Polypharmacological *In Silico* Bioactivity Profiling and Experimental Validation Uncovers Sedative-Hypnotic Effects of Approved and Experimental Drugs in Rat

Georgios Drakakis,[1] Keith A. Wafford,[2] Suzanne C. Brewerton,[2,&]

Michael J. Bodkin,[2,+] David A. Evans[2] and Andreas Bender[1,*]

1. Centre for Molecular Informatics, Department of Chemistry, University of Cambridge, Lensfield Road, Cambridge CB2 1EW, United Kingdom, Telephone: +44(0)1223 762983; Fax: +44 (0)1223 763076

2. Eli Lilly U.K., Erl Wood Manor, Windlesham, Surrey GU206PH, United Kingdom

& Current address: Human Longevity Singapore Pte. Ltd, 1 Fusionopolis Link, Nexus@One-North, North Tower, Suite #06-06, Singapore 138542

+ Current address: Evotec (UK) Ltd, 114 Innovation Drive, Milton Park, Abingdon, Oxfordshire OX14 4RZ, UK

* ab454@cam.ac.uk

Keywords: cheminformatics, *in silico* bioactivity prediction, bioactivity profiles, mechanism-of-action analysis, phenotypic deconvolution, side-effect prediction

Conflict of interest: The authors declare that they have no conflict of interest.

Acknowledgements: This research was supported by Unilever (AB), the EPSRC (GD) and Eli Lilly (GD). AB thanks the ERC for an ERC Starting Grant.

ABSTRACT

In this work, we describe the computational ('*in silico*') mode-of-action analysis of CNS-active drugs, which is taking both *multiple simultaneous hypotheses* as well as *sets of protein targets* for each mode-of-action into account, and which was followed by successful prospective *in vitro* and *in vivo* validation. Using sleep-related phenotypic readouts describing both efficacy and side-effects for 491 compounds tested in rat, we defined an 'optimal' (desirable) sleeping pattern. Compounds were subjected to *in silico* target prediction (which was experimentally confirmed for 21 out of 28 cases, corresponding to 75%), followed by the utilization of decision trees for deriving polypharmacological bioactivity profiles. We demonstrated that predicted bioactivities improved classification performance compared to using only structural information. Moreover, DrugBank molecules were processed *via* the same pipeline and compounds in many cases not annotated as sedative-hypnotic (alcaftadine, benzatropine, palonosetron, ecopipam, cyproheptadine, sertindole and clopenthixol) were prospectively validated *in vivo*. Alcaftadine, ecopipam cyproheptadine and clopenthixol were found to promote sleep as predicted, benzatropine showed only a small increase in NREM sleep, whereas sertindole promoted wakefulness. To our knowledge, the sedative-hypnotic effects of alcaftadine and ecopipam have not been previously discussed in literature. The method described extends previous single-target, single-mode-of-action models, and is applicable across disease areas.

INTRODUCTION

Drug discovery has traditionally relied on measuring phenotypic readouts on a biological system.[1] Recently phenotypic screening has experienced a resurgence primarily due to technological advances and late-stage clinical trial failures of medicines that were discovered in a target-based fashion.[2] These attrition rates were in some cases due to the lack of efficacy of compounds designed based on the "one gene, one drug, one disease"[3] paradigm, particularly in the case of complex central nervous system (CNS) disorders, such as depression and schizophrenia.[4] This led to 'selectively unselective' compounds, whose main drawback was the secondary activities elicited via the binding to off-targets.[5] In order to avoid undesirable effects, the design of compounds with very specific multi-target biological activity profiles was introduced, named designed multiple ligands[6]. The paradigm in which a compound elicits its effect by modulating multiple targets is termed polypharmacology. A common example of a drug known to bind to multiple targets (even though not rationally optimised) is the antipsychotic clozapine, with activity on several serotonin, dopamine and adrenergic receptors.[4] Its main effect however is carried out by activity on Dopamine D2 and 5-hydroxytryptamine 2A receptors.[7] Another CNS disorder seeking polypharmacological drug candidates (or alternatively a combination of compounds) is Alzheimer's disease, with targets such as acetylcholinesterase and monoamine oxidase.[8] The observed shift in drug design approach[3,4,9] has also reached *in silico* analyses, both in bioactivity prediction[10,11] and mechanism-of-action (MoA) analysis using machine learning algorithms.[12,13] One area where bioactivity against multiple targets is of relevance is sleep, which we will consider in this study.

Sleep and insomnia

Sleep can be defined as the naturally recurring state of rest, vital for energy restoration, tissue growth and healing, protein synthesis, as well as processing information in the form of memory and learning.[14–17] Sleep can be divided into two main states, namely rapid eye movement (REM) and non-rapid eye movement (non-REM) sleep.[18] REM sleep is characterised by high frequency and low amplitude electrical activity from electroencephalogram (EEG) readouts coincident with muscle atonia. Non-REM sleep is characterised by low frequency and high amplitude waves (slow wave sleep; SWS) and in humans can be broken down into several sub-stages.[19] Disruption to these sleep stages leads to sleep disorders, insomnia being the most common. Insomnia is described as the difficulty of sleep initiation/maintenance, or the experience of unsatisfactory non-restorative sleep.[20–23] Many different biological mechanisms have been targeted to optimise insomnia treatment both for efficacy (rapid sleep onset, increased NREM sleep and improved sleep consolidation) and side-effect (no impairment on normal daily performance, no addiction etc.).

The main mechanisms of small molecules supporting or causing sleep

The most common categories of hypnotics currently on the market are benzodiazepine ligands, antihistamines, and melatonin receptor agonists.[21] Benzodiazepines (BZs) such as diazepam bind to gamma-aminobutyric acid$_A$ ($GABA_A$) receptors and thus augment the inhibitory effect of the GABA neurotransmitter mediated by $GABA_A$.[24,25] Other BZ-site drugs such as zolpidem have some $GABA_A$ subtype selectivity for a1-subunit containing receptors and exert preferential effects at this subtype.[26] Furthermore, several antihistamines, antidepressants and antipsychotics have been used as sedative-hypnotics, e.g. trazodone and doxepin.[20]

In the course of developing novel sedative-hypnotics attention has turned to the natural neurotransmitters and circuits that control sleep and wake. In addition to the inhibitory GABAergic system, many of the monoaminergic transmitters such as histamine, acetylcholine, dopamine and serotonin are linked to promoting and maintaining wakefulness and inhibiting these pathways promotes sleep. In addition peptides such as orexin and neuropeptide S have been shown important in controlling sleep and wake.[27,28]

More specifically, histamine H1 receptor antagonists decrease wakefulness by blocking the action of histamine.[29] First generation H1-antagonists were more potent sedative-hypnotics due to higher blood-brain barrier penetration.[30] Histamine H1 receptor antagonism has also been identified as an off-target activity (along with muscarinic receptors) for antidepressants, whose primary action is achieved by enhancing serotonergic and noradrenergic neurotransmission.[31] Antipsychotics are also promiscuous binders for G-protein-coupled receptors (GPCRs), namely histamine, dopamine, serotonin, α-adrenergic and muscarinic receptors.[32–34] Another class of sedative-hypnotics are melatonin receptor agonists, which increase sleep and reduce sleep onset latency, but are generally less effective than benzodiazepines and related compounds.[21,35]

For purposes later in this study, we must mention the involvement of the transient receptor potential cation channel subfamily A member 1 (TRPA1) in the sleep/wake mechanism. It has not been used a therapeutic target for sleep disorders, but TRPA1 has been used to stimulate the large ventral lateral clock neurons (lLNv),[36] causing sleep loss in *Drosophila*[37]. To our knowledge, this is the only protein target in our analysis whose *individual* link to sleep in humans is not well-established.

None of the existing agents for insomnia are without problems and there is clinical need for improved sleep promoting medications. Side-effects include residual drowsiness or addiction as

in the case of benzodiazepines such as triazolam,[38] or seizures such as in the case of clozapine.[7] Melatonin receptor agonists do not have as pronounced adverse effects such as dependence and addiction[39] but appear not as effective.[21]

Hence, in this work, we used *in silico* methods to understand the polypharmacological MoA of sedative-hypnotics with the aim to arrive at more efficacious compounds in the future. Recently, *in silico* target prediction was used to explore neuroactive polypharmacology for behavioural phenotypes in zebrafish[40]. To this end, we used sleep-related measures from the Eli Lilly *SCORE2004*™ rat model to describe what we are calling an "optimal sleeping profile". More specifically, we used readouts describing compound efficacy such as minutes of sleep increase, as well as side-effect, such as decrease in locomotor intensity. We proceeded to classify compounds as for whether they elicit the desirable phenotype or not, and then to linked chemical structures to phenotypes via emerging patterns in their predicted polypharmacological bioactivity profiles. Finally, these profiles were used to assess selected (marketed and their analogues) drugs from DrugBank[41] for predicted sedative-hypnotic effects. Those predicted as promoting an optimal sleeping profile but had no literature annotation were followed up in *SCORE2004*™. Furthermore, three more drugs were investigated *in vivo* based on partially or fully matching the most populated bioactivity profile in this analysis. The overall aim is to understand sedative-hypnotic MoAs using multiple hypotheses on predicted polypharmacological bioactivity profiles.

RESULTS & DISCUSSION

In silico derivation of polypharmacological mode-of-action hypotheses for sedative-hypnotic compounds

Parameters for optimising both classification accuracy and decision tree size/depth were derived, resulting in a confidence factor of 0.75, using 10 features as attributes with the minimum number of objects in each leaf set to 7. The output comprised a decision tree of 23 nodes (total size) and depth 5 with a classification accuracy of 68.4% after 10-fold cross-validation, shown in Figure 1. At each leaf, the first number describes the total number of instances reaching the leaf, whilst the second describes the number of instances misclassified. The accuracy reported was computed based on the average performance of the 10 individual models built during cross-validation. The final tree, obtained on all data, achieves 70.9% overall accuracy. The ROC and precision/recall curves can be seen in Supporting Figures 1 and 2. It can be seen that for both the optimal and non-optimal sleep profiles, the true positive rates (0.679 and 0.689 respectively), as well as the ROC curves (area under the curve 0.694) are highly similar. This suggests that our classifier is able to predict both classes with approximately the same accuracy (highly balanced). Similarly, precision and recall in both cases are very consistent, respectively 0.665 and 0.679 for the optimal sleep profile, and 0.702 and 0.689 for non-optimal sleep.

The performance of the resulting decision tree was compared with the one derived using Molprint2D fingerprints (i.e. chemical structure) as input variables, in order to assess whether the bioactivities added information to the analysis. The results showed a balanced accuracy of 56.4% over 10-fold cross validation (ROC area 0.579), with true positive rates for optimal and non-optimal sleep compounds reported as 0.545 and 0.585 respectively. It can thus be concluded that including predicted bioactivity has enhanced the analysis, given the numbers reported above.

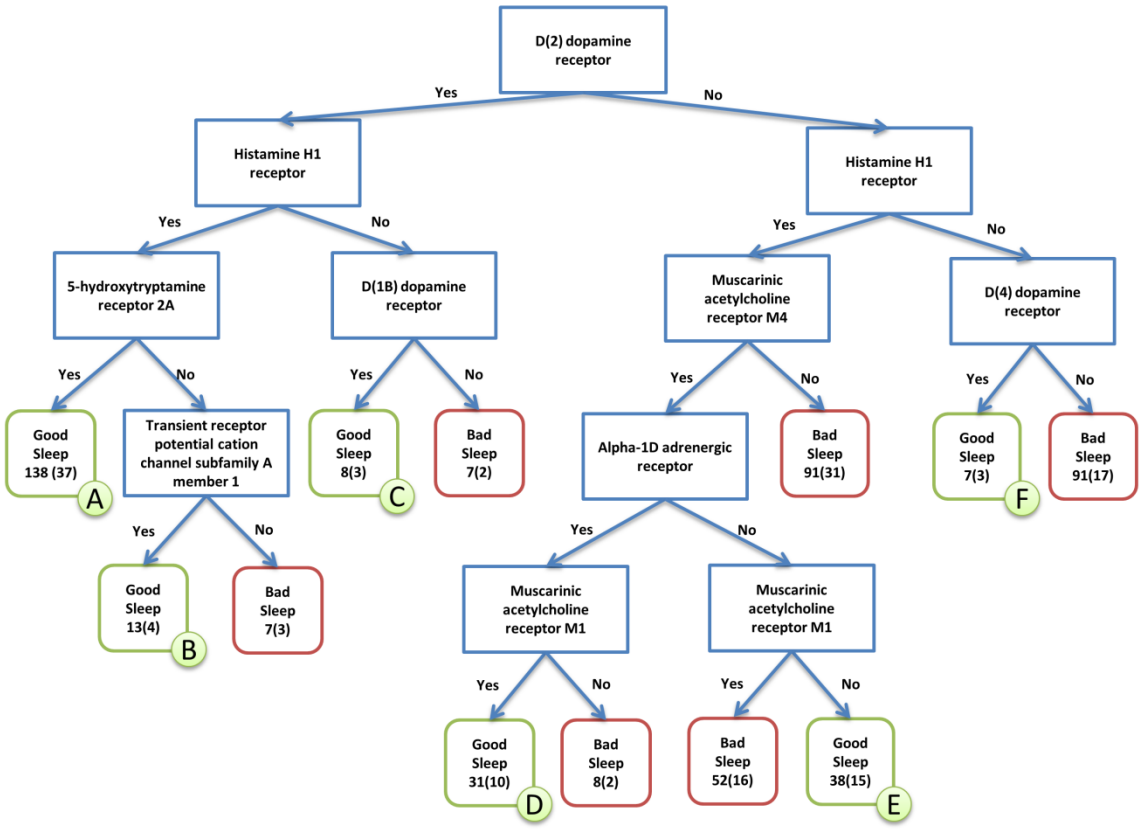

Figure 1 | C4.5 decision derived for *SCORE2004*^{TM} dataset. Bioactivity profiles of compounds are read by assessing predicted activity (yes/no on decision tree edges) on a series of protein targets (tree nodes) until reaching a binary phenotypic outcome (optimal/not sleep profile; tree leaves). The number of instances (compounds) from our training set reaching a particular phenotypic outcome is annotated in each leaf, along with the number of those misclassified in parentheses. The tree was derived from 491 compounds evaluated for 7 sleep-related metrics and resulted in 5 polypharmacological bioactivity profiles (A, B, C, D and E; F inferred single target activity) linked to promoting an optimal sleeping profile. Both sides of the tree include HRH1 as a node which signifies its importance for sedation, whereas ADRA1D and CHRM1 seem to be involved in sleep promotion in opposing combinations. The largest class (138 instances) consists of compounds with simultaneous predicted activity on DRD2, HRH1 and HTR2A.

Furthermore, as a baseline, we note the inferior performance of Naïve Bayes classifiers using both predicted bioactivities and Molprint2D fingerprints as input variables. The resulting balanced accuracy for predicted bioactivities was 65.2% with a true positive rate of 0.543 for optimal sleep and 0.751 for non-optimal sleep. Similarly, the Molprint2D-based Naïve Bayes classifier reported a 61.1% balanced accuracy, whilst the true positive rate for optimal sleep was merely 0.490, compared to 0.744 for non-optimal sleep. Both classifiers were outperformed by the one used in this work (68.4% balanced accuracy) and had a low true positive rate for optimal sleep-promoting compounds (0.543 and 0.490 compared to 0.679 of the current model).

The protein targets used in specific combinations to classify compounds as promoting an optimal sleeping pattern are the D(1B), D(2) and D(4) dopamine receptors (DRD1B, DRD2, DRD4),[32–34] histamine H1 receptor (HRH1),[29] muscarinic acetylcholine receptors M1 and M4 (CHRM1, CHRM4),[34] α1A adrenergic receptor (ADRA1D),[31] 5-hydroxytryptamine receptor 2A (HTR2A)[20] and transient receptor potential cation channel subfamily A member 1 (TRPA1),[36] which have all linked to sleep in literature. Our approach however went one step further to suggest combinations of targets involved in the mode-of-action of sedative-hypnotics using decision trees, which are summarised in Table 1.

Table 1 | Bioactivity profiles describing an optimal sleeping pattern derived from the decision tree. Each profile (A-F) is annotated with '1', '0' or 'NA', respectively meaning 'activity', 'inactivity' and 'not relevant for the profile'. The protein targets are in accordance with known targets known to modulate sleep, most of the specific combinations however appear to be novel. Dopamine D2 and histamine H1 receptors are assessed for all profiles and are therefore concluded to be the most important targets for this analysis to separate compounds promoting an optimal sleeping pattern from those which do not.

Protein Targets	Polypharmacological Bioactivity Profiles					
	A	**B**	**C**	**D**	**E**	**F**
D(2) dopamine receptor	1	1	1	0	0	0
Histamine H1 receptor	1	1	0	1	1	0
5-hydroxytryptamine receptor 2A	1	0	*NA*	*NA*	*NA*	*NA*
Transient receptor potential cation channel subfamily A member 1	*NA*	1	*NA*	*NA*	*NA*	*NA*
D(1B) dopamine receptor	*NA*	*NA*	1	*NA*	*NA*	*NA*
Muscarinic acetylcholine receptor M4	*NA*	*NA*	*NA*	1	1	*NA*
α-1A adrenergic receptor	*NA*	*NA*	*NA*	1	0	*NA*
Muscarinic acetylcholine receptor M1	*NA*	*NA*	*NA*	1	0	*NA*
D(4) dopamine receptor	*NA*	*NA*	*NA*	*NA*	*NA*	1

Even though the targets have been individually linked to sleep in literature, to the best of our knowledge most of the suggested specific combinations are novel. More specifically, the main bioactivity profile ('A'; 173 instances) suggests that simultaneous activity on DRD2, HRH1 and HTR2A should promote good sleep. The HRH1/HTR2A combination (which is the main mechanism targeted by compounds tested thus far in our *in vivo* rat model) is known, but the pairing with DRD2 is not well-established. On the other hand, activity on DRD2 and HTR2A has been discussed in previous work concerning sleep[7]. Simultaneous activity on DRD2, HRH1 and TRPA1 (Profile 'B'; 13 instances) has not been discussed before in the literature, however given the involvement of the individual receptors in promoting sleep processes, this could be a novel

polypharmacological bioactivity profile eliciting sedative-like behaviour. The combinatorial activity on DRD2/DRD1B ('C'; 8 instances) has not been explicitly stated, but the dopaminergic system is definitely of interest in sleep studies, as D1 and D2 antagonists are known separately to promote sleep.[32] Finally, profiles 'D' and 'E' (31 and 38 instances respectively) are very similar, as both include HRH1/CHRM4 activity, but are predicted to promote good sleep only when CHRM1 and ADRA1D are simultaneously targeted ('D') or missed ('E'). This means that sedation is predicted to be mainly caused by HRH1 and CHRM4 activity, whilst the individual effects from CHRM1 and ADRA1D are cancelled out when both these receptors are targeted. Profile 'F' was discarded due to predicted activity on a *single* target. Next, feature correlation analysis was performed to assess whether the multi-target bioactivity profiles found in the decision tree were the result of correlated features, and hence spurious, or whether they indeed contained complementary information (Supporting Table 1).

We recognise that the full interpretation of the decision tree is not possible without further information on the in vivo brain levels of the compounds, which is unfortunately not possible to collect in the *SCORE2014*^{TM} experiment (and which is generally not trivial to obtain). As steps towards this we present the predicted fraction unbound in brain in Supporting Table 2, and note the published experimental rat exposure data of ecopipam[42] where the cortex and striatum total concentration are quoted at between 10 and 1000 μg/l for a dose of 0.25mg/kg over a time course of 4hr, which translates to between 0.03 – 3μM. Coupled with the biochemical activities in Table 1, these values are consistent within the wide experimental error and predictive model uncertainty with sleep profile C.

Compound Selection for experimental follow-up

We proceeded to select marketed compounds from DrugBank which followed our predicted bioactivity profiles, in order to validate our polypharmacological mechanism-of-action hypotheses prospectively *in vivo*. The full table from the DrugBank molecules following the profiles A-E predicted to have sleep-promoting activity can be found in Supporting Table 3, along with their predicted blood-brain barrier permeation value.[43] From the subset of molecules not annotated with sedative-hypnotic keywords in literature, the compounds originally selected for testing in *SCORE2004*™ were benzatropine (profile A), fenoldopam and palonosetron (C), alcaftadine (D) and loratadine (E). Loratadine was not run as it had already been evaluated in *SCORE2004*™ and shown to promote sleep as expected, and had been included in the decision tree training set. Fenoldopam was replaced by blood-brain barrier permeating analogue SCH-39166 (ecopipam). We note that an additional fenoldopam analogue under consideration for testing, SCH-23390, has already been proven to promote sleep in rats in another work.[44] From the most-populated profile A, we decided to additionally test cyproheptadine (predicted to bind DRD2, HRH1 and HTR2A), as well as compounds partially fulfilling the profile: sertindole (HRH1/HTR2A) and clopenthixol (DRD2/HRH1 predicted profile), in order to pinpoint the most significant of the three targets, if possible. The latter profile also would occur in the decision tree with the enforcement of more strict parameters, such as limiting the tree depth. In a biological context, this means that predicted activities on DRD2 and HRH1 are likely more important than activity on HTR2A for separating compounds promoting an optimal sleeping profile from those which do not in our dataset. All compounds tested are also provided for the user in the appropriate leaves of the decision tree in Supporting Figure 3. We note the inferior performance of C4.5 decision trees generated using only activity against DRD2 and HRH1 (accuracy 65.2%;

ROC area 0.634) and HTH1 and HTR2A (accuracy 65.6%; ROC area 0.653) as input variables compared to the model used in our analysis (accuracy 68.4%; ROC area 0.694). Furthermore, the true positive rate for the optimal sleep compounds is merely 0.487 and 0.453 respectively (0.802 and 0.840 for non-optimal sleep), suggesting an imbalanced model, outperformed by the model used in this work (0.679 for optimal and 0.689 for non-optimal sleep). Profile B was excluded at the time due to the fact that only Phenindamine and Chlordiazpozide were present, and these are confirmed sedative-hypnotics, respectively as HRH1 and GABAA binders.

Performance of Bioactivity predictions

We proceeded to validate *in vitro* all the predicted bioactivity profiles for each molecules which were selected to be tested *in vivo* as potential sleep-promoting compounds. Overall 21 of 28 predicted protein targets were indeed correct (75%). The target predictions for benzatropine, clopenthixol and cyproheptadine were correct in all cases (3 predictions each; 9 of 9 instances correct). Two of three ecopipam predictions were correct, where HRH1 binding was not predicted while the compound showed activity in vitro, whilst three of five alcaftadine predictions were also correct, and where activity on ADRA1D and CHRM1 was predicted but not observed. Moreover, bioactivity predictions for loratadine and sertindole were mostly correct, succeeding in four out of five and two out of three cases, respectively. Palonosetron was the only compound with two of three errors in predictions, which is also reflected in the failure to modulate the sleep mechanism, discussed later in this work. TRPA1 binders were not assessed *in vitro,* as all molecules following the predicted bioactivity profile were known sedative-hypnotics. This information is summarised in Table 2.

Table 2 | Bioactivity outcomes for DrugBank molecules predicted to be sedative-hypnotic in this work. Assay performed on all relevant protein targets derived from the decision tree. Values are in uM (IC50/EC50). * indicates that the value was estimated from a single point measurement with the formula IC50(single point) = c * (100 - s) / s, where 's' is the % inhibition or stimulation at c, where c is the concentration of the single point dose in uM . Colour annotations for TP, TN, FP and FN have been added to the bottom of the table for the reader's convenience. It can be seen that 21 of 28 predictions are correct.

	Molecule	ADRA1D	CHRM1	CHRM4	DRD1	DRD2	HRH1	HTR2A	
Predicted	alcaftadine	Active	Active	Active		Inactive	Active		
	benzatropine					Active	Active	Active	
	clopenthixol					Active	Active	Inactive	
	cyproheptadine					Active	Active	Active	
	ecopipam				Active	Active	Inactive		
	loratadine	Inactive	Inactive	Active		Inactive	Active		
	sertindole					Inactive	Active	Active	
	palonosetron				Active	Active	Inactive		**Score**
Measured	alcaftadine	>99*	5.3*	0.43*		>99*	0.0030*		3 of 5
	benzatropine					0.78	<0.001	0.0070	3 of 3
	clopenthixol					0.0026	0.001	10	3 of 3
	cyproheptadine					0.036	<0.001	<0.001	3 of 3
	ecopipam				<0.001	0.197985	0.366491		2 of 3
	loratadine	10	5.8*	16*		10	0.02		4 of 5
	sertindole					<0.001	<0.001	<0.001	2 of 3
	palonosetron				>99*	20*	1.0*		1 of 3
	Annotations	TP	TN	FP	FN	Irrelevant		**Total**	21 of 28

The full table for all *in vitro* measurements for the seven compounds tested *in vivo* for this work is provided in Supporting Table 4, including the assay mode (inhibition/antagonism/agonism/potentiation), assay measurement (concentration response curve/single point), activity type (Ki, IC50, EC50, Kb, etc.), average and most potent values, unit, concentration, standard deviation and number of replicates of each compound in an assay.

It can be seen that for the majority of compounds tested, HRH1 activity is present, which may explain the hypnotic effect in some cases by itself. However, the aim of the current work was to rationalize an overall "Good sleep" profile that we have described based on seven outcome variables, which comprises four variables for efficacy and three related to side-effects. We note the balanced accuracy of a model based on solely HRH1 as input to be 59.2% (compared to 68.4% of the model used in this work) when taking those comprehensive sleep profiles into account. The model employing HRH1 as a single input variable showed a high true positive rate (of 0.880) for compounds presenting an optimal sleep profile, but a low true positive rate (of 0.331) for the non-optimal sleep set. This suggests that classification based purely on a single receptor is not sufficient for the purposes described in this work, especially when comparing to the true positive rates for optimal and non-optimal sleep in the model used in this work, which were 0.679 and 0.689, respectively.

Furthermore, we compared our predictions to the available in-house Lilly bioactivity data for 30 compounds in the SCORE2004TM data set which are also contained in DrugBank. More specifically, the available experimental activity values on CHRM1, CHRM4, ADRA1D, DRD1B, DRD2, HTR2A and HRH1 were compared to our predictions (shown in Supporting Table 5). DrugBank compounds are shown with their maximum Tanimoto coefficient (nearest neighbour based on Molprint2D fingerprints) compared to all molecules in the training file. The performance is shown in Supporting Table 6 both for overall (76.92% TP, 76.39% TN) and for when compounds in the training set were removed regardless of their annotations (64.71% TP, 78.79% TN).

Prospective *in vivo* validation

For each molecule selected for testing, its predicted profile was assessed *in vitro* for confirmation of the predicted protein targets. The comparative pairs for predicted and measured outcomes are shown in Table 2, with 21 of 28 predictions being correct. Additionally, the molecules tested were also assessed for similarity with those in the *SCORE2004*^{TM} training set, using the Tanimoto similarity coefficient (Tc). Apart from loratadine (already stated to be in the training set; Tc = 1), the nearest neighbour in the training set ranged from 0.17 to 0.38 for all other molecules. This information, along with the therapeutic classifications for each molecule, is provided in Supporting Table 7. Furthermore, the nearest neighbour containing a sleep-related keyword/indication (sedation, somnolence, etc.) and the Tanimoto coefficient for each molecule is shown in Supporting Table 8. We note the prevalence of DrugBank compounds annotated with a sleep-related keyword (sleep, drowsiness, sedation, insomnia, somnolence, etc.) to be 6% at the time this work was carried out, to establish the chance of random selection of a sleep-promoting agent.

It was found in our prospective *in vivo* testing that 5 out of 7 compounds suggested for experimental follow-up promoted a sedative-hypnotic effect. All compounds are visualised on the appropriate leaves which follows their predicted bioactivity profile on the decision tree in Supporting Figure 3, including their *in vitro* and *in vivo* outcomes. The post-dosing changes in all the sleep measures used for each of the compounds are shown in Figure 2. These changes are compared to the optimal and acceptable ranges for each sleep-related output variable detailed in the Methods section of this work.

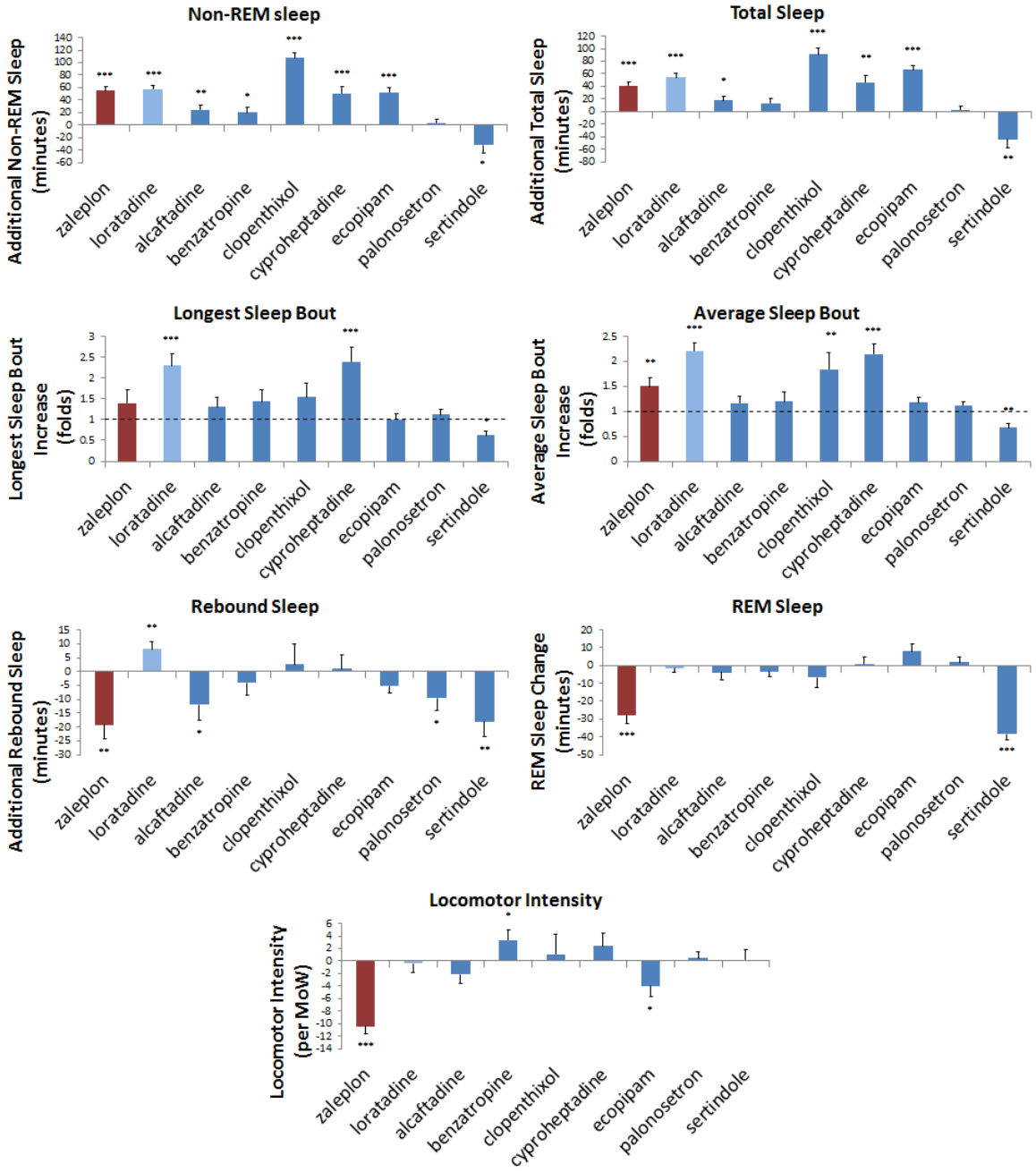

Figure 2 | Sedative-hypnotic effects for each compound in the test set as well loratidine (training) and zaleplon (marketed sedative-hypnotic). (a) Increase in NREM sleep over 6 hrs (b) total sleep over 6 hrs (c) fold change in longest and average (d) sleep bout over 6 hours (e) change in NREM sleep between hrs 6-9 post dose (f) change in REM sleep over 12 hours (g) change in locomotor activity per minute of wake. Asterisks highlight the significance of the readout, where the p-value is less than 0.05, 0.01 and 0.001 for 1, 2 and 3 asterisks respectively. All molecules except palonosetron have an effect on the sleep mechanism, three of which even outperform the marketed sedative-hypnotic zaleplon for both efficacy and side-effect. Hence, this provides validation for the bioactivity prediction-based decision tree model for the understanding of polypharmacological drug mechanisms complex in a CNS disorder.

Alcaftadine showed a significant increase in non-REM sleep and total sleep in the first 6 hours post-dosing (24.4 and 17.8 minutes respectively), but not in the sleep bout measures. Both sleep bout measures however were in the acceptable range. Benzatropine showed an increase in non-REM sleep (19.74 minutes), with no other significant efficacy effects. Furthermore, benzatropine readouts were in the optimal range across all side-effects. Palonosetron had no effect on either efficacy or side-effect parameters. Ecopipam significantly increased non-REM sleep by 52.7 minutes and total sleep by 66.6 minutes in the first 6 hours. Ecopipam was also found to significantly decrease locomotor intensity, but not outside the acceptable range. The only variables for which ecopipam did not report a successful outcome were on the sleep bout measures, which were not significantly affected. Cyproheptadine augmented accumulated non-REM sleep significantly (by 50.6 minutes) in the first 6hrs post-dosing with increased sleep bout continuity. Here, the longest bout was increased 2.4-fold and the average bout was increased by 2.15 fold. Cyproheptadine did not have an effect on the side-effect variables, leaving REM sleep unaffected. For the majority of sleep measures, clopenthixol was in the optimal range, significantly increasing non-REM sleep by 107 minutes, as well as total sleep by 91.4 minutes. Clopenthixol also increased average sleep bout by 1.84 fold with no significant effect on REM sleep. Finally, contrary to our predictions, sertindole *increased* wakefulness by 44.9 minutes, and disrupted longest and average sleep bouts (0.62 and 0.67 fold compared to vehicle). It profoundly reduced REM sleep by 38.6 minutes, but had no effect on locomotor activity. The post-dosing changes in all the sleep measures used for each of the compounds are shown in Figure 2, against a marketed sedative-hypnotic, zaleplon, which was also run in *SCORE2004*^{TM} and used as a benchmark for comparison. Raw outcomes can be found in Supporting Table 9. It can be seen that although Zaleplon has reported increases in non-REM and total sleep (54.9 and

40.9 minutes respectively), it significantly inhibits REM sleep by 27.9 minutes and decreases locomotor activity by 10.5 counts per MoW. Furthermore, it caused a loss of 19.3 minutes of non-REM sleep between 6 and 9 hours post-dosing (rebound insomnia), which is on the border between our acceptable and unacceptable thresholds outlined in the Methods section. The inclusion of measurable side-effects means that zaleplon would actually *fail* to pass our criteria defined earlier in this work for promoting optimal effects on sleep. On the other hand, loratadine and cyproheptadine both matched our predicted bioactivity profile for optimal sleep promoting effects, and led to readouts which are within the optimal range for all seven sleep variables. We note that other side effects produced by cyproheptadine via its anticholinergic activity would likely preclude it from being used for sleep medication. Finally, DrugBank compounds missed by our analysis but annotated as sedative-hypnotic or have a secondary sedative effect are shown in Supporting Table 10.

To summarise, in this work we have presented an experimentally-validated computational analysis pipeline, which can be employed across disease areas and target classes for the prediction of phenotypically active drugs with different pharmacologies in a complex CNS disorder. Moreover, we have demonstrated the added merit of using *in silico* bioactivity predictions, by showing in this case that models built using predicted bioactivities outperform those using only structural information, when predicting a phenotypic effect.

METHODS

Sleep/Wake analysis

For details on the animal experiments please refer to the Supporting Information (Supporting Text Box 1).

The compound set from *SCORE2004*™

The dataset used for this work comprised 845 compound-phenotypic effect pairs in rat, for which outcomes were measured at different doses and administration routes. This data set represented all compounds tested in the CT-18 format, which has been run consistently for several years at Lilly and Hypnion, excluding those from ongoing drug discovery projects at the time of commencing the project. It is therefore representative of particular chemotypes of target classes prone to inducing sleep, rather than a random sampling of drug-like space. For each unique compound, the best sleep outcome was retained based on the set thresholds, resulting in 491 unique compound-effect pairs.

For those compounds, seven parameters were measured using the *SCORE2004*™ system (4 for 'efficacy' and 3 for 'side-effect') following dosing at CT-18, and used to create a simplified description of an 'optimal sleep-promoting profile', based on arbitrary fixed thresholds for each outcome variable determined using a combination of literature and proprietary compounds. The variables used for describing increases in sleep are; increase in non-rapid eye movement (NREM) sleep over the 6hrs post dosing vs. vehicle, total sleep for 6hrs post-dosing vs. vehicle, the longest sleeping bout and the average of the sleep bouts in the 6 hours period averaged over the treatment group. With respect to the vehicle (control) animal, the first two measures were calculated in minutes, whereas the latter two in x-fold relative to vehicle (due to their log-normal

distribution). The variables used to describe 'side-effect' were rebound insomnia (cumulative NREM sleep 6-9hrs post-dosing vs. vehicle), REM inhibition (cumulative REM sleep for 12 hours post-dosing vs. vehicle) and locomotor activity during wake, for 6 hours post-dosing. The first two measures were calculated in minutes compared to vehicle, whereas locomotor activity was measured as counts per minute of wake (MoW). Locomotor activity used in this context does not address the next day sedation, but to the acute motor deficits linked to problems on wakening during the night, such as falls and disorientation. The names of these variables, as well as their description is summarised in Table 3.

Table 3 | Names, brief descriptions and types of the seven output variables from *SCORE2004*™ used to define optimal sleeping patterns and side-effects.

Variable Name	Variable Description	Variable Type
NREM6hr	Cumulative non-REM sleep in the first 6 hours post dosing compared to vehicle	Efficacy
Sleep6hr	Cumulative total sleep in the first 6 hours post dosing compared to vehicle	Efficacy
LBout	Longest sleep bout in the first 6 hours post dosing compared to vehicle	Efficacy
AvgAvgBout	Average of the first 6 average hourly sleep bouts post dosing compared to vehicle	Efficacy
RebIns	Rebound insomnia; the cumulative non-REM sleep between hours 6-9 hours post dosing compared to vehicle	Side-effect
REMinh	REM sleep inhibition; the cumulative REM-sleep in the first 12 hours post dosing compared to vehicle	Side-effect
LMinh	Locomotor inhibition; the cumulative locomotor Activity per minute of Wake (MOW) time in the first 6 hours compared to vehicle	Side-effect

Analysis of covariance (ANCOVA) was used to estimate drug-effects using the equivalent baseline sleep outcome value as a covariate. LS mean differences to vehicle for each treatment level were calculated across the relevant time periods for each parameter. Two fixed thresholds were defined for all variables, one optimal and one acceptable. The thresholds were based on

expert opinion by comparison to compounds that have been advanced to clinical studies. For each variable compounds are scored between zero and one. Compounds received a score of zero for a variable if they were below the acceptable threshold, between 0.5 and 1 for values between the acceptable and the optimal threshold and 1 for anything equal or better than the optimal threshold. Compounds with a score of 5.5 or more were considered as optimal sleep-promoting molecules. This ensured that compounds would be in the optimal range for at least 4 variables (majority of outcomes) and in the acceptable range for all the rest. The thresholds and measures (minutes, folds) for each variable are summarised in Table 4. This resulted in 218 compounds passing our overall threshold for an "optimal sleep profile" against 273 which did not (out of a total 491), which gave rise to a rather balanced dataset overall (with 44.4% of compounds in one class vs 55.6% in the other class) which has rather amenable to further algorithmic classification.

Table 4 | List of optimal and acceptable thresholds for the *in vivo* rat sleep model readouts (*SCORE2004*™), explanations of their usage and the units in which they are measured.

Variable Name	Optimal Threshold	Acceptable Threshold	Usage and Units Explanation
NREM6hr	>=30	>= 10	<Threshold> minutes or more increase compared to vehicle
Sleep6hr	>=25	>=10	<Threshold> minutes or more increase compared to vehicle
LBout	>=1.75	>=1.2	<Threshold>-fold or more increase compared to vehicle
AvgAvgBout	>=1.5	>=1.1	<Threshold>-fold or more increase compared to vehicle
RebIns	>=-10	>=-20	No more than <Threshold> minutes decrease compared to vehicle
REMinh	>=-5	>=-10	No more than <Threshold> minutes decrease compared to vehicle
LMIinh	>=-2.5	>=-5	No more than <Threshold> counts per MoW decrease compared to vehicle

Compound pre-processing and i*n silico* target prediction

The 491 unique compounds were subjected to a KNIME[45] pipeline described and utilised in recent studies.[12,13] In particular, using MOE[46] nodes all structures were converted to 2D, neutralised, salts were stripped and molecules were checked for tautomeric duplicates. Once compounds were standardised, Molprint2D[47] fingerprints of depth=3 were generated using the OpenBabel[48] extension. Stereochemistry was not addressed in this work as it is not captured by Molprint2D circular fingerprints, although it is known that different stereoisomers may in some cases have different effects on proteins. This is due to the nature of the method employed here, which generally aims to provide reasonable on-target activity hypotheses, which then can (and need to) be followed up experimentally.

Bioactivities were predicted with the use of a Laplacian-modified Naïve Bayes classifier described by Koutsoukas *et al.*,[10] trained on ~189k ligand-target pairs spanning over 477 human protein targets. Data was extracted from ChEMBL 14 according to the original publication criteria,[10] with the exception of the minimum number of instances which was set to 50, for which the binding affinity threshold was 1μM or better.[13] Class-specific score thresholds were employed in order to improve prediction accuracy according to Drakakis *et al.*,[49] which have recently been used in several studies.[12,13]

Decision tree generation and bioactivity profiles

Compounds were classified based on whether they promote an optimal sleeping pattern or not, with the use of predicted bioactivities as variables. The algorithm chosen was the WEKA[50] implementation of the C4.5[51] classification tree due to its interpretable output (distinct distinct polypharmacological bioactivity profiles derived by reading node sequences from the root to each leaf; such as if compound active on targets x, y, and z then ‘optimal sleeping profile).

However, it was necessary to optimise the algorithmic parameter selection (number of attributes, minimum number of instances in leaves, confidence score) in order to improve both the accuracy and interpretability of the output, meaning the number of correctly classified instances, as well as size of the final decision tree, respectively. Confidence scores of 0.25, 0.5 and 0.75 were assessed for classification accuracy, number of leaves in tree and overall tree size, over a wide range of features (number of predicted targets) to be used as attributes. More specifically, the number of features was varied exhaustively in bins of ten, starting from a minimum number of ten targets. Next, the optimal minimum number of instances for each leaf from one to nine was then assessed. Finally, based on fixed confidence score and minimum number of instances in leaves, the number of features used was optimised globally for both classification accuracy and tree depth, using 10-fold cross-validation. All polypharmacological bioactivity profiles (predicted activity on more than one protein target) from the decision tree leading to an 'optimal sleeping profile' classification were recorded. Feature correlation analysis was performed in order to assess whether the multi-target bioactivity profiles were a result of correlated features or not. This was carried out in KNIME using Cramér's V[52]. This association measure comprises Pearson's chi square test[53] normalized to the [0, 1] range. It is calculated as the square root of the chi-squared statistic divided by the number of instances and the smallest dimension (rows or columns) minus one.

Assessing bioactivity predictions on DrugBank molecules

From the bioactivity profiles recorded, the available in-house Lilly compound-protein target data was used to assess the bioactivity prediction accuracy. More specifically, all available experimental activity data for 30 of these DrugBank[41] molecules against CHRM1, CHRM4, ADRA1D, DRD1B, DRD2, HTR2A and HRH1 were compared to the *in silico* predictions.

Furthermore, for each compound selected for *in vivo* testing described later in this work, the full predicted bioactivity profile was subjected to prospective *in vitro* testing. It is noted that even though rat *in vitro* models are generally less available, the analysis carried out by Kruger and Overington[54] showed robustness of small molecule binding across species. Specifically, using 2782 compound instances over 151 pairs of orthologous proteins they reported a significant linear relationship between bioactivities measured against human and rat targets (Pearson's correlation coefficient $r=0.71$, $p<2e\text{-}16$). In our experience the correlation between rat and human activity is high enough to justify using the far larger sources of human data to train models.

Compound selection for *in vivo* testing

All marketed molecules in DrugBank[41] were subjected to the same pipeline as our training set (described earlier) for molecule pre-processing and target prediction. Based on their predicted bioactivity spectra, compounds which followed the recorded bioactivity profiles were predicted as sleep-promoting (or at least to affect the sleep mechanism, as agonism/antagonism are not taken into account for the *in silico* bioactivity predictions). An extensive literature search was performed for all compounds in DrugBank, SIDER and SciFinder using keywords such as sleep, hypnotic, sedation, drowsiness, somnolence etc. Finally, compounds were ranked according to selectivity (smallest number of total predicted targets), and assessed for their blood-brain barrier permeation score according to the model developed by Kortagere *et al.*[43] Two sets of compounds were selected for subsequent *in vivo* validation in *SCORE2004*™. The first set consisted of compounds both partially and fully matching the most populated polypharmacological profile, for investigation of the most important targets in the main derived bioactivity profile. The second comprised a subset of those with no sleep-related annotations in literature.

Prospective *in vivo* validation

The drugs selected for experimental follow-up and further discussion were alcaftadine, benzatropine, cyproheptadine, ecopipam, clopenthixol, palenosetron, serindole and loratidine. Compounds tested in this set were run at a single dose. Dose, route and vehicle were selected based on available literature data suggesting *in vivo* activity and include the following; loratidine (confirming sleep-promoting activity) 30mg/kg, PO in 0.25% methylcellulose, alcaftadine 30mg/kg IP in 0.25% hydroxyethylcellulose, benzatropine 1mg/kg IP in 0.25% methylcellulose, cyproheptadine 10mg/kg PO in 0.25% methylcellulose, ecopipam 0.3mg/kg SC in 0.25% hydroxyethylcellulose, palenosetron 3mg/kg PO in 0.25% methylcellulose, sertindole 3mg/kg PO in 0.25% methylcellulose, clopenthixol 2mg/kg IP in 0.25% hydroxyethylcellulose. We note that we used a relatively high dose of 30 mg/kg IP for alcaftadine (which has been approved as a topical drug) to investigate its systemic effects in the experiments conducted here. However, while this can still be seen as a validation of the method described in this work, this compound and dose are unlikely to have practically relevant implications. Studies were run in a parallel dosing paradigm, with vehicle control groups made for comparison to drug treated animals. Pretreatment baseline data was used as the covariate in the analysis so was utilized for comparing drug effects. Animals were given a washout period of at least one week before any further treatment

SUPPORTING INFORMATION

The Supporting Information is available free of charge on the ACS Publications website at DOI:-

- Supporting Tables 1 – 10, Figures 1, 2, 3 and details on the sleep/wake analysis carried out in this work. (PDF)

References

1. Koutsoukas, A., Simms, B., Kirchmair, J., Bond, P. J., Whitmore, A. V., Zimmer, S, Young, M. P., Jenkins, J. L., Glick, M, Glen, R. C. & Bender, A. From in silico target prediction to multi-target drug design: Current databases, methods and applications. *J Proteomics* **74,** 2554–2574 (2011).

2. Feng, Y., Mitchison, T. J., Bender, A., Young, D. W. & Tallarico, J. A. Multi-parameter phenotypic profiling : using cellular effects to characterize small-molecule compounds. *Nat Rev Drug Discov* **8,** 567–578 (2009).

3. Hopkins, A. L. Network pharmacology : the next paradigm in drug discovery. *Nat Chem Biol* **4,** 682–690 (2008).

4. Roth, B. L., Sheffler, D. J. & Kroeze, W. K. Magic shotguns versus magic bullets: selectively non-selective drugs for mood disorders and schizophrenia. *Nat Rev Drug Discov* **3,** 353–359 (2004).

5. Morphy, R., Kay, C. & Rankovic, Z. From magic bullets to designed multiple ligands. *Drug Discov Today* **9,** 641–651 (2004).

6. Morphy, R. & Rankovic, Z. Designed multiple ligands. An emerging drug discovery paradigm. *J Med Chem* **48,** 6523–6543 (2005).

7. Anighoro, A., Bajorath, J. & Rastelli, G. Polypharmacology : Challenges and Opportunities in Drug Discovery. *J Med Chem* **57,** 7874–7887 (2014).

8. Youdim, M. B. H. & Buccafusco, J. J. Multi-functional drugs for various CNS targets in the treatment of neurodegenerative disorders. *Trends Pharmacol Sci* **26,** 27–35 (2005).

9. Hopkins, A. L., Mason, J. S. & Overington, J. P. Can we rationally design promiscuous drugs? *Curr Opin Struct Biol* **16,** 127–136 (2006).

10. Koutsoukas, A., Lowe, R., KalantarMotamedi, Y., Mussa, H. Y., Klaffke, W., Mitchell, J. B. O., Glen, R. C. & Bender, A. In silico target predictions: defining a benchmarking data set and comparison of performance of the multiclass Naïve Bayes and Parzen-Rosenblatt window. *J Chem Inf Model* **53,** 1957–1966 (2013).

11. Mervin, L.H., Afzal, A.M., Drakakis, G., Lewis, R., Engkvist, O. & Bender, A. Target prediction utilising negative bioactivity data covering large chemical space. *J Cheminform* **7** (51), (2015)

12. Drakakis, G., Hendry, A. E., Hanson, K. M., Brewerton, S. C., Bodkin, M. J., Evans, D. A., Wheeler, G. N. & Bender, A. Comparative Mode-of-Action Analysis Following

Manual and Automated Phenotype Detection in Xenopus laevis. *Med Chem Commun* **5,** 386–396 (2014).

13. Liggi, S., Drakakis, G., Hendry, A. E., Hanson, K. M., Brewerton, S. C., Wheeler, G. N., Bodkin, M. J., Evans, D. A. & Bender, A. Extensions to In Silico Bioactivity Predictions Using Pathway Annotations and Differential Pharmacology Analysis: Application to Xenopus laevis Phenotypic Readouts. *Mol Inf* **32,** 1009–1024 (2013).

14. McKenna, J. T., Cordeira, J. W., Christie, M. A., Tartar, J. L., McCoy, J. G., Lee, E., McCarley, R. W. & Strecker, R. E. Assessing sleepiness in the rat: a multiple sleep latencies test compared to polysomnographic measures of sleepiness. *J Sleep Res* **17,** 365–375 (2009).

15. Siegel, J. M. Clues to the functions of mammalian sleep. *Nature* **437,** 1264–1271 (2005).

16. Vassalli, A. & Dijk, D. J. Sleep function: Current questions and new approaches. *Eur J Neurosci* **29,** 1830–1841 (2009).

17. Sejnowski, T. J. & Destexhe, A. Why do we sleep? *Brain Res* **886,** 208–223 (2000).

18. Hobson, J. A. & Pace-Schott, E. F. The cognitive neuroscience of sleep: neuronal systems, consciousness and learning. *Nat Rev Neurosci* **3,** 679–693 (2002).

19. Rechtschaffen, A. & Kales, A. *A manual of standardized terminology, techniques and scoring system for sleep stages of human subjects*. (U.S. Dept. of Health Education and Welfare Public Health Services-National Institutes of Health National Institute of Neurological Diseases and Blindness Neurological Information Network, 1968).

20. Ebert, B., Wafford, K. A. & Deacon, S. Treating insomnia: Current and investigational pharmacological approaches. *Pharmacol Ther* **112,** 612–629 (2006).

21. Wafford, K. a & Ebert, B. Emerging anti-insomnia drugs: tackling sleeplessness and the quality of wake time. *Nat Rev Drug Discov* **7,** 530–540 (2008).

22. American Academy of Sleep Medicine. *International Classification of Sleep Disorders: Diagnostic and Coding Manual*. (American Academy of Sleep Medicine, 2005).

23. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders, DSM-5*. (American Psychiatric Publishing, 2013).

24. Whiting, P. J. GABA-A receptors: a viable target for novel anxiolytics? *Curr Opin Pharmacol* **6,** 24–29 (2006).

25. Costa, E., Auta, J., Grayson, D. R., Matsumoto, K., Pappas, G. D., Zhang, X., Guidotti, A. GABA-A receptors and benzodiazepines : a role for dendritic resident subunit mRNAs. *Neuropharmacology* **43,** 925–937 (2002).

26. Fitzgerald, A. C., Wright, B. T. & Heldt, S. A. The behavioral pharmacology of zolpidem: Evidence for the functional significance of α1-containing GABA(A) receptors. *Psychopharmacology (Berl).* **231,** 1865–1896 (2014).

27. Brown, R. E., Basheer, R., McKenna, J. T., Strecker, R. E. & McCarley, R. W. Control of Sleep and Wakefulness. *Physiol Rev* **92,** 1087–1187 (2012).

28. Saper, C. B. & Scammell, T. E. Emerging therapeutics in sleep. *Ann Neurol* **74,** 435–440 (2013).

29. Simons, F. E. R. Advances in H1-antihistamines. *New Engl J Med* **351,** 2203–2217 (2004).

30. Monti, J. M. & Monti, D. Histamine H1 Receptor Antagonists in the Treatment of Insomnia. *CNS Drugs* **13,** 87–96 (2000).

31. Feighner, J. P. Mechanism of action of antidepressant medications. *J Clin Psychiat* **60,** 4–11 (1999).

32. Ongini, E., Bonizzoni, E., Ferri, N., Milani, S. & Trampus, M. Differential Antagonist Effects of Dopamine D-1 and D-2 Receptor Antipsychotics on Sleep-Wake Patterns in the Rat. *J Pharmacol Exp Ther* **266,** 726–731 (1993).

33. Blin, O. A comparative review of new antipsychotics. *Can J Psychiatry* **44,** 235–244 (1999).

34. Monti, J. M. & Monti, D. Sleep in schizophrenia patients and the effects of antipsychotic drugs. *Sleep Med Rev* **8,** 133–148 (2004).

35. Arendt, J. & Rajaratnam, S. M. W. Melatonin and its agonists: an update. *Brit J Psychiat* **193,** 267–269 (2008).

36. Kryger, M. H., Roth, T. & Dement, W. C. in *Principles and Practice of Sleep Medicine: Expert Consult Premium Edition - Enhanced Online Features* (Elsevier Health Sciences, 2010).

37. Sheeba, V., Fogle, K. J., Kaneko, M., Rashid, S., Chou,Y.-T., Sharma, V. K., & Holmes, T. C. Large Ventral Lateral Neurons Modulate Arousal and Sleep in Drosophila. *Curr Biol* **18,** 1537–1545 (2009).

38. Morin, C. M. & Wooten, V. Psychological and pharmacological approaches to treating insomnia: critical issues in assessing their separate and combined effects. *Clin Psychol Rev* **16,** 521–542 (1996).

39. Cardinali, D. P., Srinivasan, V., Brzezinski, A. & Brown, G. M. Melatonin and its analogs in insomnia and depression. *J Pineal Res* **52,** 365–375 (2012).

40. McCarroll, M. N., Gendelev, L., Keiser, M. J. & Kokel, D. Leveraging Large-scale Behavioral Profiling in Zebrafish to Explore Neuroactive Polypharmacology. *ACS Chem Biol* **11** (4), 842-849 (2016)

41. Wishart, D. S., Knox, C., Guo, A. C., Cheng, D., Shrivastava, S., Tzur, D., Gautam, B. & Hassanali, M. DrugBank: a knowledgebase for drugs, drug actions and drug targets. *Nucleic Acids Res* **36,** D901–D906 (2008).

42. Hietala, J., Seppälä, T., Lappalainen, J. & Syvälahti, E. Quantification of SCH 39166, a novel selective D1 dopamine receptor antagonist, in rat brain and blood. *Psychopharmacology (Berl)* **106**(4), 455-458 (1992)

43. Kortagere, S., Chekmarev, D., Welsh, W. J. & Ekins, S. New predictive models for blood-brain barrier permeability of drug-like molecules. *Pharm Res* **25,** 1836–1845 (2008).

44. Trampus, M. & Ongini, E. The D1 dopamine receptor antagonist SCH 23390 enhances REM sleep in the rat. *Neuropharmacology* **29,** 889–893 (1990).

45. Berthold, M. R., Cebron, N., Dill, F. & Gabriel, T. R. in *Studies in Classification, Data Analysis, and Knowledge Organization (GfKL 2007)* **11,** 319–326 (Springer, 2007).

46. Chemical Computing Group Inc. *Molecular Operating Environment (MOE)*. (2011).

47. Bender, A., Mussa, H. Y., Glen, R. C. & Reiling, S. Similarity searching of chemical databases using atom environment descriptors (MOLPRINT 2D): evaluation of performance. *J Chem Inf Comp Sci* **44,** 1708–1718 (2004).

48. O'Boyle, N. M., Banck, M., James, C. A., Morley, C., Vandermeersch, T. & Hutchison, G. R. Open Babel: An open chemical toolbox. *J. Cheminform.* **3,** 33–46 (2011).

49. Drakakis, G., Koutsoukas, A., Brewerton, S. C., Bodkin, M. J., Evans, D. A. & Bender, A. Comparing Global and Local Likelihood Score Thresholds in Multiclass Laplacian-Modified Naïve Bayes Protein Target Prediction. *Comb Chem High Throughput Screen* **18,** 323–330 (2015).

50. Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P & Witten, I. A. The WEKA Data Mining Software: An Update. *SIGKDD Explor Newsl* **11,** (2009).

51. Quinlan, J. R. *C4.5: Programs for Machine Learning*. (Morgan Kaufmann Publishers, 1993).

52. Cramér, H. *Mathematical Methods of Statistics*. (Princeton University Press, 1946)

53. Pearson, K. On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling. *Philosophical Magazine Series 5* **50** (302), 157–175 (1900)

54. Kruger, F.A. & Overington, J.P., Global Analysis of Small Molecule Binding to Related Protein Targets. *PLoS Comput Biol* **8**(1), e1002333 (2012)