nature reviews drug discovery

https://doi.org/10.1038/s41573-026-01496-2

Supplementary information

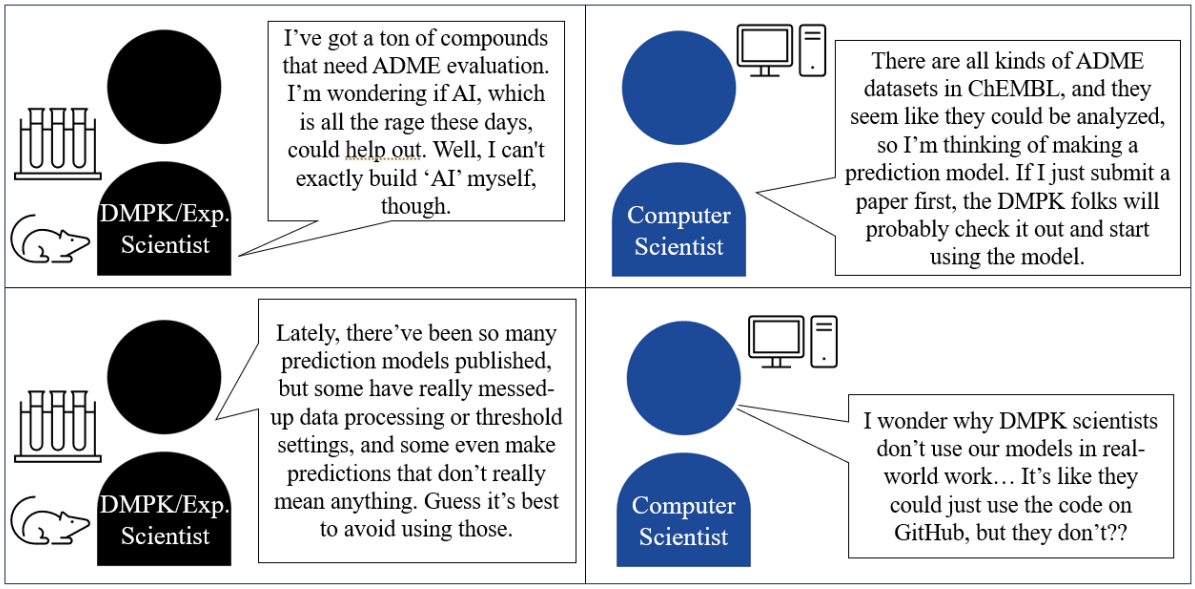
Artificial intelligence in drug discovery — what it is, where we stand and the path forward

In the format provided by the authors and unedited

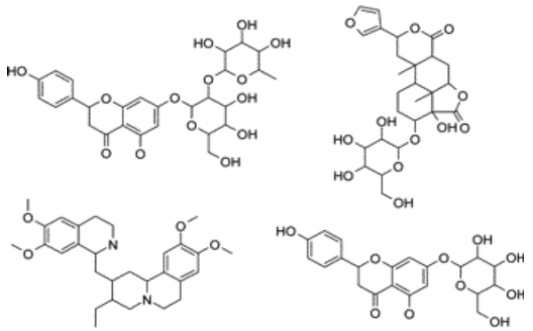
Supplementary Table 1 | The differences between ligand and drug discovery

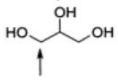
Physicochemical Properties	Relevance for efficacy and safety and relationship to other compound properties	Considered in ligand design	Needs to be in suitable range to become a drug
Molecular Weight	Contributing to various target binding (e.g. lipophilicity drives affinity) and ADME properties (molecular weight and hydrogen bonding is related to blood-brain barrier permeation, lipophilicity is related to clearance, solubility needs to be achieved for absorption in GI environment *etc.*)	yes	
H-Bond Donors/Acceptors			
Rotatable Bonds			
logP/logD			
Solubility (at different pH, in buffer, etc.)		potentially	
....			
Potency and Selectivity			
On-target	On-target activity needs to be compatible with compound dosing and safety	sometimes binding only	
Mutants	E.g. in cancer one might aim to target different mutants at once	if explicitly measured and optimized for	
Different assay endpoints	Compounds usually have different activities in different types of assays (and it is often not clear which assay is predictive for *in vivo*)		
Off-targets within target family	Selectivity needs to be sufficient for tolerable safety at dose in man		
Other off-targets	Selectivity needs to be sufficient for tolerable safety at dose in man		
Cellular toxicity	Unspecific cellular toxicity is linked to organ toxicity	no	
hERG inhibition	If too high risk of cardiac safety events		
AMES test	If positive risk of genetic mutations when applied *in vivo*		
...			
***In vitro* ADME**			
Animal hepatic stability	Compounds with insufficient hepatic stability do not achieve therapeutically effective dose in plasma (or administration is resulting in unsuitable dosing schedule)		
Human hepatic stability			
Caco-2 permeability	Insufficient permeation into cells if too low/compounds are actively effluxed		
CYP inhibition and induction	If unsuitable profile (e.g. strong inhibitors, inducers etc.) higher risk of side effects and drug-drug interactions		
hERG inhibition	If too high risk of cardiac safety events		
AMES test	If positive risk of genetic mutations when applied *in vivo*		
...			
***In Vivo* ADME/PK (often multiple species)**			
Clearance	If too high insufficient blood plasma concentration over time to achieve efficacy/unsuitable dosing schedule		
Volume of distribution	Contributing to overall ADME profile (usually not parameter to be optimized in itself, needs to be in suitable range)		
Half life	If too low insufficient blood plasma concentration over time to achieve efficacy/unsuitable dosing schedule		
Maximum plasma concentration	If too low no efficacy is achieved; if too high risk of safety events		
Bioavailability	If too low no oral administration possible		
Blood-brain barrier permeation and efflux	If too low brain concentration intrinsic compound activities *in vitro* will not translate to expected effects *in vivo*		
...			

This is an extended version of the figure in Box 1 in the main article. It illustrates that for ligand discovery, only a subset of parameters are considered compared to drug discovery.



Supplementary Figure 1 | Illustration of the mental gap between experimental scientists and computer scientists when creating and using predictive models in drug discovery. The area of drug metabolism and pharmacokinetics (DMPK) has been chosen in this illustration, since it reaches far into the in vivo domain of the data generated (which is heavily conditional), and hence often poses problems for people from other domains to understand appropriately to arrive at fit-for-use models in a given practical situation. ADME, absorption, distribution, metabolism, excretion. Adapted with permission from REF. 1).



Fragment Number	Fragment	Frequency Bitter Dataset	Frequency Nonbitter Dataset	Information Gain	Relative Frequency Bitter Dataset	Relative Frequency Nonbitter Dataset
1		120	317	0.0139	18.49%	2.34%
2		76	76	0.0139	11.71%	0.56%
3		116	342	0.0125	17.87%	2.53%

Supplementary Figure 2 | Illustration of the pitfalls of 'Explainable AI' (XAI) in the domain of chemical datasets. In the selected study[2], a dataset of bitter compounds and non-bitter compounds was compiled (examples of bitter compounds are shown on the left), and structures were represented in circular fingerprint format. Subsequently, the information gain associated with features with respect to their ability to discriminate between both classes was calculated, and the highest-ranked features characteristic for the bitter class are shown on the right. It can be seen that features selected to be characteristic for bitterness by the feature selection criterion were found to represent parts of sugar rings, which does not represent a sufficient mechanistic rationale for the observed compound property. Rather, this represents a confounding factor of the underlying dataset, given that natural products are frequently glycosylated[3] as well as bitter. While the selection of spurious features might be obvious in this case, a similar situation is very often encountered in chemical datasets, given the various biases present in the data (such as synthesis bias, analogue bias, confirmation bias and project bias), where the confounding effect of the bias in the data used is often more difficult to detect. (It should be noted that sugar rings can indeed contribute to the bioactivity of natural products[4] and 'bitter sugars' are indeed able to bind to bitter receptors[5] and hence the above relationship is not necessarily wrong in every single case. However, attributing bitterness of natural products to the presence of sugar rings in molecules cannot be seen as a sufficient explanation of this property.)

References

1. Handa, K. et al. Computational approaches to DMPK: A realistic assessment of current methods and their practical impact. Part I: Physicochemical and in vitro properties. *Drug Discov. Today* **30**, 104422 (2025).

2. Rodgers, S., Glen, R.C. & Bender, A. Characterizing bitterness: identification of key structural features and development of a classification model. *J. Chem. Inf. Model.* **46**, 569–576 (2006).

3. He, B. et al. Glycosyltransferases: Mining, engineering and applications in biosynthesis of glycosylated plant natural products. *Synth. Syst. Biotechnol.* **7**, 602–620 (2022).

4. Méndez, C., Salas, J.A. Altering the glycosylation pattern of bioactive compounds. *Trends Biotechnol.* **19**, 449–456 (2001).

5. Fierro, F., Giorgetti, A., Carloni, P., Meyerhof, W. & Alfonso-Prieto M. Dual binding mode of "bitter sugars" to their human bitter taste receptor target. *Sci Rep.* **9,** 8437 (2019).