and Neuropathic Rats. *Life Sci.* **1998**, *62*, 947–952.

(68) Filardi, Marco; Pizza, Fabio; Tonetti, Lorenzo; Antelmi, Elena; Natale, Vincenzo; Plazzi, Giuseppe. Attention Impairments and ADHD Symptoms in Adult Narcoleptic Patients with and without Hypocretin Deficiency. *PLoS One* **2017**, *12*, e0182085.

(69) Gentile, Taylor A.; Simmons, Steven J.; Watson, Mia N.; Connelly, Krista L.; Brailoiu, Eugen; Zhang, Yanan; Muschamp, John W. Effects of Suvorexant, a Dual Orexin/Hypocretin Receptor Antagonist, on Impulsive Behavior Associated with Cocaine. *Neuropsychopharmacology* **2018**, *43*, 1001–1009.

(70) Azizi, Hossein; Mirnajafi-Zadeh, Javad; Rohampour, Kambiz; Semnanian, Saeed. Antagonism of Orexin Type 1 Receptors in the Locus Coeruleus Attenuates Signs of Naloxone-Precipitated Morphine Withdrawal in Rats. *Neurosci. Lett.* **2010**, *482*, 255–259.

(71) Hughes, J. P.; Rees, S.; Kalindjian, S. B.; Philpott, K. L. Principles of Early Drug Discovery. *Br. J. Pharmacol.* **2011**, *162*, 1239–1249.

(72) Komori, Shinji; Kasumi, Hiroyuki; Sakata, Kazuko; Koyama, Koji. The Role of Androgens in Spermatogenesis. *Soc. Reprod. Fertil. Suppl.* **2007**, *63*, 25–30.

(73) Milatiner, D.; Halle, David; Huerta, Michael; Margalioth, Ehud J.; Cohen, Yoram; Ben-Chetrit, Avraham; Gal, Michael; Mimoni, Tzvia; Eldar-Geva, Talia. Associations between Androgen Receptor CAG Repeat Length and Sperm Morphology. *Hum. Reprod.* **2004**, *19*, 1426–1430.

(74) Melo, C. O. A.; Danin, A. R.; Silva, D. M.; Tacon, J. A.; Moura, K. K. V. O.; Costa, E. O. A.; Da Cruz, A. D. Association between Male Infertility and Androgen Receptor Mutations in Brazilian Patients. *funpecrp.com.br Genet. Mol. Res. Genet. Mol. Res* **2010**, *9*, 128–133.

(75) Bachelot, Anne; Meduri, Géri; Massin, Nathalie; Misrahi, Micheline; Kuttenn, Frédérique; Touraine, Philippe. Ovarian Steroidogenesis and Serum Androgen Levels in Patients with Premature Ovarian Failure. *J. Clin. Endocrinol. Metab.* **2005**, *90*, 2391–2396.

(76) Shiina, H.; Matsumoto, T.; Sato, T.; Igarashi, K.; Miyamoto, J.; Takemasa, S.; Sakari, M.; Takada, I.; Nakamura, T.; Metzger, D.; Chambon, P.; Kanno, J.; Yoshikawa, H.;

Kato, S. Premature Ovarian Failure in Androgen Receptor-Deficient Mice. *Proc. Natl. Acad. Sci.* **2006**, *103*, 224–229.

(77) Fay, Kellie A.; Villeneuve, Daniel L.; Swintek, Joe; Edwards, Stephen W.; Nelms, Mark D.; Blackwell, Brett R.; Ankley, Gerald T. Differentiating Pathway-Specific From Nonspecific Effects in High-Throughput Toxicity Data: A Foundation for Prioritizing Adverse Outcome Pathway Development. *Toxicol. Sci.* **2018**, *163*, 500–515.

(78) Brown, Elliot G.; Wood, Louise; Wood, Sue. The Medical Dictionary for Regulatory Activities (MedDRA). In *Drug Safety*; John Wiley & Sons, Ltd: Chichester, UK, 1999; Vol. 20, pp 109–117.

(79) Pedregosa, Fabian; Varoquaux, Gaël; Gramfort, Alexandre; Michel, Vincent; Thirion, Bertrand; Grisel, Olivier; Blondel, Mathieu; Louppe, Gilles; Prettenhofer, Peter; Weiss, Ron; Dubourg, Vincent; Vanderplas, Jake; Passos, Alexandre; Cournapeau, David; Brucher, Matthieu; Perrot an Edouard Duchesnay Pedregosa, Matthieu; David Cournapeau, Al; Perrot matthieuperrot, Matthieu; Edouard Duchesnay. Scikit-Learn: Machine Learning in Python. *J. Mach. Learn. Res.* **2011**, *12*, 2825–2830.

(80) Cahill, Nathan D. Normalized Measures of Mutual Information with General Definitions of Entropy for Multimodal Image Registration. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*; 2010; Vol. 6204 LNCS, pp 258–268.

(81) Li, Wentian. Mutual Information Functions Versus Correlation Functions in Binary Sequences. *J. Stat. Phys.* **2012**, *60*, 249–252.

(82) Fisher, R. A. On the Interpretation of χ 2 from Contingency Tables, and the Calculation of P. *J. R. Stat. Soc.* **1922**, *85*, 87.

(83) Jones, Eric; Oliphant, Travis; Peterson, Pearu. SciPy: Open source scientific tools for Python http://www.scipy.org/ (accessed Jun 4, 2018).

(84) Grimes, David A.; Schulz, Kenneth F. Refining Clinical Diagnosis with Likelihood Ratios. *Lancet* **2005**, *365*, 1500–1505.

(85) Wall, R. J.; Shani, M. Are Animal Models as Good as We Think? *Theriogenology* **2008**, *69*, 2–9.

(86) Bailey, Jarrod; Balls, Michael. Recent Efforts to Elucidate the Scientific Validity of Animal-Based Drug Tests by the Pharmaceutical Industry, pro-Testing Lobby Groups, and Animal Welfare Organisations. *BMC Medical Ethics*. BioMed Central Ltd. March 2019.

(87) Chien, Patrick F. W.; Khan, Khalid S. Evaluation of a Clinical Test. II: Assessment of Validity. *BJOG An Int. J. Obstet. Gynaecol.* **2003**, *108*, 568–572.

(88) Murrell, Daniel S.; Cortes-Ciriano, Isidro; Van Westen, Gerard J. P.; Stott, Ian P.; Bender, Andreas; Malliavin, Thérèse E.; Glen, Robert C. Chemically Aware Model Builder (Camb): An R Package for Property and Bioactivity Modelling of Small Molecules. *J. Cheminform.* **2015**, *7*, 45–55.

(89) Berthold, Michael R.; Cebron, Nicolas; Dill, Fabian; Gabriel, Thomas R.; Kötter, Tobias; Meinl, Thorsten; Ohl, Peter; Thiel, Kilian; Wiswedel, Bernd. KNIME - the Konstanz Information Miner. In *ACM SIGKDD Explorations Newsletter*; Preisach C., Burkhardt H., Schmidt-Thieme L., Decker R., Ed.; Springer: Berlin, Heidelberg, 2009; Vol. 11, p 26.

(90) Krejsa, Cecile M.; Horvath, Dragos; Rogalski, Sherri L.; Penzotti, Julie E.; Mao, Boryeu; Barbosa, Frédérique; Migeon, Jacques C. Predicting ADME Properties and Side Effects: The BioPrint Approach. *Curr. Opin. Drug Discov. Devel.* **2003**, *6*, 470–480.

(91) Bento, a. Patrícia; Gaulton, Anna; Hersey, Anne; Bellis, Louisa J.; Chambers, Jon; Davies, Mark; Krüger, Felix a.; Light, Yvonne; Mak, Lora; McGlinchey, Shaun; Nowotka, Michal; Papadatos, George; Santos, Rita; Overington, John P. The ChEMBL Bioactivity Database: An Update. *Nucleic Acids Res.* **2014**, *42*, D1083-90.

(92) Tang, Jing; Tanoli, Zia ur Rehman; Ravikumar, Balaguru; Alam, Zaid; Rebane, Anni; Vähä-Koskela, Markus; Peddinti, Gopal; van Adrichem, Arjan J.; Wakkinen, Janica; Jaiswal, Alok; Karjalainen, Ella; Gautam, Prson; He, Liye; Parri, Elina; Khan, Suleiman; Gupta, Abhishekh; Ali, Mehreen; Yetukuri, Laxman; Gustavsson, Anna Lena; et al. Drug Target Commons: A Community Effort to Build a Consensus Knowledge Base for Drug-Target Interactions. *Cell Chem. Biol.* **2018**, *25*, 224–229.e2.

(93) Siramshetty, Vishal B.; Eckert, Oliver Andreas; Gohlke, Björn-Oliver; Goede, Andrean; Chen, Qiaofeng; Devarakonda, Prashanth; Preissner, Saskia; Preissner, Robert. SuperDRUG2: A One Stop Resource for Approved/Marketed Drugs. *Nucleic Acids Res.* **2018**, *46*, D1137–D1143.

(94) Heller, Stephen R.; McNaught, Alan; Pletnev, Igor; Stein, Stephen; Tchekhovskoi, Dmitrii. InChI, the IUPAC International Chemical Identifier. *J. Cheminform.* **2015**, *7*, 23.

(95) Maglott, Donna; Ostell, Jim; Pruitt, Kim D.; Tatusova, Tatiana. Entrez Gene: Gene-Centered Information at NCBI. *Nucleic Acids Res.* **2011**, *39*, D52-7.

(96) Magrane, Michele; Consortium, Uni Prot. UniProt Knowledgebase: A Hub of Integrated Protein Data. *Database* **2011**, *2011*, bar009.

(97) Toad for mysql - MariaDB Knowledge Base http://www.toadworld.com/products#mysql (accessed Feb 4, 2019).

(98) Burdett, T.; Jupp, Simon; Malone, James; Williams, Eleanor; Keays, Maria; Parkinson, Helen; Trust, Wellcome; Campus, Genome. Zooma2 - A repository of annotation knowledge and curation API http://www.ebi.ac.uk/spot/zooma/index.html (accessed Jun 4, 2018).

(99) Schriml, Lynn M.; Mitraka, Elvira; Munro, James; Tauber, Becky; Schor, Mike; Nickle, Lance; Felix, Victor; Jeng, Linda; Bearer, Cynthia; Lichenstein, Richard; Bisordi, Katharine; Campion, Nicole; Hyman, Brooke; Kurland, David; Oates, Connor Patrick; Kibbey, Siobhan; Sreekumar, Poorna; Le, Chris; Giglio, Michelle; et al. Human Disease Ontology 2018 Update: Classification, Content and Workflow Expansion. *Nucleic Acids Res.* **2019**, *47*, D955–D962.

(100) Smith, Cynthia L.; Eppig, Janan T. The Mammalian Phenotype Ontology: Enabling Robust Annotation and Comparative Analysis. *Wiley Interdiscip. Rev. Syst. Biol. Med.* **2009**, *1*, 390–399.

(101) Robinson, Peter N.; Köhler, Sebastian; Bauer, Sebastian; Seelow, Dominik; Horn, Denise; Mundlos, Stefan. The Human Phenotype Ontology: A Tool for Annotating and Analyzing Human Hereditary Disease. *Am. J. Hum. Genet.* **2008**, *83*, 610–615.

(102) Malone, James; Holloway, Ele; Adamusiak, Tomasz; Kapushesky, Misha; Zheng, Jie; Kolesnikov, Nikolay; Zhukova, Anna; Brazma, Alvis; Parkinson, Helen. Modeling Sample Variables with an Experimental Factor Ontology. *Bioinformatics* **2010**, *26*, 1112–1118.

(103) Vasant, Drashtti; Chanas, Laetitia; Malone, James; Hanauer, Marc; Olry, Annie; Jupp, Simon; Robinson, Peter N.; Parkinson, Helen; Rath, Ana. ORDO: An Ontology Connecting Rare Disease, Epidemiology and Genetic Data. In *Phenotype data at ISMB2014*; 2014.

(104) Ceusters, Werner; Smith, B.; Goldberg, L. A Terminological and Ontological Analysis of

the NCI Thesaurus. *Methods Inf. Med.* **2005**, *44*, 498–507.
(105) Huang, Jingshan; Dang, Jiangbo; Borchert, Glen M.; Eilbeck, Karen; Zhang, He; Xiong, Min; Jiang, Weijian; Wu, Hao; Blake, Judith A.; Natale, Darren A.; Tan, Ming. OMIT: Dynamic, Semi-Automated Ontology Development for the MicroRNA Domain. *PLoS One* **2014**, *9*, e100855.
(106) He, Yongqun; Sarntivijai, Sirarat; Lin, Yu; Xiang, Zuoshuang; Guo, Abra; Zhang, Shelley; Jagannathan, Desikan; Toldo, Luca; Tao, Cui; Smith, Barry. OAE: The Ontology of Adverse Events. *J. Biomed. Semantics* **2014**, *5*, 29.
(107) Mungall, Christopher J.; Koehler, Sebastian; Robinson, Peter; Holmes, Ian; Haendel, Melissa. K-BOOM: A Bayesian Approach to Ontology Structure Inference, with Applications in Disease Ontology Construction. *bioRxiv* **2019**, 048843.
(108) Mohammed, Osama; Benlamri, Rachid; Fong, Simon. Building a Diseases Symptoms Ontology for Medical Diagnosis: An Integrative Approach. In *1st International Conference on Future Generation Communication Technologies, FGCT 2012*; IEEE, 2012; pp 104–108.
(109) Ceusters, Werner; Smith, Barry. Foundations for a Realist Ontology of Mental Disease. *J. Biomed. Semantics* **2010**, *1*, 10.
(110) Schofield, P. N.; Gruenberger, M.; Sundberg, John P. Pathbase and the MPATH Ontology: Community Resources for Mouse Histopathology. *Vet. Pathol.* **2010**, *47*, 1016–1020.
(111) Dönitz, Jürgen; Wingender, Edgar. The Ontology-Based Answers (OBA) Service: A Connector for Embedded Usage of Ontologies in Applications. *Front. Genet.* **2012**, *3*, 197.
(112) Visser, Ubbo; Abeyruwan, Saminda; Vempati, Uma; Smith, Robin P.; Lemmon, Vance; Schürer, Stephan C. BioAssay Ontology (BAO): A Semantic Description of Bioassays and High-Throughput Screening Results. *BMC Bioinformatics* **2011**, *12*, 257.
(113) Koscielny, Gautier; An, Peter; Carvalho-Silva, Denise; Cham, Jennifer A.; Fumis, Luca; Gasparyan, Rippa; Hasan, Samiul; Karamanis, Nikiforos; Maguire, Michael; Papa, Eliseo; Pierleoni, Andrea; Pignatelli, Miguel; Platt, Theo; Rowland, Francis; Wankar, Priyanka; Bento, A. Patrícia; Burdett, Tony; Fabregat, Antonio; Forbes, Simon; et al. Open Targets: A Platform for Therapeutic Target Identification and Validation. *Nucleic Acids Res.* **2017**, *45*, D985–D994.
(114) opentargets - Python client for targetvalidation.org — opentargets 2.0.0 documentation https://opentargets.readthedocs.io/en/stable/ (accessed Oct 8, 2018).
(115) Zerbino, Daniel R.; Achuthan, Premanand; Akanni, Wasiu; Amode, M. Ridwan; Barrell, Daniel; Bhai, Jyothish; Billis, Konstantinos; Cummins, Carla; Gall, Astrid; Girón, Carlos García; Gil, Laurent; Gordon, Leo; Haggerty, Leanne; Haskell, Erin; Hourlier, Thibaut; Izuogu, Osagie G.; Janacek, Sophie H.; Juettemann, Thomas; To, Jimmy Kiang; et al. Ensembl 2018. *Nucleic Acids Res.* **2018**, *46*, D754–D761.
(116) Mungall, Christopher J.; McMurry, Julie A.; Köhler, Sebastian; Balhoff, James P.; Borromeo, Charles; Brush, Matthew; Carbon, Seth; Conlin, Tom; Dunn, Nathan; Engelstad, Mark; Foster, Erin; Gourdine, J. P.; Jacobsen, Julius O. B.; Keith, Dan; Laraway, Bryan; Lewis, Suzanna E.; NguyenXuan, Jeremy; Shefchek, Kent; Vasilevsky, Nicole; et al. The Monarch Initiative: An Integrative Data and Analytic Platform Connecting Phenotypes to Genotypes across Species. *Nucleic Acids Res.* **2017**, *45*, D712–D722.

(117) Requests: HTTP for Humans™ — Requests 2.21.0 documentation http://docs.python-requests.org/en/master/ (accessed Mar 14, 2019).
(118) About the HGNC: HUGO Gene Nomenclature Committee https://www.genenames.org/about/overview (accessed Oct 8, 2018).
(119) Smith, Cynthia L.; Blake, Judith A.; Kadin, James A.; Richardson, Joel E.; Bult, Carol J.; Mouse Genome Database Group. Mouse Genome Database (MGD)-2018: Knowledgebase for the Laboratory Mouse. *Nucleic Acids Res.* **2018**, *46*, D836–D842.
(120) Jensen, Lars Juhl; Julien, Philippe; Kuhn, Michael; von Mering, Christian; Muller, Jean; Doerks, Tobias; Bork, Peer. EggNOG: Automated Construction and Annotation of Orthologous Groups of Genes. *Nucleic Acids Res.* **2008**, *36*, D250-4.
(121) pandas: a Foundational Python Library for Data Analysis and Statistics | R (Programming Language) | Database Index https://www.scribd.com/document/71048089/pandas-a-Foundational-Python-Library-for-Data-Analysis-and-Statistics (accessed Jun 4, 2018).
(122) Ahlberg, Christopher. Spotfire: An Information Exploration Environment. *ACM SIGMOD Rec.* **1996**, *25*, 25–29.