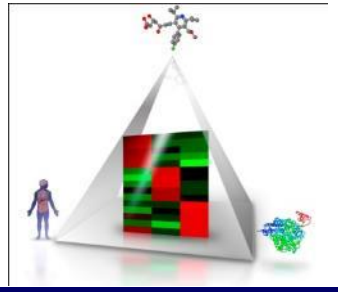
Using Chemical and Biological Data and ML/AI For Predictive Safety – Concepts, Readouts, Applications, Limitations

Andreas Bender, PhD

Natural Philosopher for Molecular Informatics
Department of Chemistry, University of Cambridge

Chief Technology & Information Officer
PangeAI, part of Pangea Botanica, London/Berlin

UNIVERSITY OF CAMBRIDGE



Any statements made during this talk are
in my capacity as an academic

Slides available from http://www.andreasbender.de

Further reading: Artificial Intelligence in Drug Discovery – What is Realistic, What are Illusions? (Parts 1 and 2)

Andreas Bender and Isidro Cortes-Ciriano

Drug Discovery Today 2021

On data, endpoints, models, and predictions

- Preamble: Will ‘AI’ and computational models eat the world? Some historical context
- What we need: *in vivo* relevance
- Data (labelling, conditionality, *in vivo* relevance)
- Models and Validation
  - Descriptors, machine learning, validation
  - “Questions to ask your friend, the modeller”
- Readouts and Applications
  - Cell Painting Assay
  - Applications to mitochondrial toxicity
  - Interpreting high-dimensional cell morphology readouts

On data, endpoints, models, and predictions

- Preamble: Will ‘AI’ and computational models eat the world? Some historical context
- What we need: *in vivo* relevance
- Data (labelling, conditionality, *in vivo* relevance)
- Models and Validation
  - Descriptors, machine learning, validation
  - “Questions to ask your friend, the modeller”
- Readouts and Applications
  - Cell Painting Assay
  - Applications to mitochondrial toxicity
  - Interpreting high-dimensional cell morphology readouts

The 3rd wave of computers in drug discovery (80s, 2000, today) – time for realistic assessment has come

Fortune cover 1981

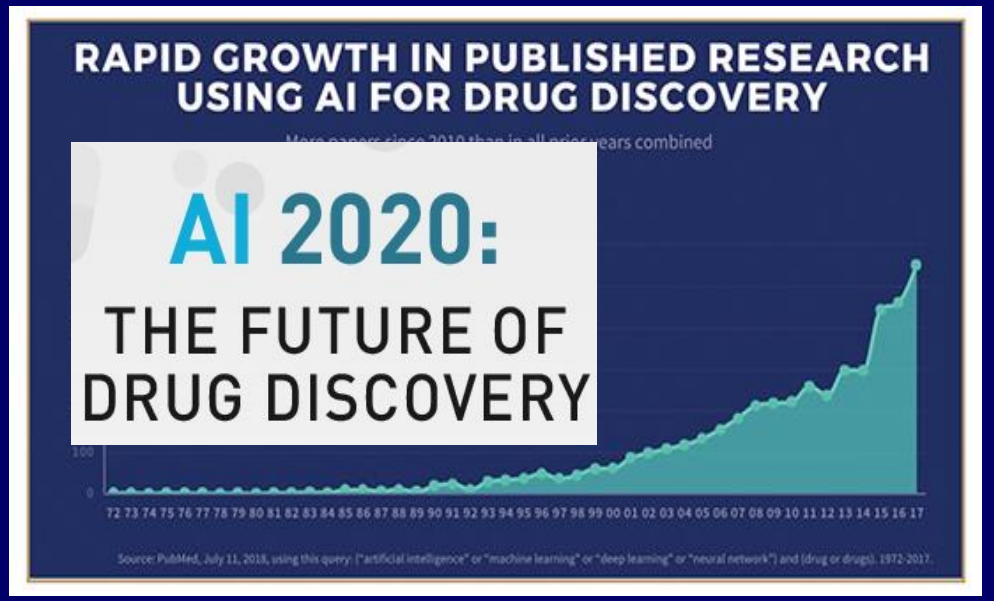
Recent headlines (2018-2020)

SPOTLIGHT · 30 MAY 2018

How artificial intelligence is changing drug discovery

World first breakthrough in AI drug discovery

By **Emma Morriss** - January 30, 2020



Old enough to remember 2000 biotech bubble, Human Genome Project, etc.

T. Reiss, Trends in Biotechnology, 2001:

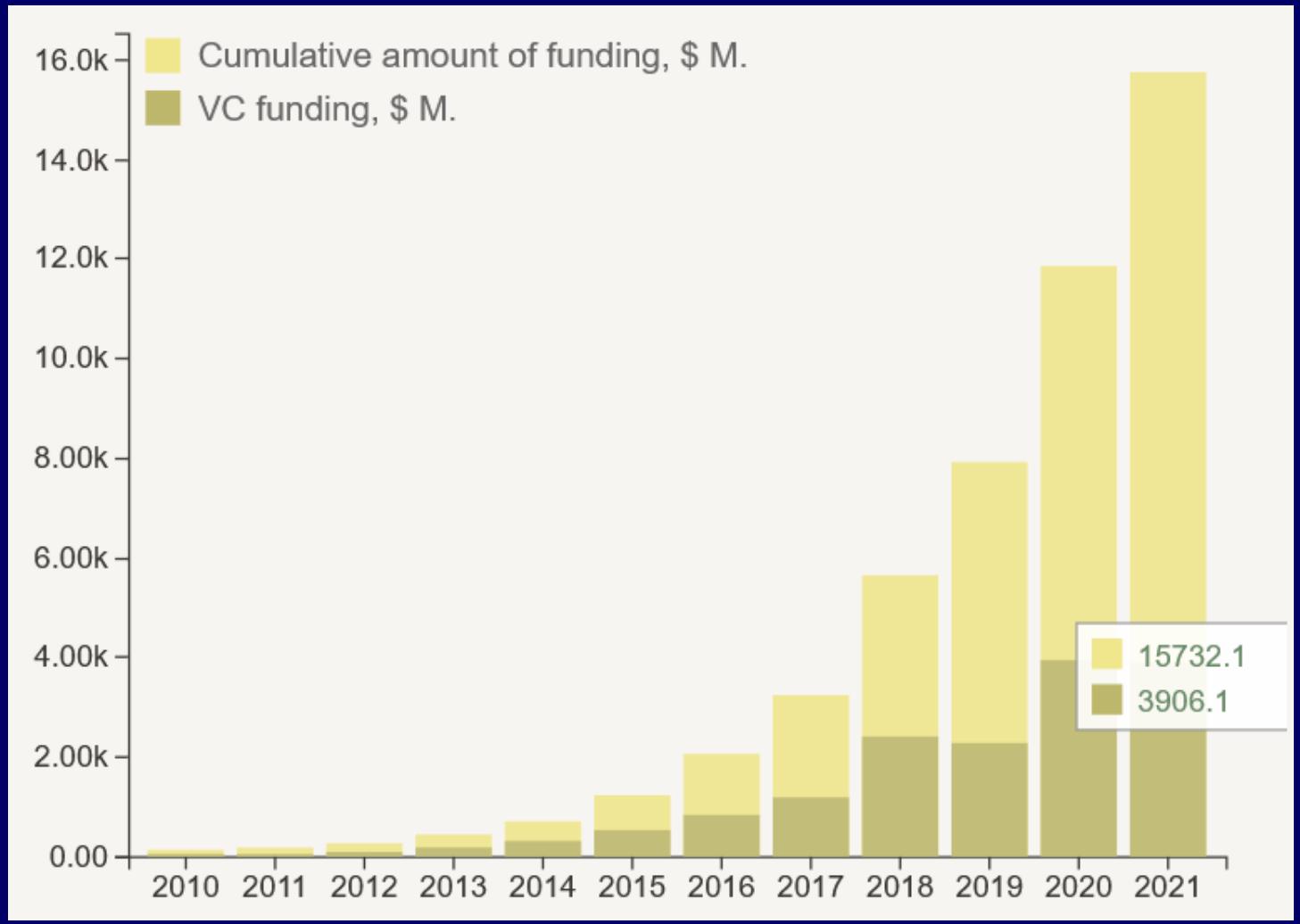
“The number of drug targets will increase by at least one order of magnitude and target validation will become a high-throughput process.”

“More drug targets… 3,000–10,000 targets compared with 483”

Recent (NRDD 2017) estimates of drug targets put the number currently at around *667*

-> How to go from *technology and potential* to *applications/better decisions*?

Funding going into AI in drug discovery 2021: ~$4bn VC funding, $16bn total



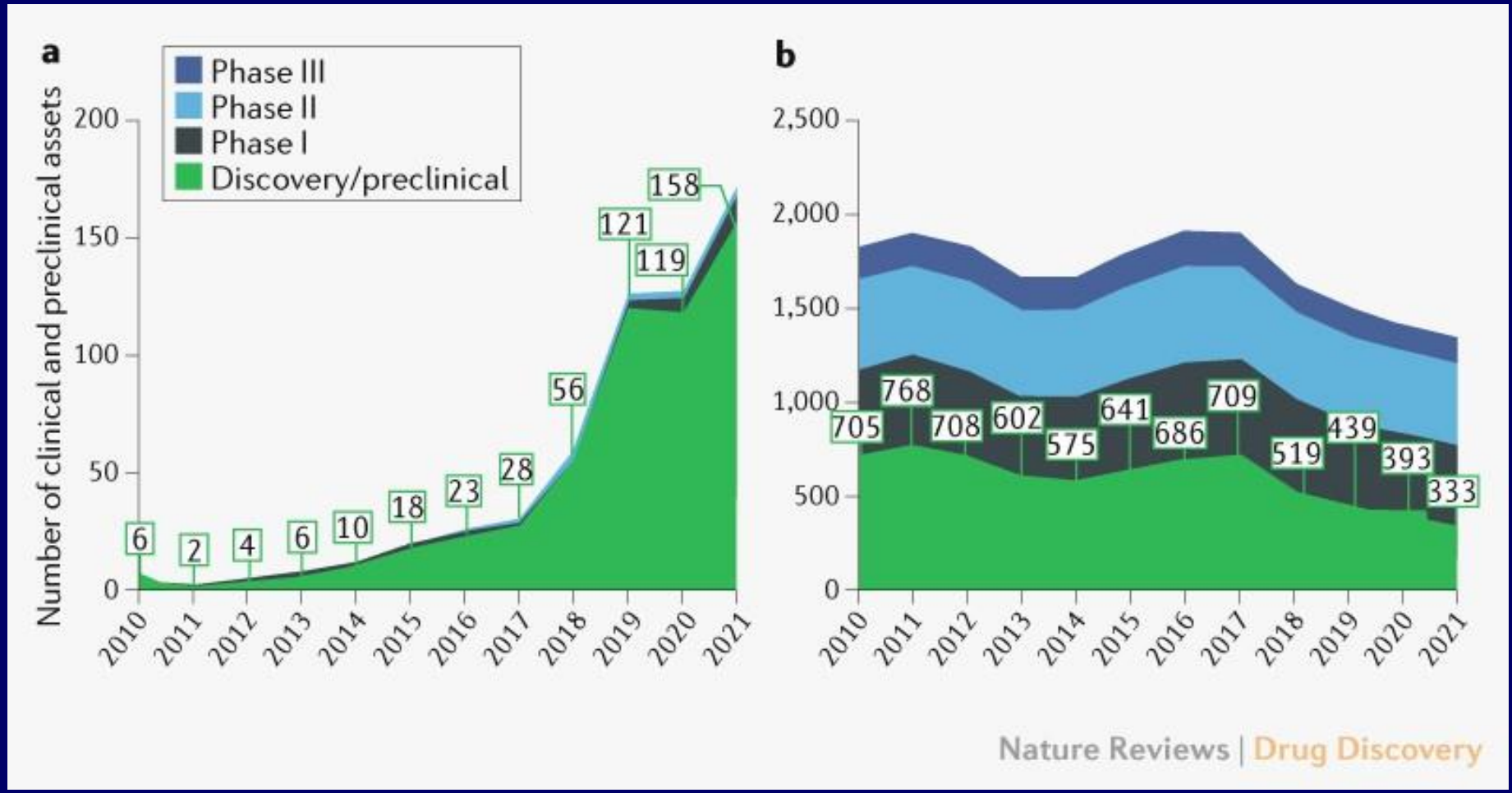
https://www.biopharmatrend.com/post/397-pharmaceutical-artificial-intelligence-in-2021-key-developments-so-far/

Current discovery pipeline: AI-based start-ups vs big pharma

‘AI-native companies’ Top 20 pharma

Significant *number* of *discovery/ preclinical* programs of AI companies (~160 vs ~330)

Very little Phase 1, less Phase 2, no Phase 3



-> Little *in vivo* safety (Phase 1) data *yet*; virtually no *in vivo* efficacy (Phase 2/3) data *yet*

Jayatunga et al., AI in small-molecule drug discovery: a coming wave? *Nature Reviews Drug Discovery* 7 Feb 2022

Conclusion about the world as it is

- No *in vivo* relevant *discovery* coming out of ‘AI’ confirmed *so far*
- Lots of activity in early stage pipeline of AI-first companies, but often already explored targets, close analogues (*Jayatunga et al.)*
- Data is often limiting factor – in both chemical and target space (leads to work on well-explored targets, with more data, less complex pharmacology)
- Appropriate question to ask: Where is the novelty?

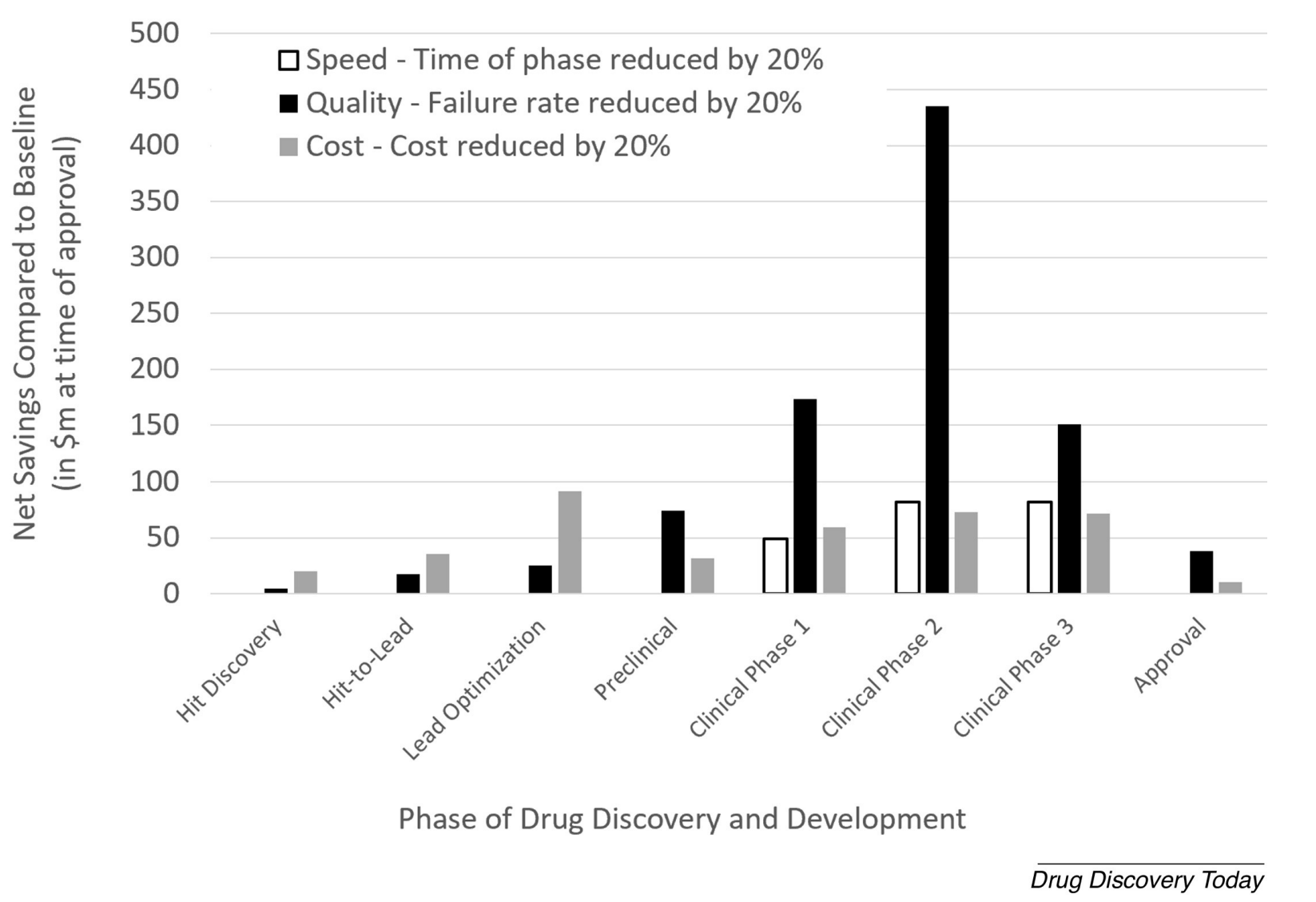
- Is *input* (e.g. funding) success, or *output*?
- The first ‘AI-designed drug’ will be celebrated by the media, but…

… tens of billions went into funding AI in drug discovery, so even the null model would lead to an expected tens of approved drugs

On data, endpoints, models, and predictions

- Preamble: Will ‘AI’ and computational models eat the world? Some historical context
- What we need: *in vivo* relevance
- Data (labelling, conditionality, *in vivo* relevance)
- Models and Validation
  - Descriptors, machine learning, validation
  - “Questions to ask your friend, the modeller”
- Readouts and Applications
  - Cell Painting Assay
  - Applications to mitochondrial toxicity
  - Interpreting high-dimensional cell morphology readouts

The *quality* of *in vivo-relevant* decisions matters more than *speed* and *cost*!



Bender and Cortes, Drug Discovery Today 2021

On data, endpoints, models, and predictions

- Preamble: Will ‘AI’ and computational models eat the world? Some historical context
- What we need: *in vivo* relevance
- Data (labelling, conditionality, *in vivo* relevance)
- Models and Validation
  - Descriptors, machine learning, validation
  - “Questions to ask your friend, the modeller”
- Readouts and Applications
  - Cell Painting Assay
  - Applications to mitochondrial toxicity
  - Interpreting high-dimensional cell morphology readouts

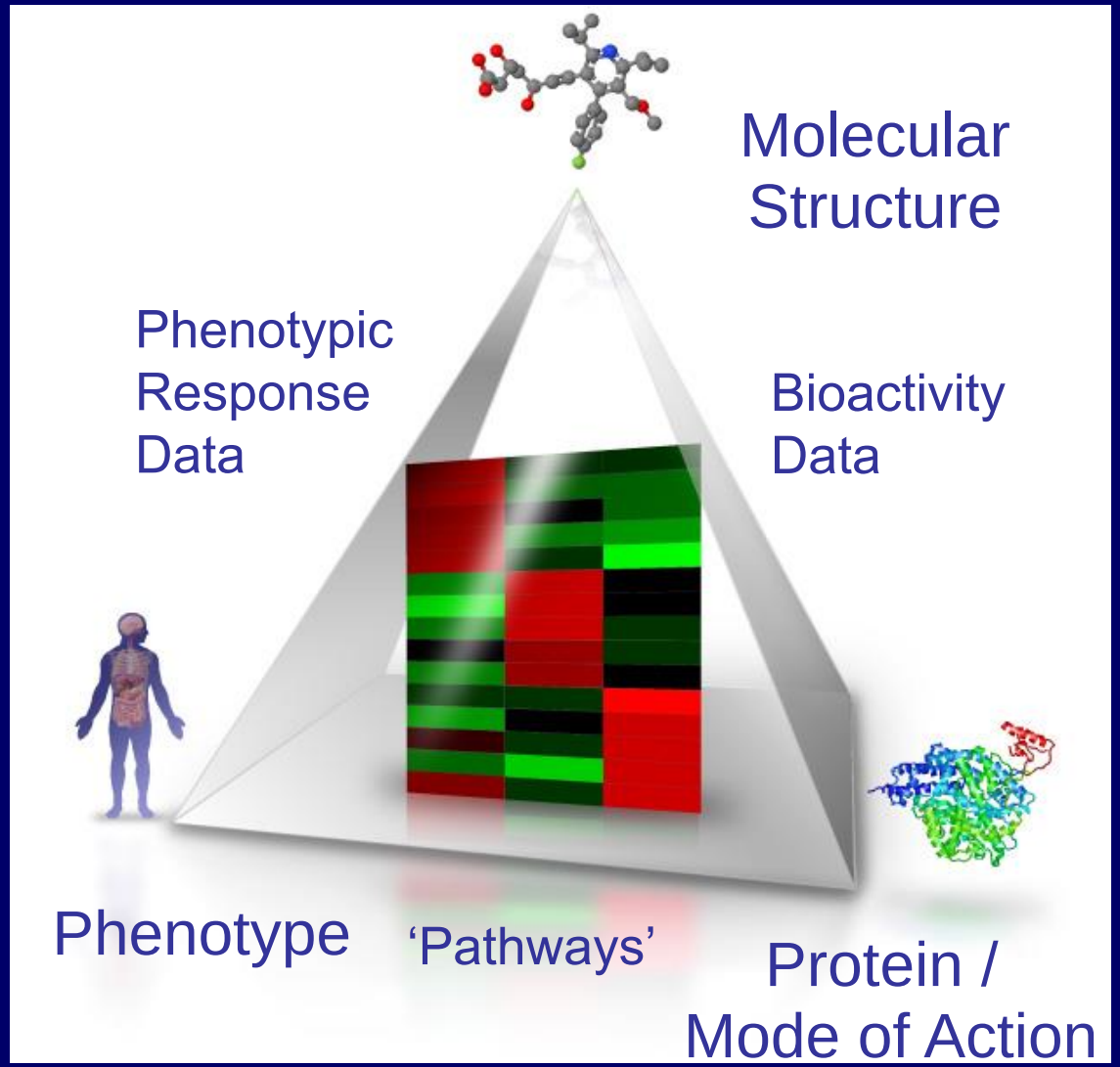
Key point: We often cannot label our data properly in the life sciences

- Machine learning/AI knows *unsupervised* or *supervised* methods
- Predictive methods are (usually) supervised, and need data points with *labels* (active/not active; or quantitative labels, etc.)
- Those labels need to come from *experiments*

- Experiments (and hence labels) often either fall into the ‘large-scale, but little *in vivo* relevance’ or ‘*in vivo* relevant, but small scale and conditional’ category
- **This is a problem** for AI/ML in drug discovery and safety

- So should we use and analyze our data? Absolutely!
- But we need to work towards *in vivo* relevance of data, jointly

A simple view on the world: Linking Chemistry, Phenotype, Targets / Mode of Action (myself, until *ca.* 2010)



a.k.a. "The world is flat"

= "We believe our labels"

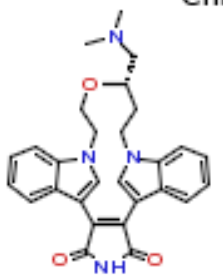
"Compound A is toxic",
"Compound B binds target X",
"Compound C treats disease Y", …

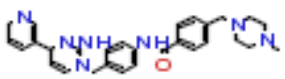
Works in cases where data is large-scale, and homogenous, and we have meaningful labels

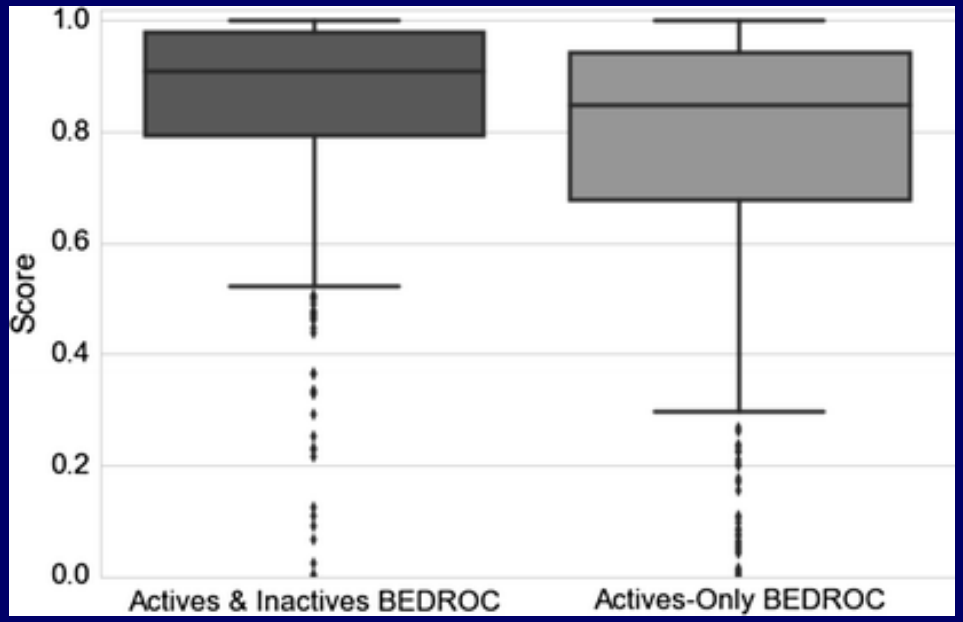
Does not consider data conditionality, e.g. dose, PK, translatability from model system to *in vivo* setup, endotype, genotype, *etc. etc.*

The ‘flat earth’ view can *still* help! Eg Public target prediction model, based on ~200 mio data points

- E.g. work of Lewis Mervin, with AstraZeneca
- 2015, *J. Cheminformatics* (7) 51
- ChEMBL actives (~300k), PubChem inactives (~200m); 1,080 targets
- Can be retrained on in-house data
- https://github.com/lhm30/PIDGIN

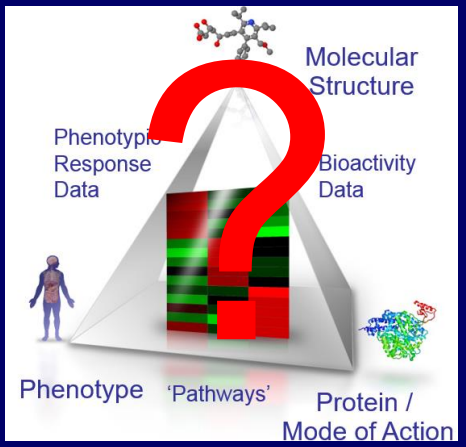
Molecule	Targets	Scores
Chiral 	PRKCB1	95.81
	CAMK2G	87.48
	PRKCG	66.35
	PRKCA	56.99
	PRKCD	52.44
	PRKCH	51.41
	PRKCE	50.42
	PRKCZ	42.48

Molecule	Targets	Scores
	ABL1	46.50
	PDGFRB	28.99
	KIT	22.02
	CDK9	21.30
	BRAF	16.13
	FLT1	13.09
	PLK1	8.05
	BTK	5.44



Also data publicly available

BUT…The world is not flat. What now?



- Links between drugs/targets/diseases are quantitative, incompletely characterized
- Subtle differences in eg compound effects (partial vs full agonists, off-targets, residence times, biased signalling, etc.)
- 'Pathways' from very heterogenous underlying information; dynamic elements not captured etc.
- Effects are state-dependent (variation between individuals, age, sex, co-medication…) – PK is often rather neglected in AI approaches
- Endotyping is not sufficient – *how do we characterize disease/phenotypes*?
- We don't understand biology ('the system'), we don't know what we *should* label, and measure, hence …
- We label what we *can* measure: 'T*echnology push*' vs '*science pull*' (!)
- **Are our labels – 'drug treats disease X', 'ligand is active against target Y', … - meaningful?**
- **Conditionality: Causality, confidence, quantification, ….?**
- **Computer science is tremendously powerful… but is our data?**

Example of labelling problems: adverse reactions

- ***"Does* drug Y cause adverse reaction Z? Yes, or no?"**
- Pharmacovigilance Department: Yes, *if* we have…
  - A patient with this *genotype* (which is generally unknown)
  - Who has this *disease endotype* (which is often insufficiently defined)
  - Who takes *dose X* of *drug Y* (but sometimes also forgets to take it)
  - With known targets 1...n, but also unknown targets (n+1…z)
  - Then we see *adverse reaction (effect) Z* …
  - But only in x*% of all cases* and
  - With *different severity* and
  - *Mostly if co-administered with a drug from class C*, and then
  - More frequently in *males* and
  - Only *long-term*
  - (Etc.)
- **So – does drug Y cause adverse event Z?**

Data/'AI' in early discovery vs efficacy/safety

Early discovery/proxy space (usually *in vitro*)

- Often 'simple' readouts (eg protein activity), hence…
- Large number of data points for training models

- *Models have clear labels* (within limits of model system, eg 'ligand is active against protein at IC50<10uM', or solubilities, logP, or the like)
- Good for model generation: *Many*, *clearly categorized* data points

Efficacy/safety (usually *in vivo*)

- Quantitative data (dose, exposure, …)

- More complex models (to generate data), *fuzzy labels* (classes 'depend', on exposure, multiple eg histopathological endpoints) – hence…

- *Less, and less clearly labelled data*: Difficult from machine learning angle

- Data: Difficult to generate, eg animal data tricky, even within single company (confounding factors abound)

Problem setting in early discovery vs safety

Early discovery/proxy space

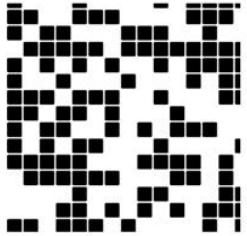
- Discovery setting – ‘find me suitable 100s or 1000s out of a million’ (eg screening)
- Anything fulfilling (limited) set of criteria will do ‘for now’, predicting *presence of something*
- Computationally *generative* models often fine

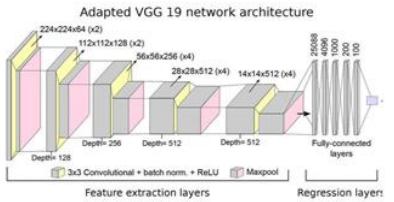
Efficacy/safety

- Need to predict for *this particular data point, quantitatively!*
- *Long list of criteria to rule out*, *based on limited data*… predicting *absence of ‘everything’* (eg different modes of toxicity)
- *Predictive* models (more tricky than generative!)

Object | Representation | Model | Object Label

Image Domain





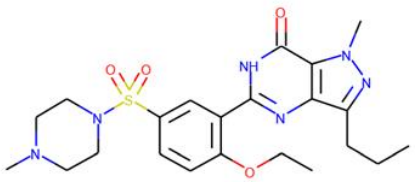
ResNet?
AlexNet?
CapsuleNet?

→ Cat

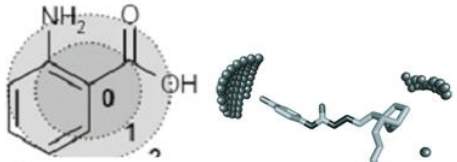
Largely *Unconditional* labels

Representation and model are *intrinsically linked* (ie, model uses native object representation by pixels)

Drug Discovery: Chemical Domain



?



logP = ...
Molecular Weight = ...
Molar Refractivity = ...

?

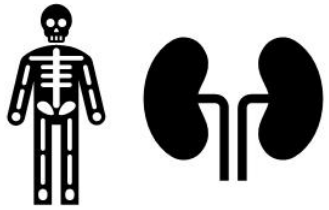
Artificial Neural Network/DNN? Support Vector Machine? Random Forest? Bayesian Classifier?...

→ Property A

Conditional labels (eg dependent on assay system, genotype, dose, endotype, sex, age, comedications, lifestyle, ...) ?

Both representation and modelling approach are *largely trial and error* (and *not* intrinsic to the chemical domain)

Drug Discovery: Biological Domain



?

Transcript-/proteomics? High-content imaging? Epigenetics? Histopathology? ....

?

Artificial Neural Network? Support Vector Machine? Random Forest?

→ State/Effect B

Heavily conditional labels (eg dependent on genotype, dose, endotype, sex, age, comedications, lifestyle, ...) ?

Both representation and modelling approach are *largely trial and error* (in particular the information content of biological readouts has only been established for particular cases)

Bender & Cortes, Drug Discovery Today, 2021

Much of the data we have has been generated with proxy assays. Why is this a problem for AI in drug discovery?

- There is *what we are really interested in* - say, mitochondrial safety, Drug-Induced Liver Injury (DILI), …
- And there is what we *measure as an assay endpoint* – say, cytotoxicity in a Glu/Gal (differential cytotoxicity) assay to *approximate* mitochondrial safety; Bile Salt Export Pump (BSEP) inhibition to *approximate* DILI, …

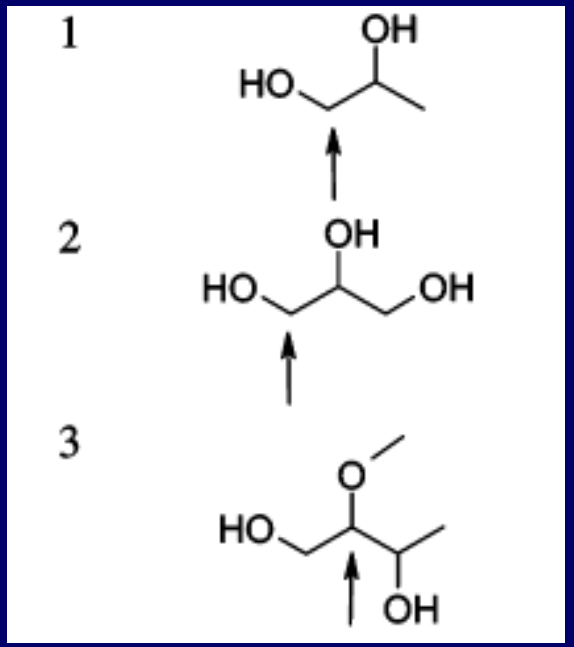
- Take-away: ‘Proxy’ assays measure only part of reality, in a particular assay, with particular conditions
- Not to be confused with property itself (!)
- Problem: Proxy endpoint (a) taken as ‘ground truth’ in AI in drug discovery, (b) embedding into project context neglected

Key problem in chemical datasets: Biases! Influences all explainable AI approaches (!)

- Chemical space is 10^{63} - however, our data (large is 10^{6} compounds) clusters tremendously
  - Drugs? Fast followers, analogues
  - Published literature? Series (for SAR)
  - *Etc*

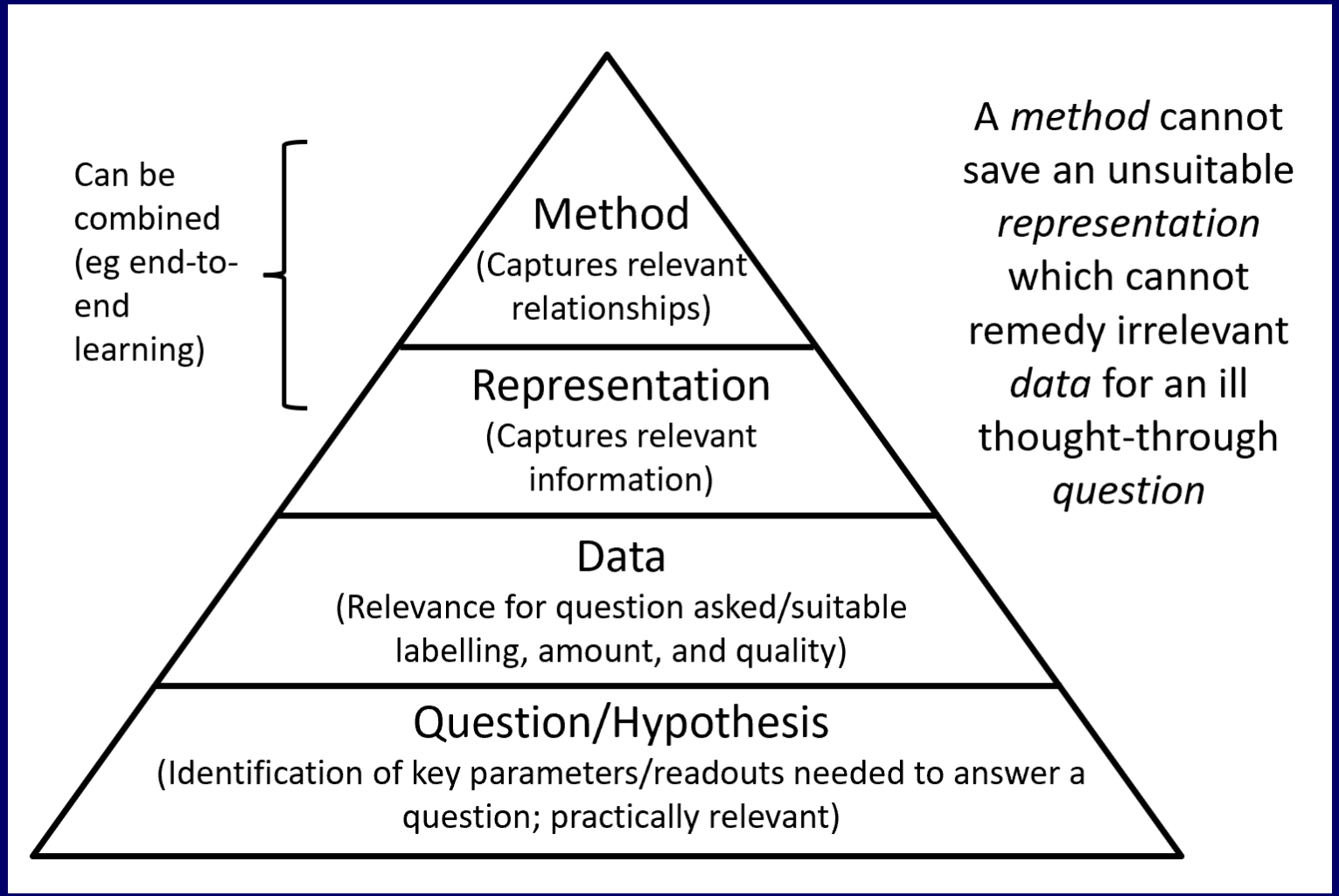
- Example (from own work): 649 bitter compounds vs 13k compounds from MDL Drug Data Repository
- Characteristic features for bitter compounds?

  Sugar rings! (due to glycosylation of natural products, which are often bitter; shown are fingerprint features which capture parts of those rings)



Rodgers, *J. Chem. Inf. Model.* 2006, 46, 569.

The *question* needs to come first… and then the data, then the representation, and then the modelling method! http://www.DrugDiscovery.NET/HowToLie



Lots of attention currently here…

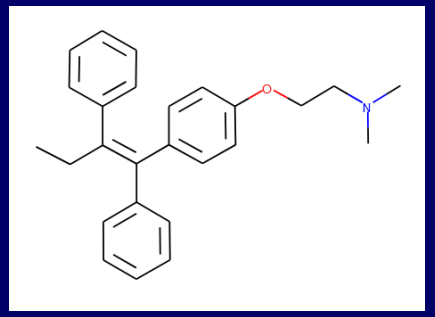
But we need to care more about this

On data, endpoints, models, and predictions

- Preamble: Will ‘AI’ and computational models eat the world? Some historical context
- What we need: *in vivo* relevance
- Data (labelling, conditionality, *in vivo* relevance)
- Models and Validation
  - Descriptors, machine learning, validation
  - “Questions to ask your friend, the modeller”
- Readouts and Applications
  - Cell Painting Assay
  - Applications to mitochondrial toxicity
  - Interpreting high-dimensional cell morphology readouts

What is a computational model?

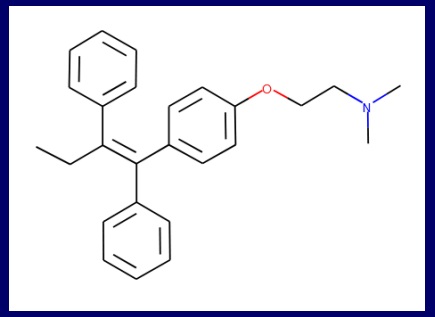
We have (from experiments): Molecule -> Endpoint

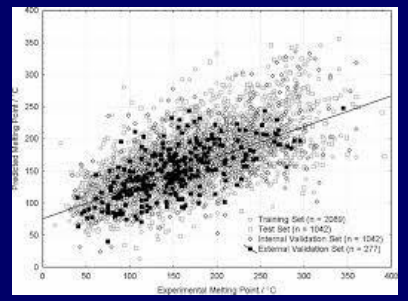


Measured (*condition: experiment*)

IC50= ..nM

We model: Molecule -> Descriptor -> Model -> Endpoint

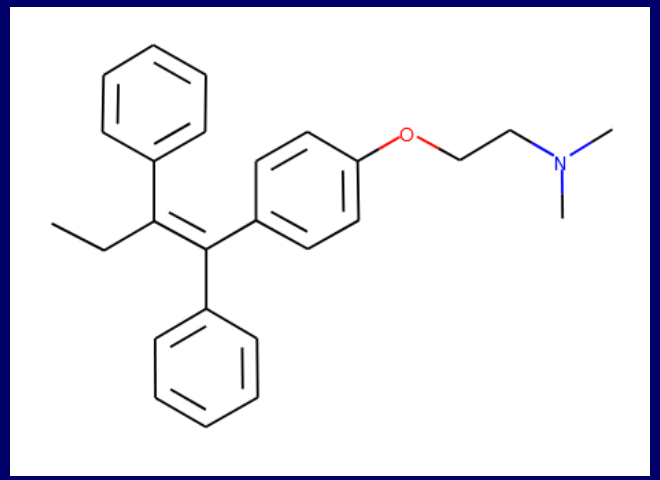


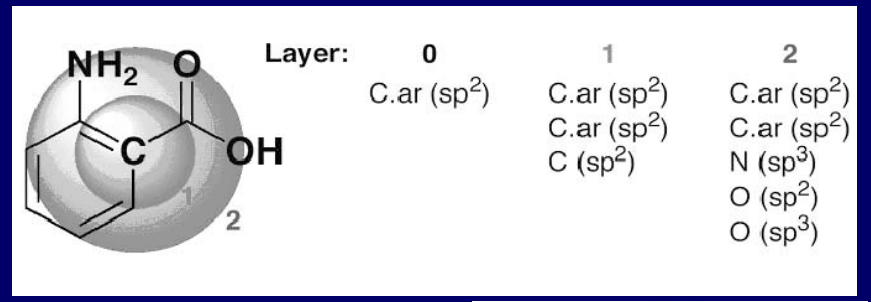


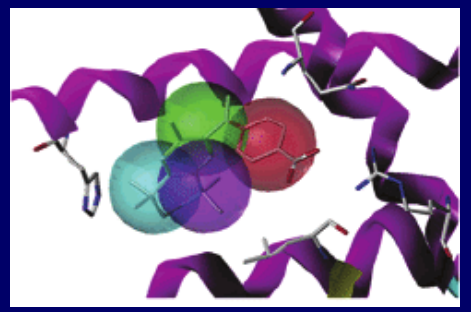
IC50= ..nM

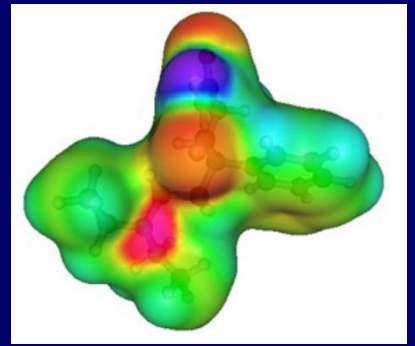
Descriptors

- Provide an *information-preserving* representation of input data (e.g. structures) for the model
- Either knowledge-based (e.g. reactive groups), or (usually) ‘trial and error’
- Can also be biological readouts (gene expression, cell morphology)
- *Can* be learned from data, but only *if there is enough data, and we can meaningfully label!*







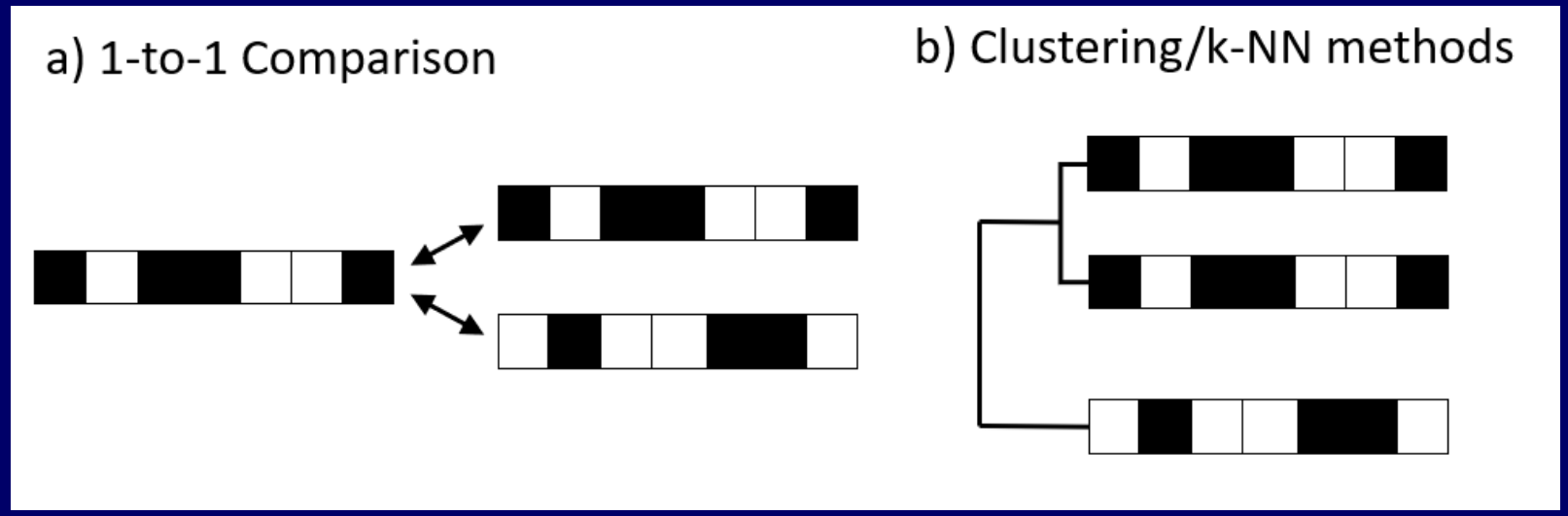


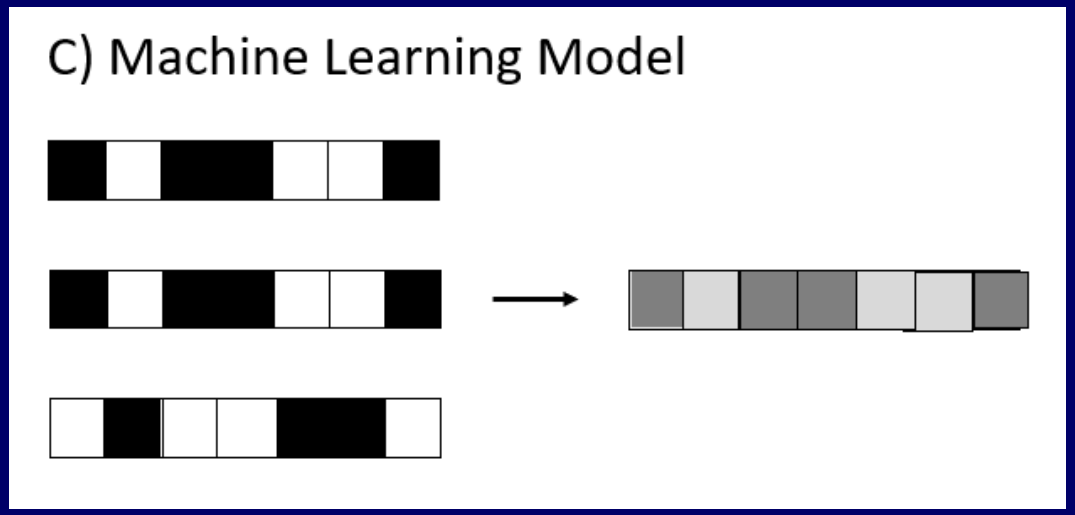
0100101010000…

Fingerprints, pharmacophores, surface properties, substructures/ functional groups, shapes, physchem properties *etc.*

Types of models (all of which can involve feature selection)

- Similarity-based (single neighbour, 1-NN)
- Clustering-based (multiple neighbour, k-NN)
- Machine learning models



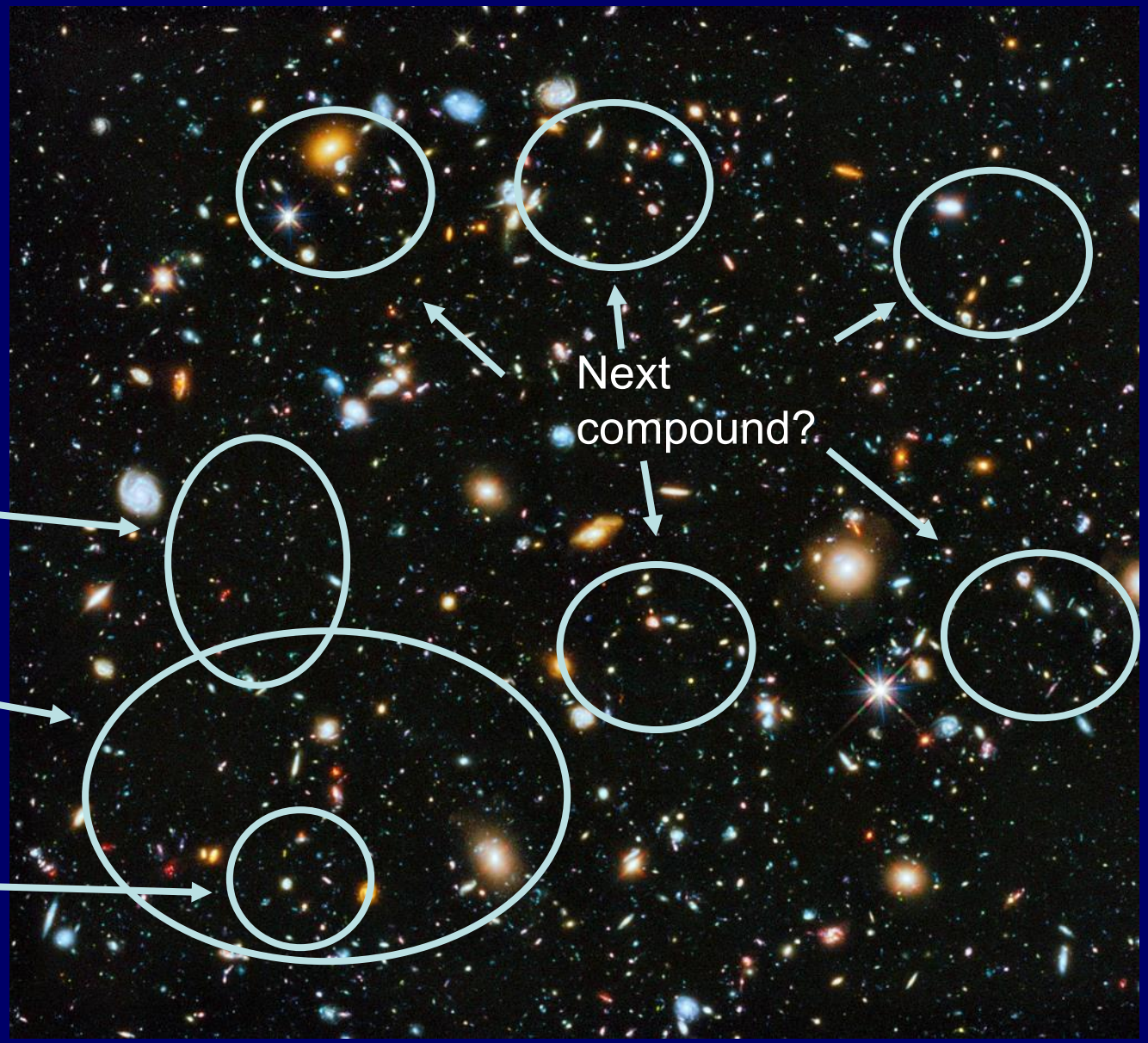


How do we know that something *works*? What is ‘validation’?

- Core question in science, core question for start-ups
- In *theory* we establish a method, use a benchmark, and know how well the method works
- In *practice* this doesn’t really work with *in vivo* data –
  - Labels are either mostly only *in vitro*-relevant, or conditional (‘depend’ on dose, *etc*)
  - Validation is costly (e.g. phase II studies for efficacy; *plus controls*), *little prospective data*
  - **Difficult to sample distribution in chemistry/’project’ space well (diversity, number), so performance *depends heavily on test set***

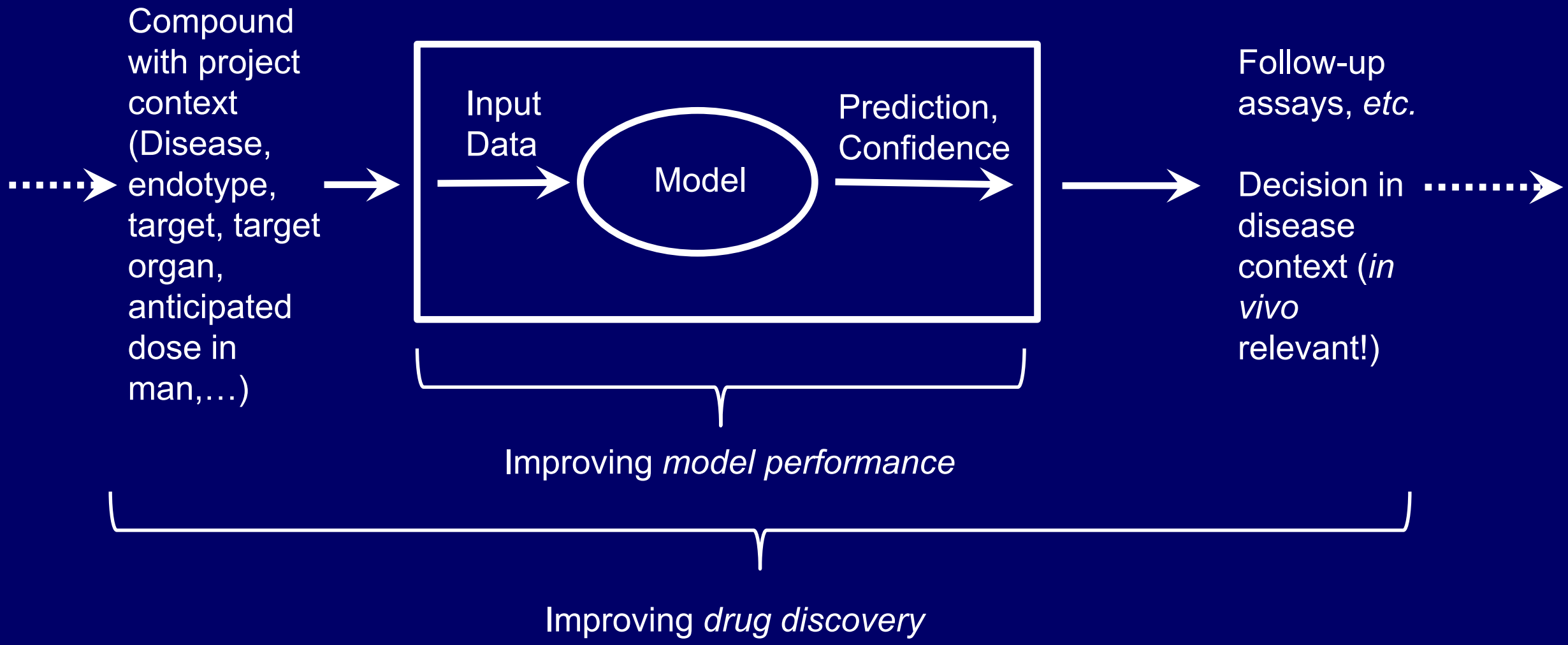
- Retrospective validation is all we can do (but no prospective discovery, predictivity for future projects unknown, all behave differently)

Why ‘validation’ of a model is tricky: You get the numbers you want (depending on the question you ask/data set you use!)

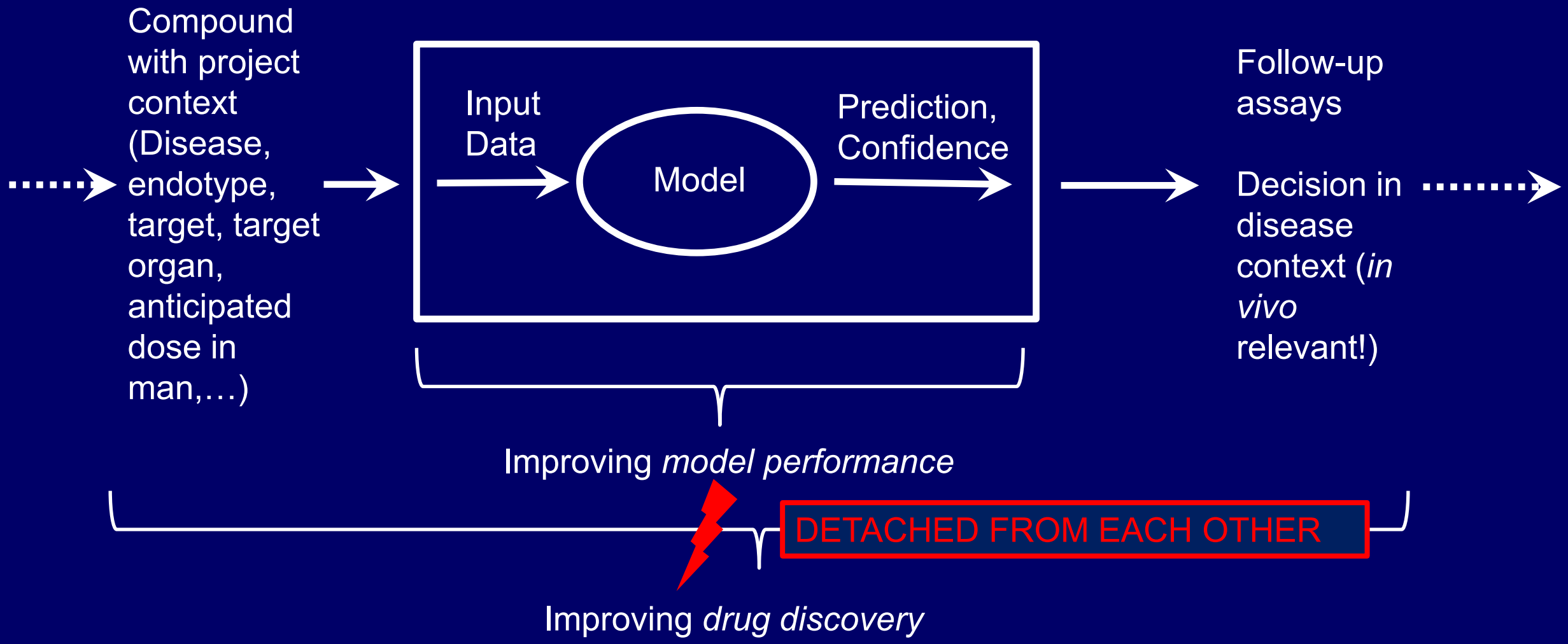


- Chemical space is large; data sets are small
- Model is unable to generalize to unseen spaces
- Effect of changes is conditional on scaffold/context
- Sampling of data is generally insufficient
- *“Every model is a local model”*

Model validation vs process validation (e.g. compound structure-based property predictions)



Using computational models for decision making often disappoints since (a) model validation is decoupled from process validation, and (b) many (most!) models use only proxy data ('model of models')

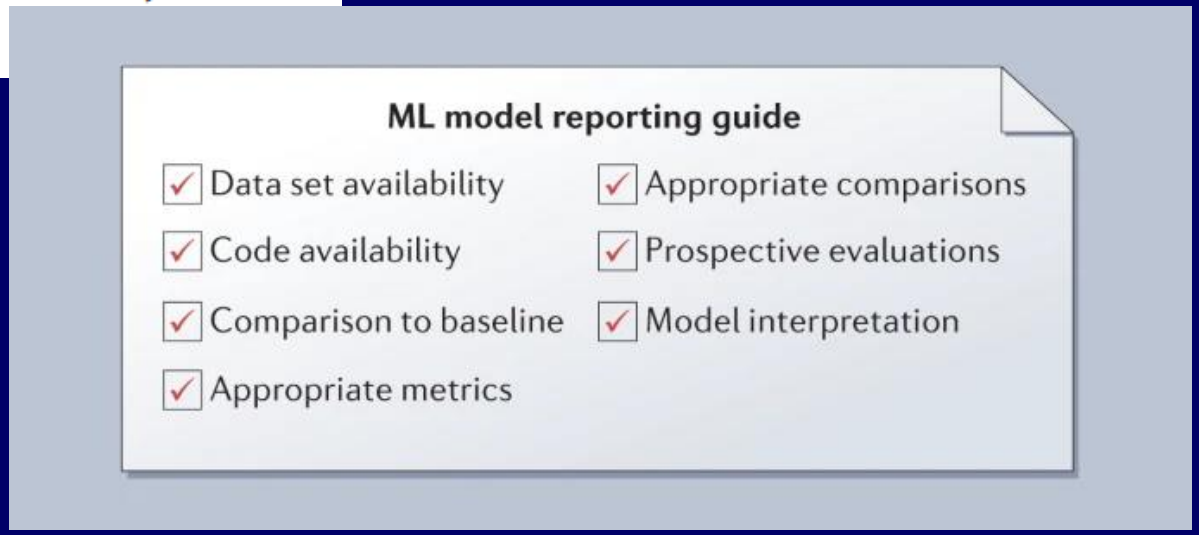


Model validation – two resources

1. http://www.drugdiscovery.net/HowToLie
2. Nature Reviews Chemistry 2022 article

Evaluation guidelines for machine learning tools in the chemical sciences

Andreas Bender, Nadine Schneider, Marwin Segler, W. Patrick Walters, Ola Engkvist and Tiago Rodrigues



Questions to ask your friend, the modeler (1/2)

- Key goal: How good is the prediction for *my new compound?*
- Data
  - What is the *number* of data points in the model, and is chemical space coverage *relevant* for my application?
  - *Performance is a function of space! Less space… gives a (numerically) better model! Performance/applicability domain is a trade-off!*
  - What is the closest neighbour (according to mechanistically interpretable space; model space; similarity space), and *is it relevant, given the particular question being asked*?
- Descriptors
  - How was the descriptor chosen, and is there a mechanistic rationale for its choice? (depends on understanding of system; e.g. reactive substructures, bioactivity-based, generic similarity, …)

Questions to ask your friend, the modeler (2/2)

- Models
  - Was there an external test set used in model validation (and was it large, diverse, *relevant to new compound predictions*)?
  - Does model performance change, depending on parameter choices (indicates model instability), and training/test set splits (indicates overfitting)?
  - Is there an applicability domain/confidence that the model assigns – and *does it actually work on the external test set (rather often it does not!)*?
- If all of this is answered satisfactorily, then (a) data in the model covers my new molecule, with (b) a suitable descriptor, and provides (c) a confidence with the prediction

On data, endpoints, models, and predictions

- Preamble: Will ‘AI’ and computational models eat the world? Some historical context
- What we need: *in vivo* relevance
- Data (labelling, conditionality, *in vivo* relevance)
- Models and Validation
  - Descriptors, machine learning, validation
  - “Questions to ask your friend, the modeller”
- Readouts and Applications
  - Cell Painting Assay
  - Applications to mitochondrial toxicity
  - Interpreting high-dimensional cell morphology readouts

Applications

- Cell Painting – What is it?
- Predicting mitochondrial toxicity
- Merging chemical structural and cell painting information
- Predicting mitochondrial toxicity of PROTACs
- Representing and understanding high-dimensional feature spaces

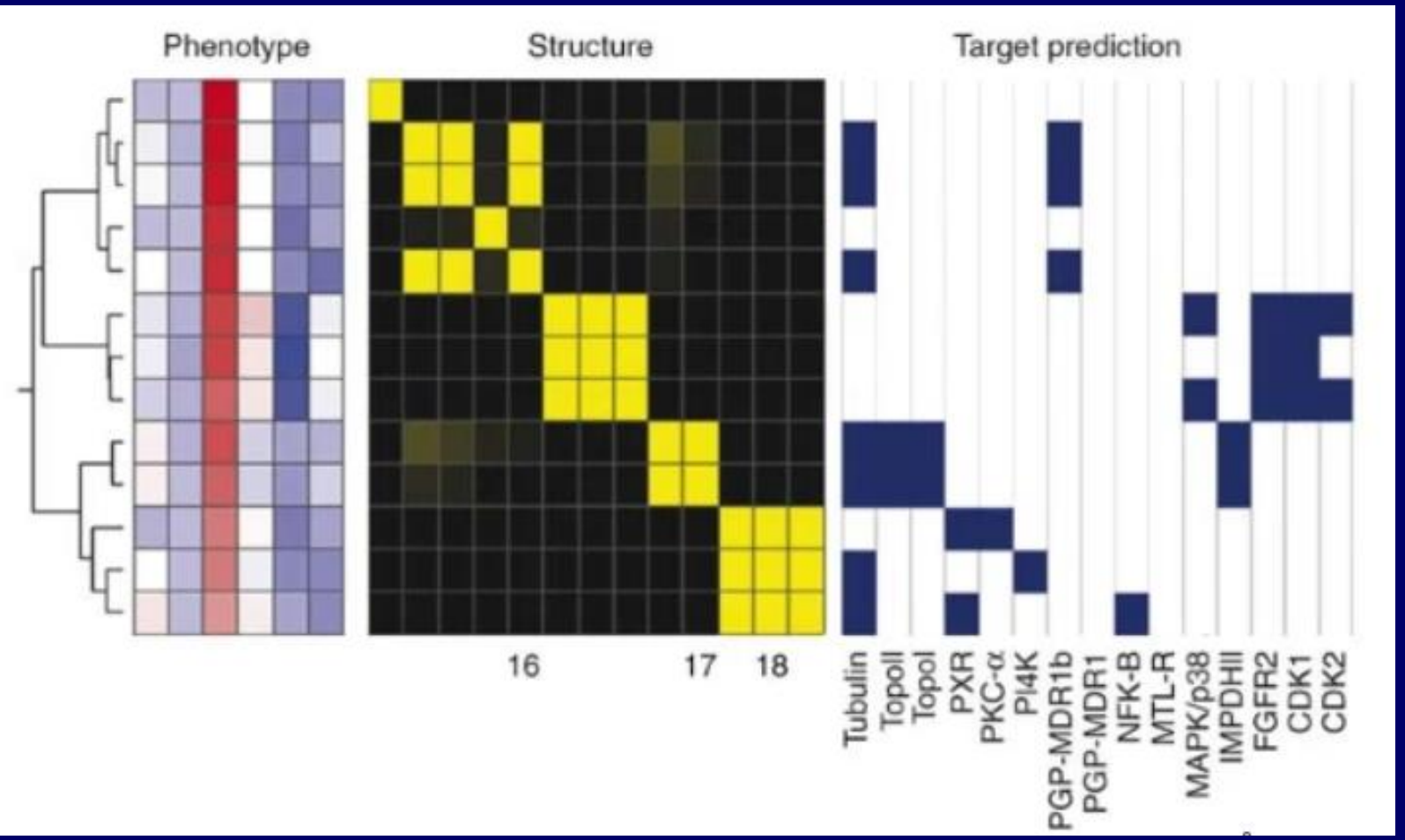
Problem

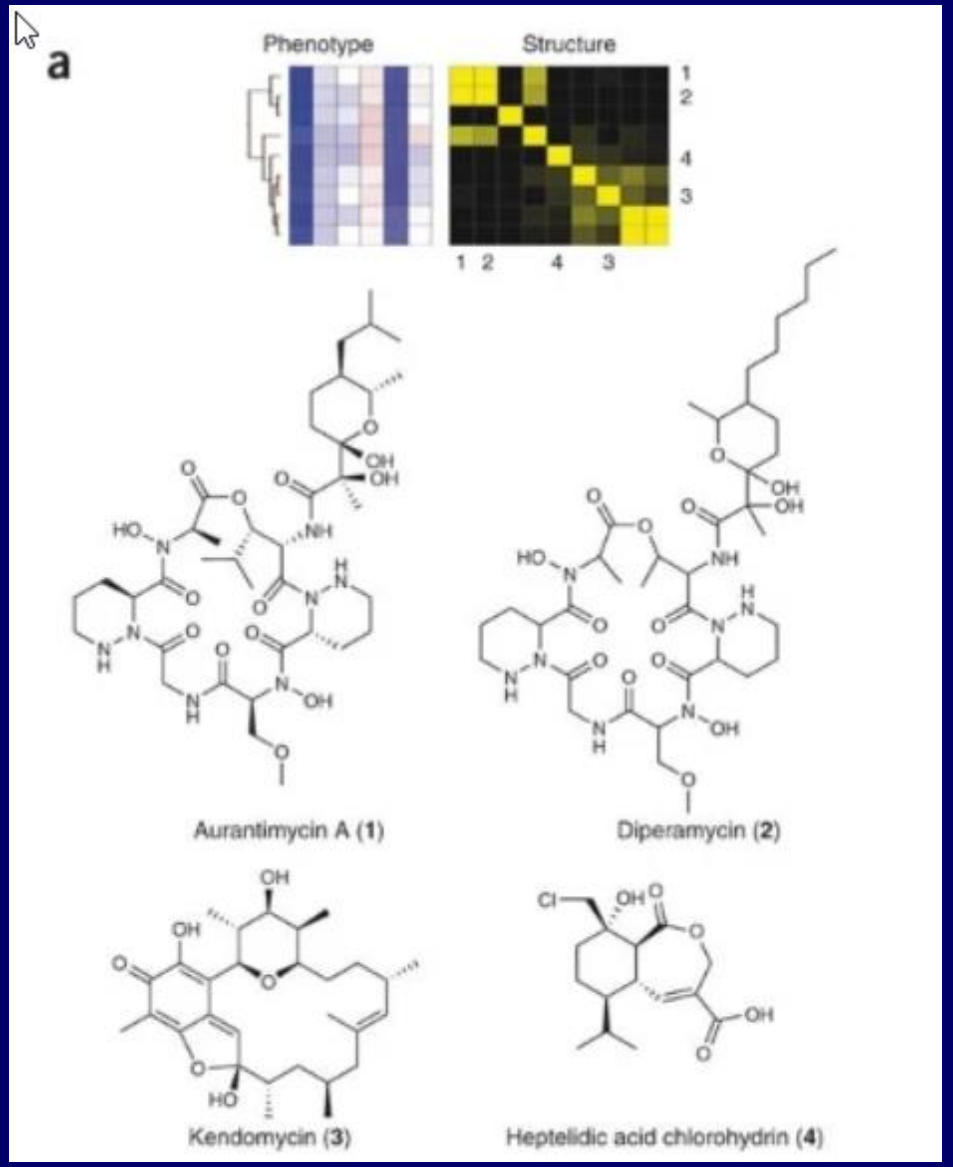
- In many (most?) cases we don’t understand how something works (i.e., biology)

- If we understand how something works we can do *hypothesis-driven, science-pull* driven data generation

- If we *don’t* understand how something works we need to revert to *hypothesis-free, technology-push* driven data generation and describe *variance*
- In this case we need *independent* pieces of information, and we need to *retro-fit* to what is *relevant*

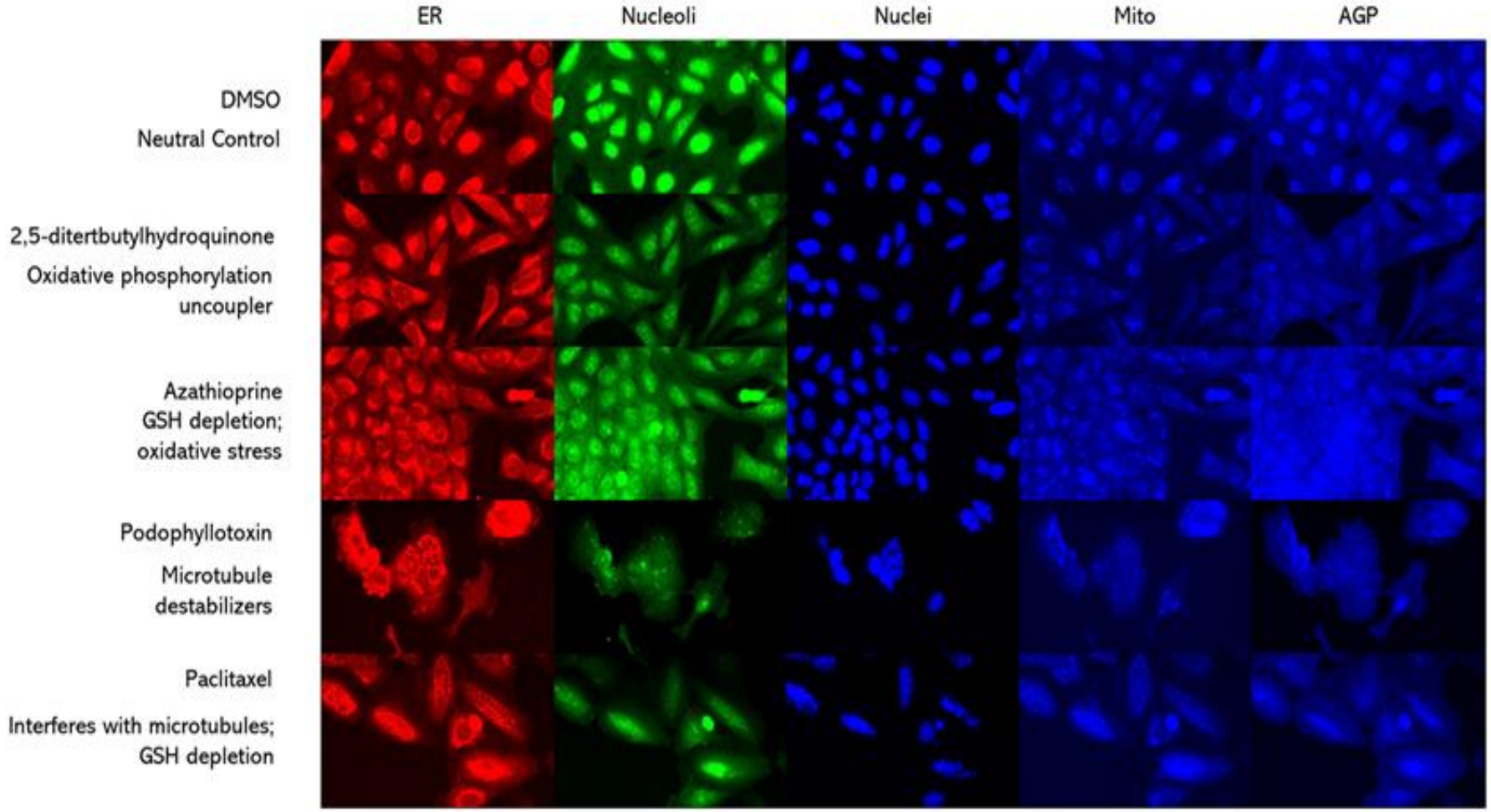
Why –omics, why cell morphology, … if we have the structure?





D. W. Young et al., Integrating high-content screening and ligand-target prediction to identify mechanism of action, Nature Chem. Biol. 2008

Cell Painting cell morphology assays:
Six stains/five channels/eight compartments

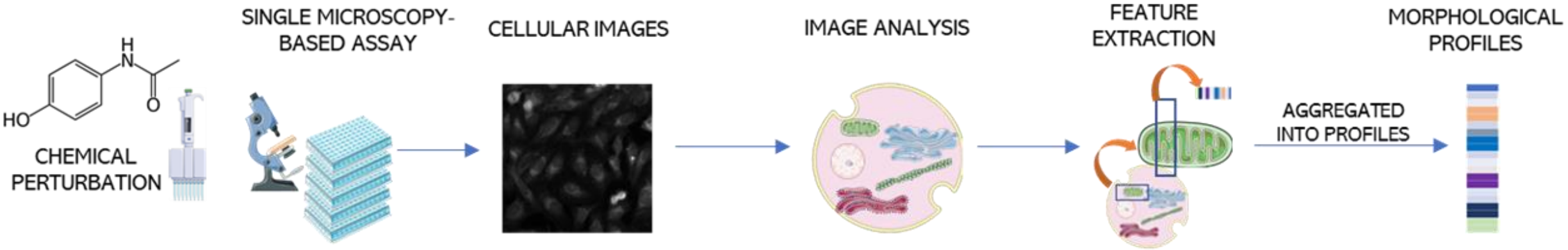


Anika Liu and Srijit Seal et al

Features of the Cell Painting Assay – Form basis (input variables) of machine learning model

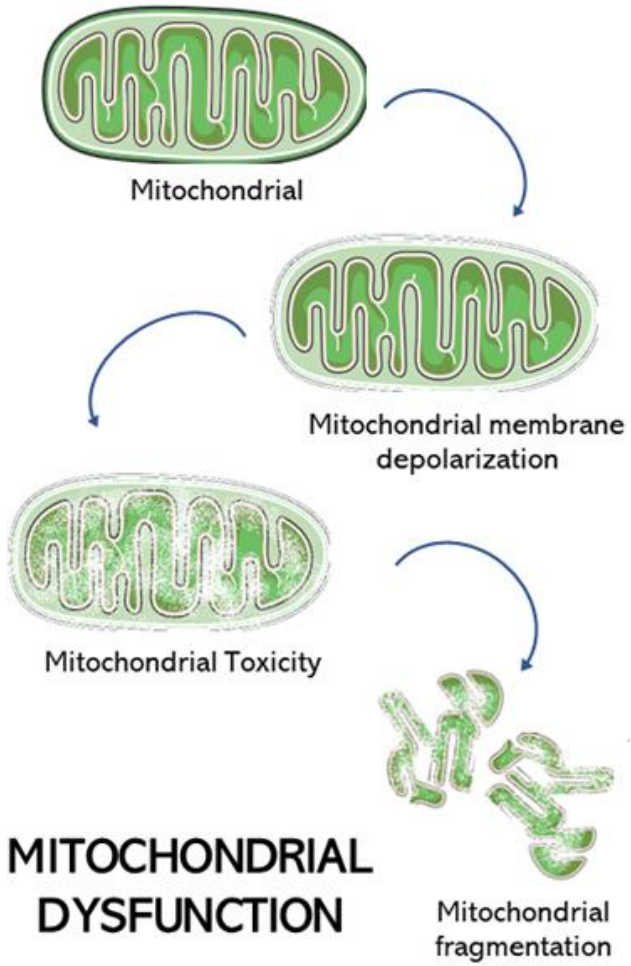
For each identified compartment, measurements include:

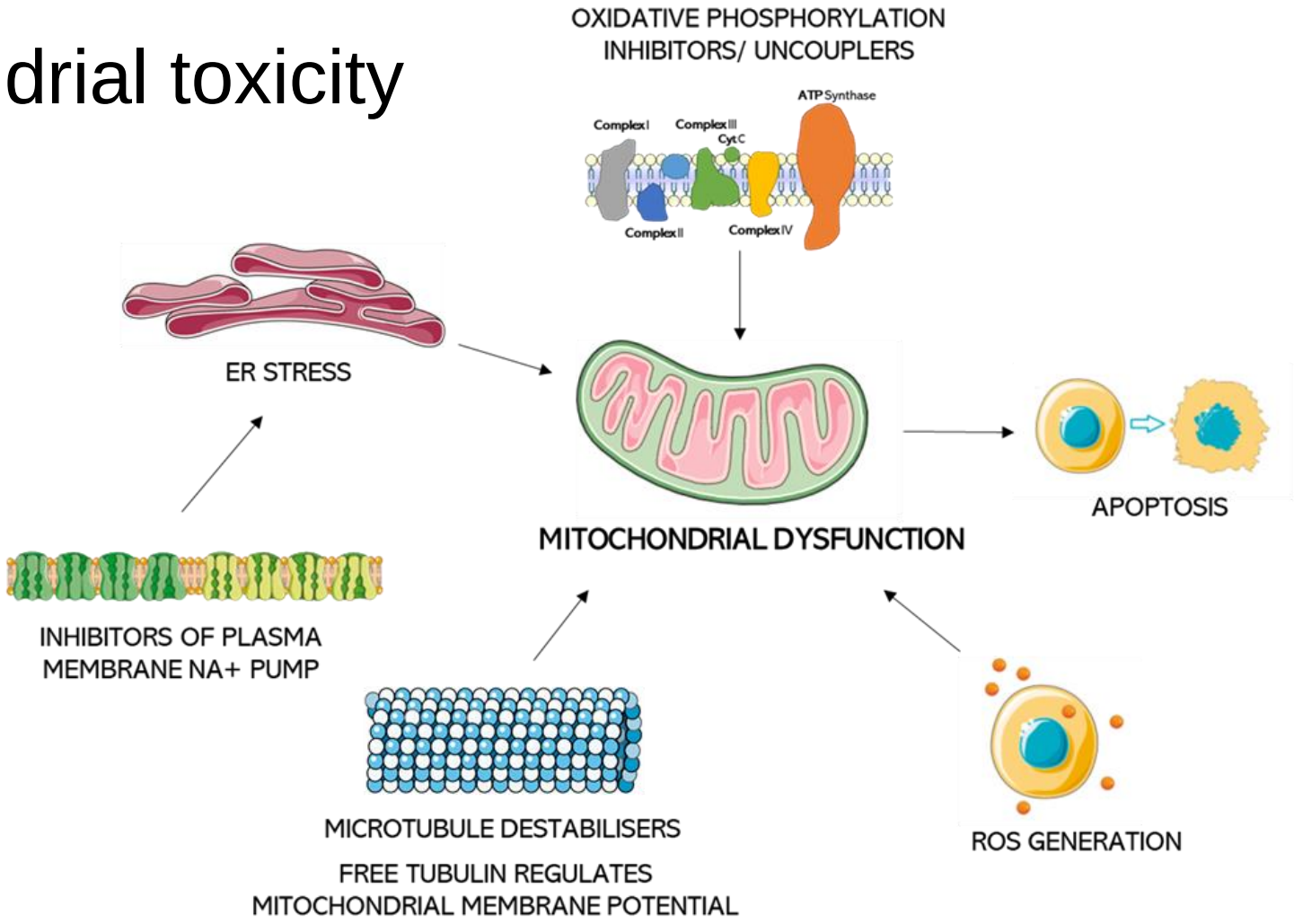
- Counts
- Size: area, volume, perimeter, diameter
- Shape
- Texture (smoothness)
- Intensity
- Spatial relationships between features



Bray, M. A.; Gustafsdottir, S. M et al., A Dataset of Images and Morphological Profiles of 30 000 Small-Molecule Treatments Using the Cell Painting Assay. GigaScience. Oxford University Press December 1, 2017, 1–5.

Mitochondrial toxicity





Toxicants act on multiple pathways to exhibit mitochondrial toxicity, mostly inhibition of mitochondrial respiratory chain or uncoupling of oxidative phosphorylation.

Dataset

Training Dataset:

- Tox21 Mitochondrial membrane potential disruption assay hit calls (summary assay)
- 382 compounds
- 62 Mitotoxic

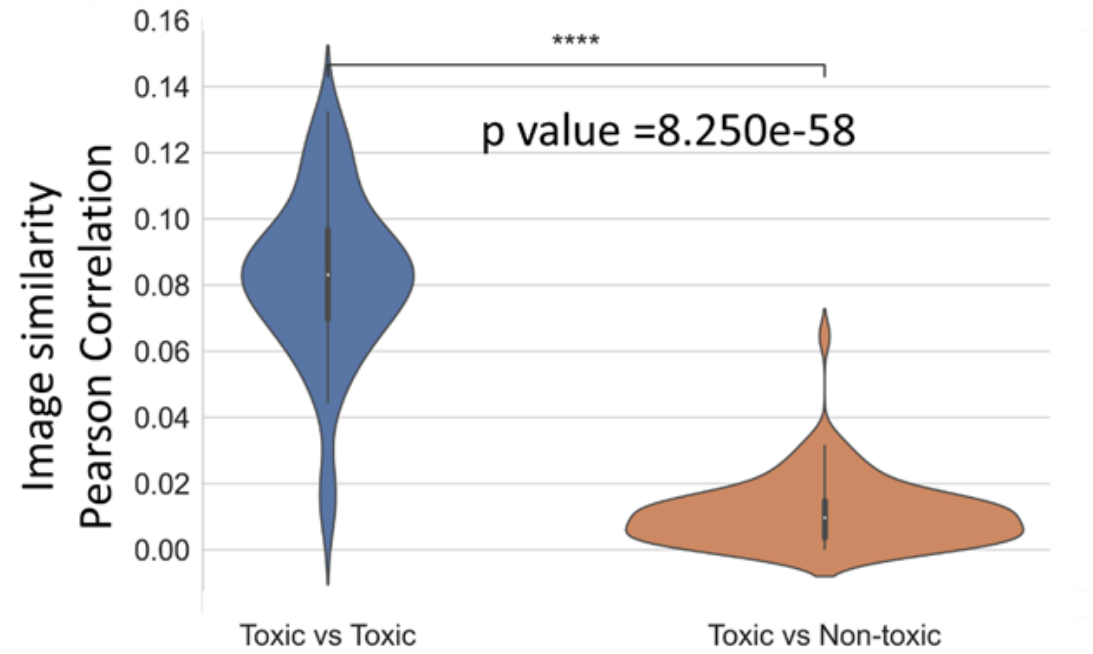
External Test:

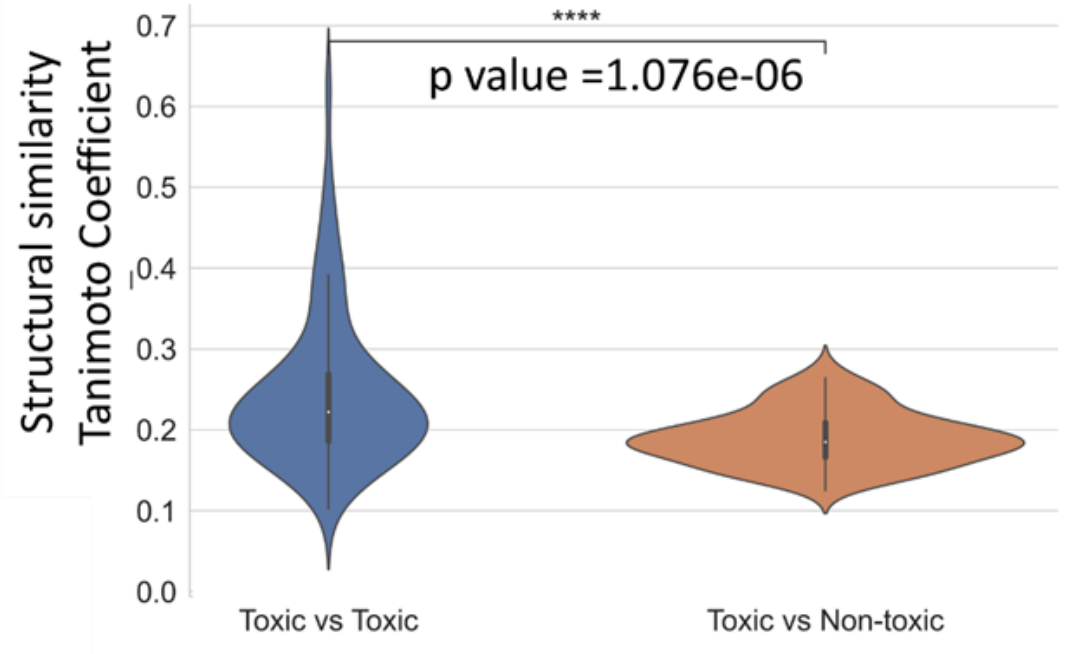
- Additional mitotox assays from CHEMBL, PubChem, Mitotox Database relevant to mitochondrial potential
- 244 compounds
- 47 Mitotoxic

Hemmerich, J., Troger, F., Füzi, B. & F.Ecker, G. Using Machine Learning Methods and Structural Alerts for Prediction of Mitochondrial Toxicity. Mol. Inform. 39, (2020)

Toxic compounds are more similar in morphology space than fingerprint space

Morphological space is more able to discriminate between mitochondrial toxicants and non-toxicants than structural fingerprints.

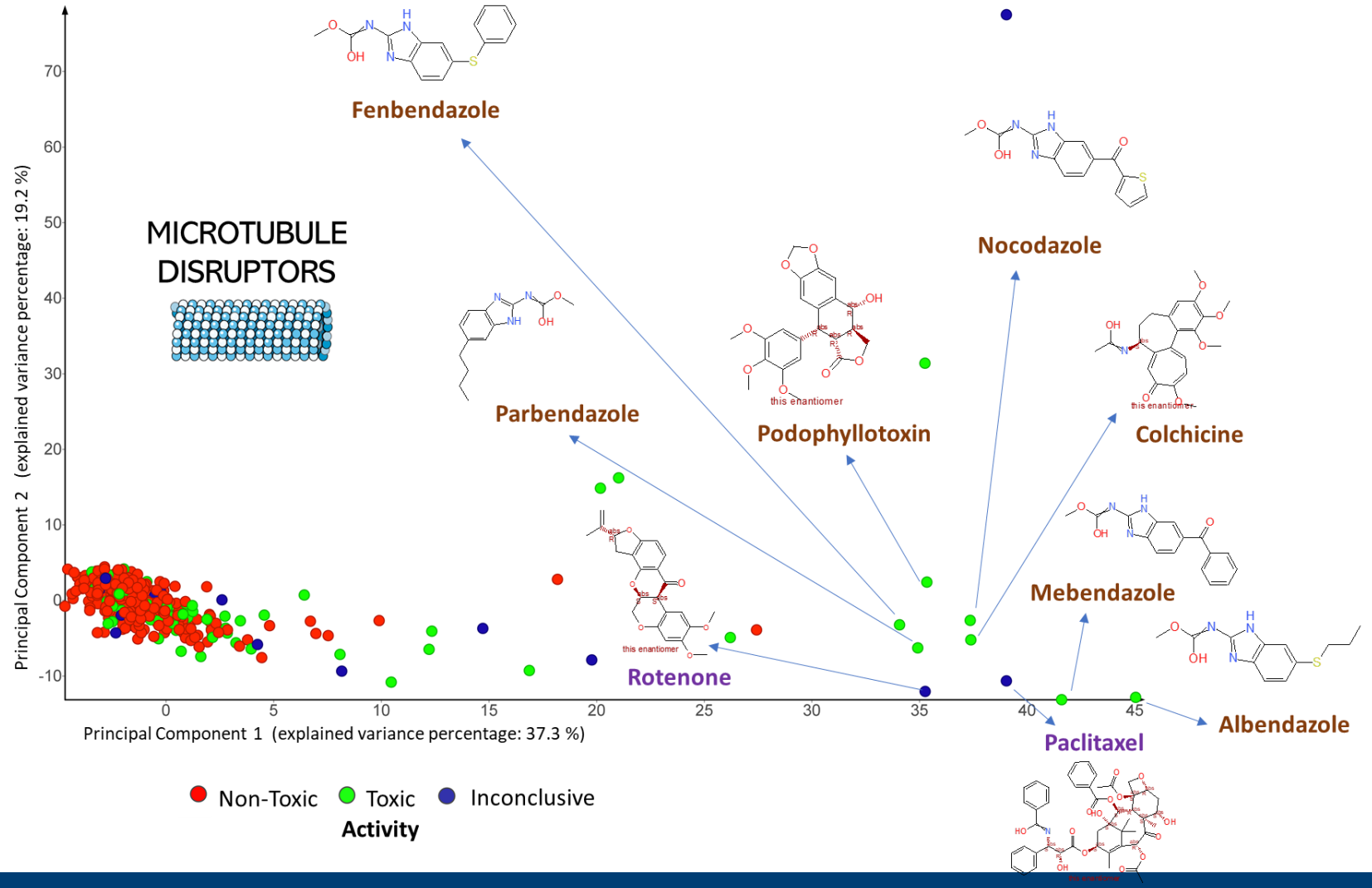




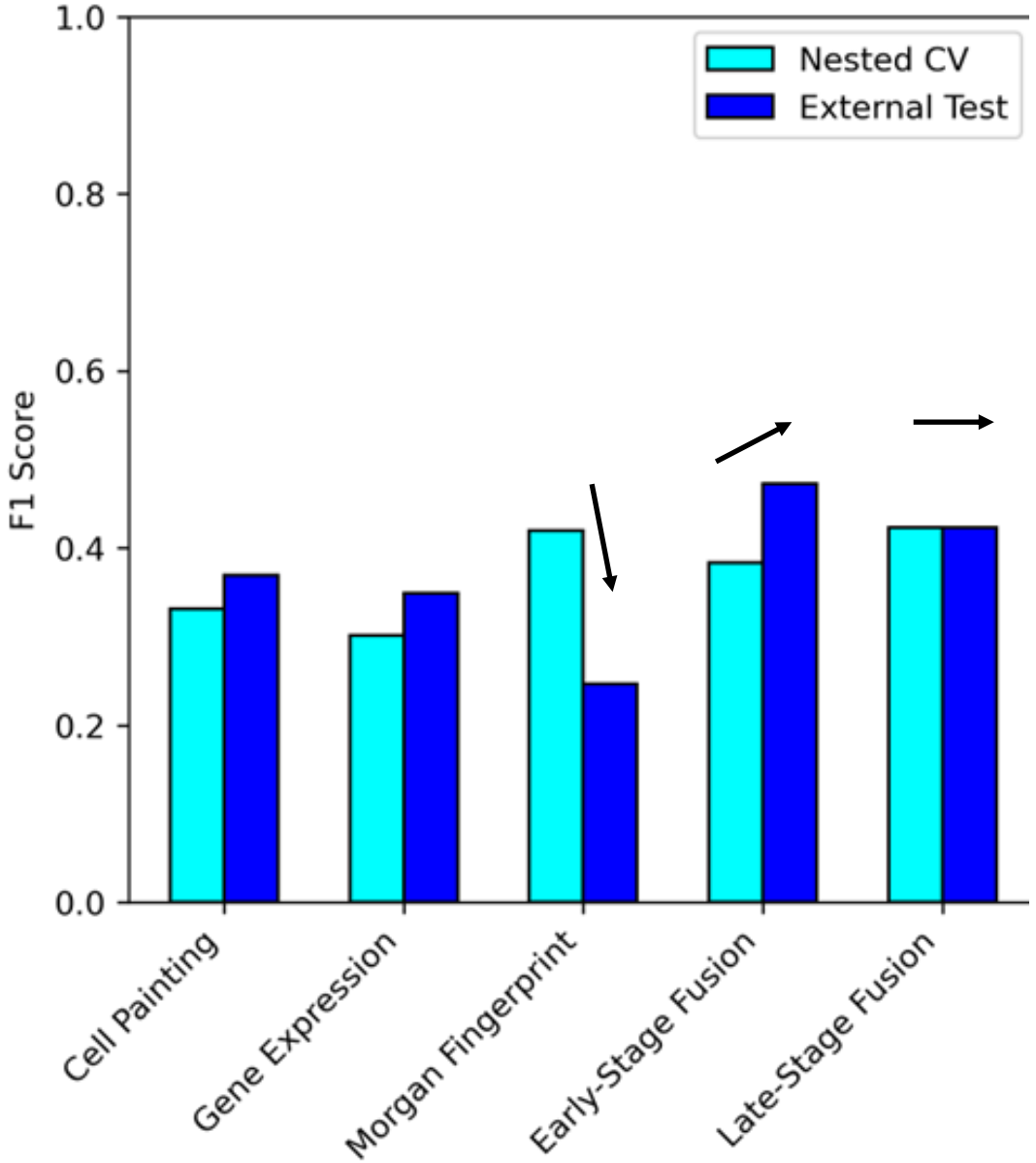
Intra- and inter-class pairwise similarity for 486 compounds (85 mitotoxic)

Morphology space clusters compounds with similar mechanisms

Compounds clustered further away from the distribution of majority of compounds having similar mechanisms of actions, for example, microtubule disruptors



Principal Component Analysis of 542 compounds in 110-dimensional Cell Painting feature space.



Fusion models perform better on external test set

- External test set: F1 Score increases by 60% (0.25 to 0.42 in absolute terms) when using fusion models compared to Morgan fingerprints.
- Our method achieve higher sensitivity (0.79 in our study vs 0.37 in Apredica MitoMass) with comparable balanced accuracies (0.69 in our study vs 0.65 in Apredica MitoMass).

Hallinger, D. R., Lindsay, H. B., Friedman, K. P., Suarez, D. A. & Simmons, S. O. Respirometric screening and characterization of mitochondrial toxicants within the toxcast phase i and II chemical libraries. Toxicol. Sci. 176, 175–192 (2020)

Cell Painting features related to mitotoxicity

Biological significance of Cell Painting features with respect to Mitochondrial Toxicity :

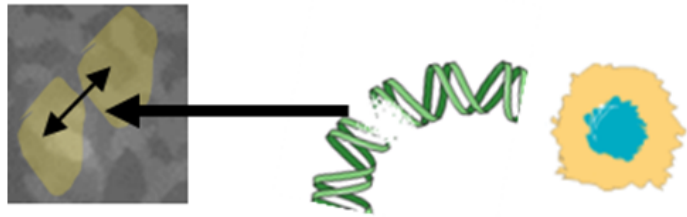
MITOCHONDRIAL FEATURES

Cells Intensity MaxIntensityEdge Mito (PPV 0.83)

Edge of segmented object potentially indicates loss of membrane integrity

FEATURES FROM OTHER IMAGE CHANNELS

Cells Correlation Costes DNA AGP (PPV 0.52)



Potentially indicates DNA fragmentation and entering apoptosis or cell death

Application to mitochondrial toxicity of PROTACs

ACS chemical biology

pubs.acs.org/acschemicalbiology

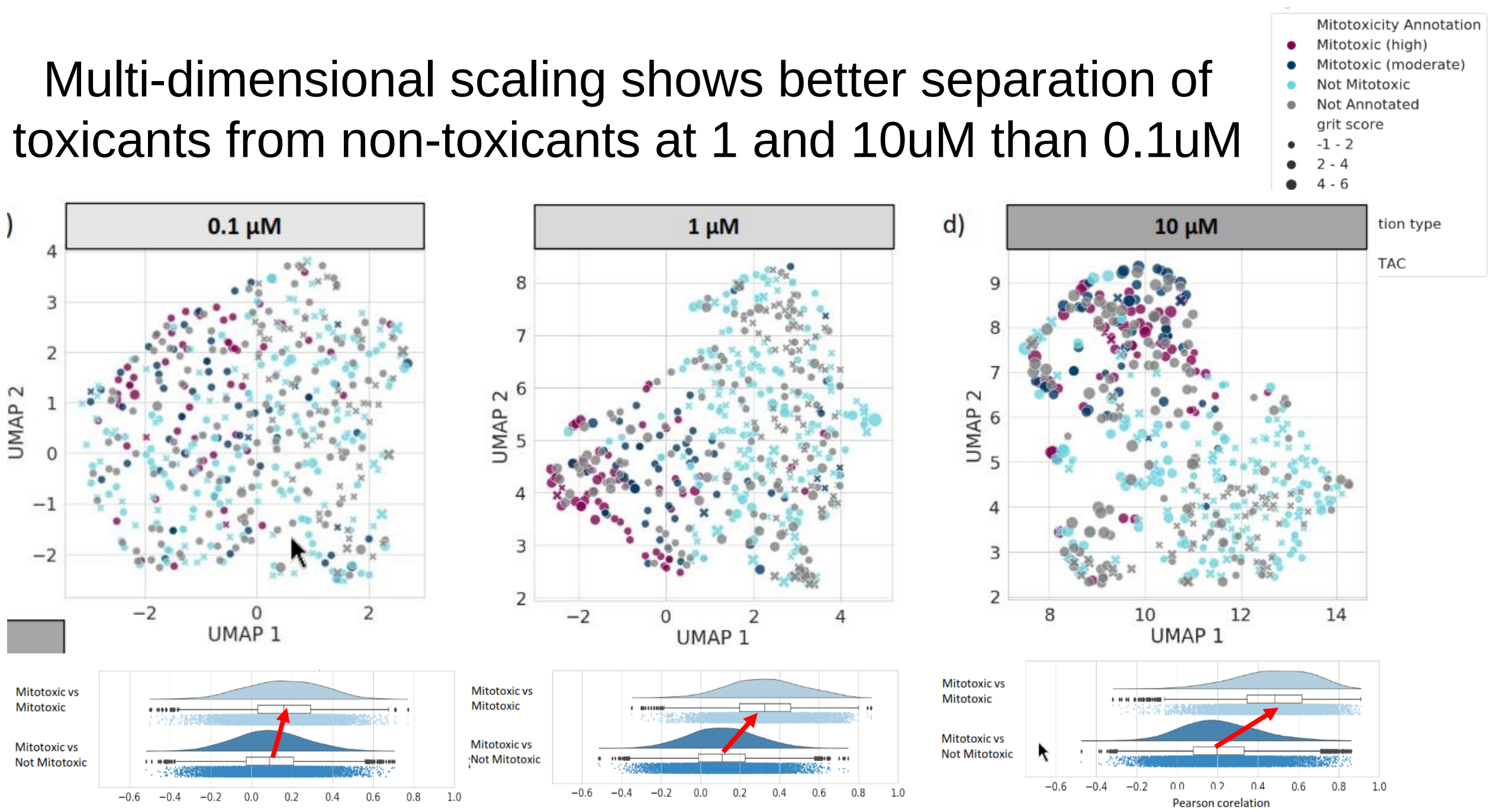
Articles

Cell Morphological Profiling Enables High-Throughput Screening for PROteolysis TArgeting Chimera (PROTAC) Phenotypic Signature

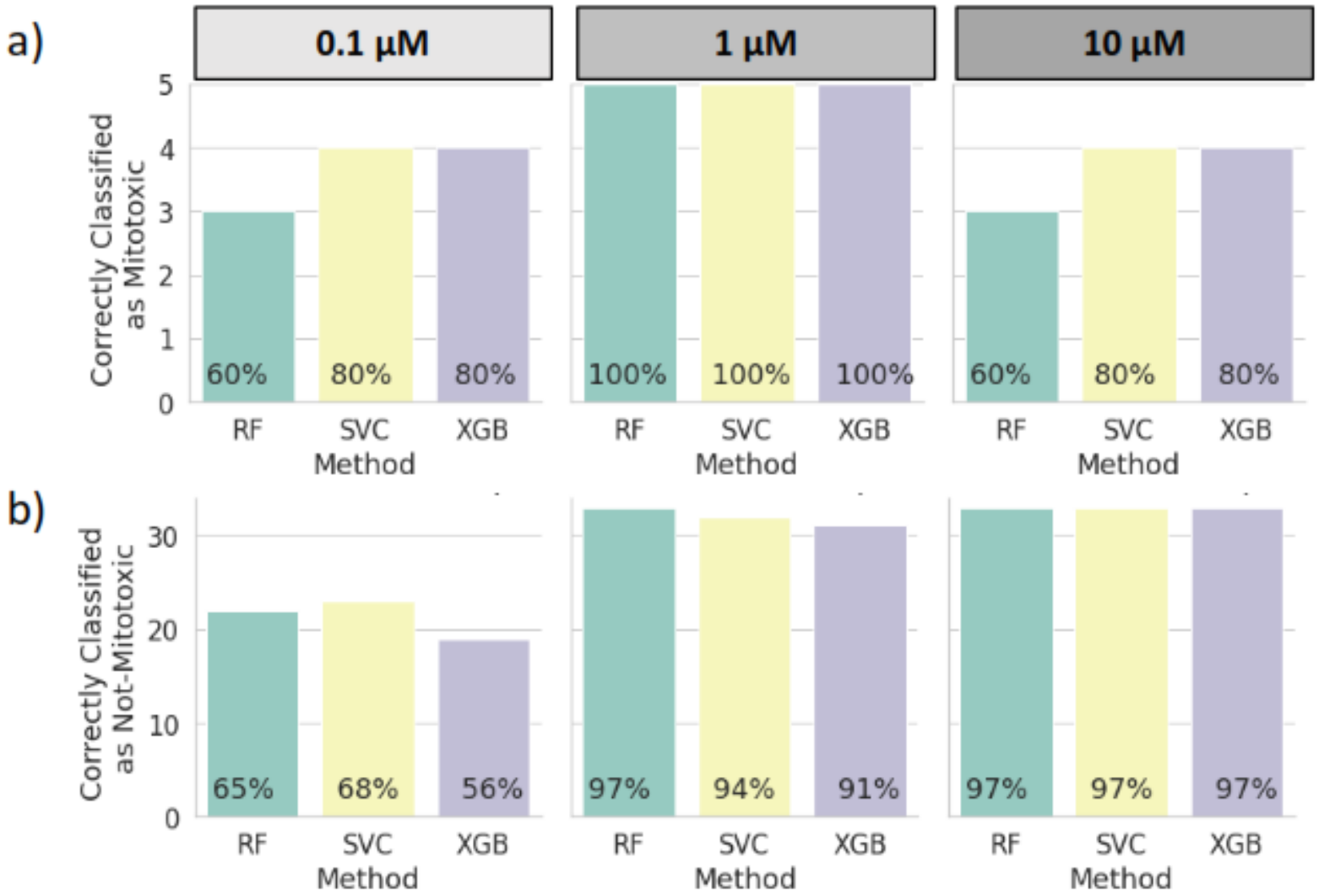
Maria-Anna Trapotsi, Elizabeth Mouchet, Guy Williams, Tiziana Monteverde, Karolina Juhani, Riku Turkki, Filip Miljković, Anton Martinsson, Lewis Mervin, Kenneth R. Pryde, Erik Müllers, Ian Barrett, Ola Engkvist, Andreas Bender, and Kevin Moreau*

- Work by Maria-Anna Trapotsi, Kevin Moreau, and others
- With AstraZeneca

Multi-dimensional scaling shows better separation of toxicants from non-toxicants at 1 and 10uM than 0.1uM

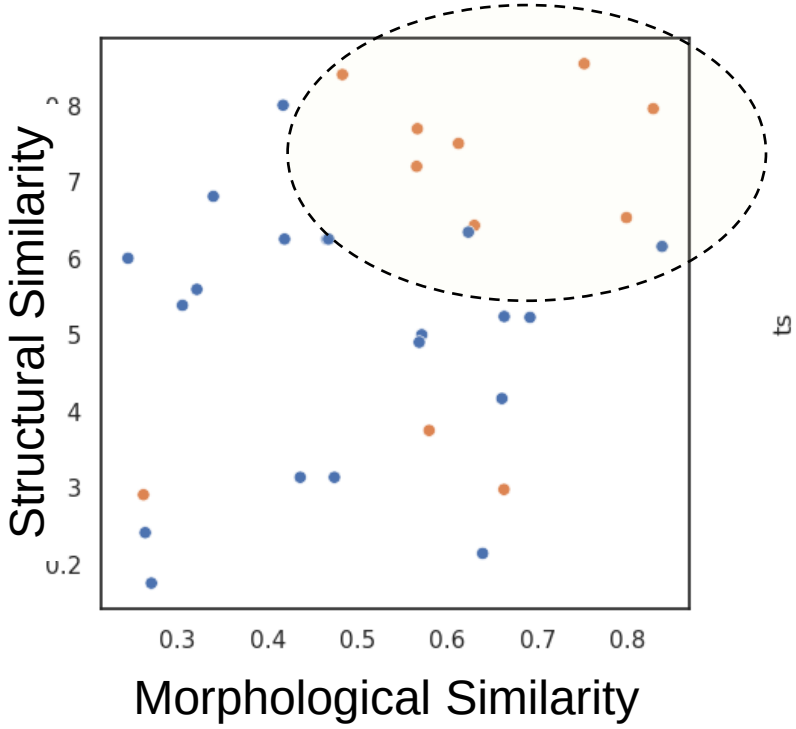


Prospective validation of mitotoxicants successful

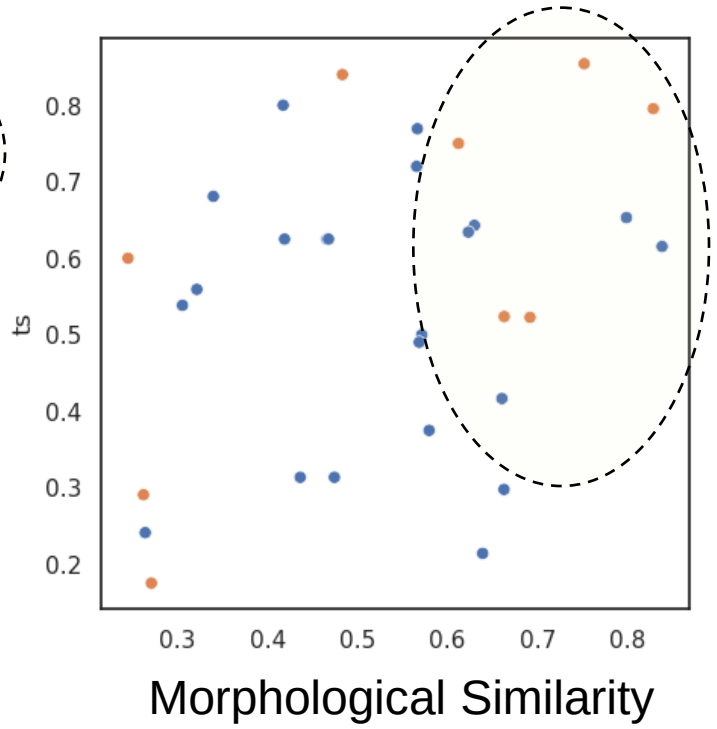


Merged Models can improve applicability domain (here for Tox21 endpoint, separate work)

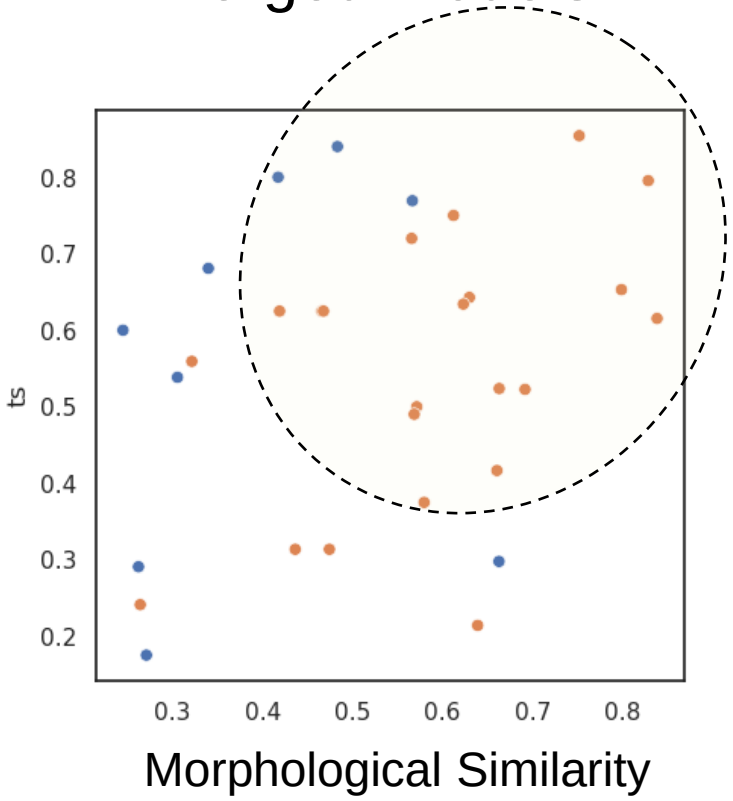
Morgan Fingerprint



Cell Painting



Merged Models



But how to interpret Cell Painting space, which is highly correlated?

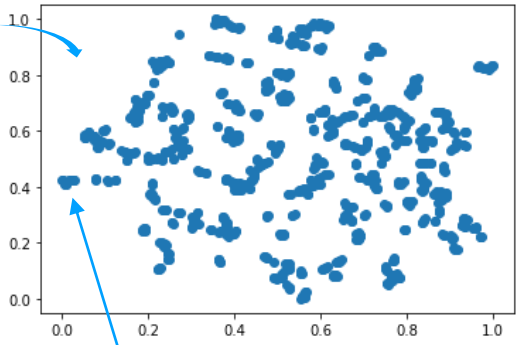
- Cell morphological readouts contain information on several bioactivity endpoints
- Features are highly correlated – we *can* remove some of them, but then we lose biologically meaningful information
- We obtain here feature maps which group correlated features, which have importance for a particular endpoint
- We can obtain per-endpoint and per-compound importance heatmaps using Grad-CAM.

Method

1. Prepare Feature Map

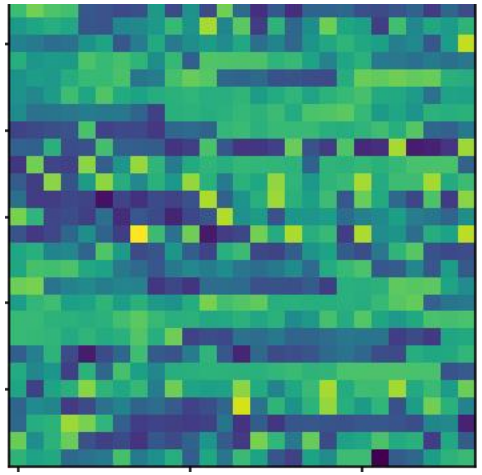
	0	1	2	3	4
Metadata_profile_id	profile_0	profile_1	profile_2	profile_3	profile_4
Metadata_cell_line	A549	A549	A549	A549	A549
Metadata_pert_name	AKT1-1	AKT1-2	ARID1B-1	ARID1B-2	ATF4-1
Cells_AreaShape_Center_Y	-0.18016	0.370572	-0.360905	0.26245	-0.110264
Cells_AreaShape_Compactness	-0.155631	-0.247842	0.79474	0.480421	-0.074895
...	...	...	...	...	...
Nuclei_Texture_SumVariance_DNA_5_0	0.923143	0.504751	-0.497296	0.063444	0.594059
Nuclei_Texture_Variance_AGP_5_0	0.944998	0.407462	-0.748232	-0.560178	0.674015
Nuclei_Texture_Variance_DNA_10_0	0.984938	0.522251	-0.51524	-0.062851	0.140325
Nuclei_Texture_Variance_DNA_20_0	1.122724	0.64437	-0.42144	0.085026	0.29123
Nuclei_Texture_Variance_DNA_5_0	0.961945	0.519441	-0.526734	0.026056	0.417465

t-SNE of Feature Map



Cell_Texture_SumAvergage_AGP_10_0

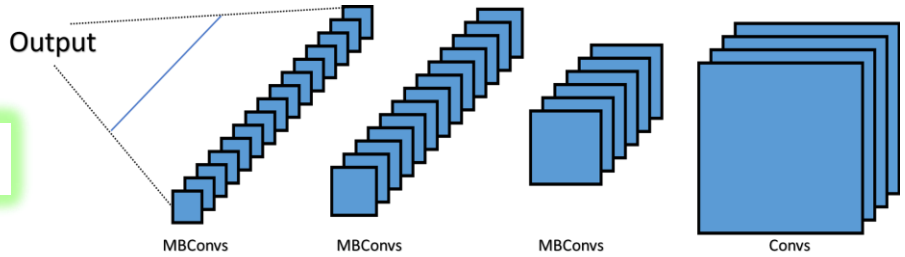
Jonker-Volgenant algorithm



Tox21 Assays

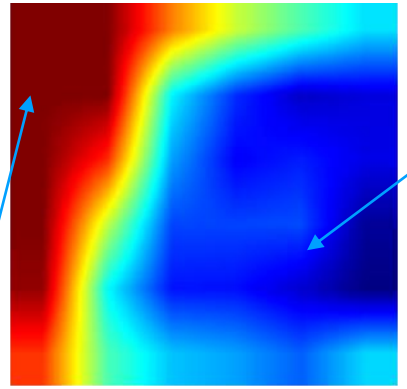
e.g. ER Stress

2. Predict Endpoint of test set



Model: EfficientNet B0

3. Interpretation using Grad-CAM



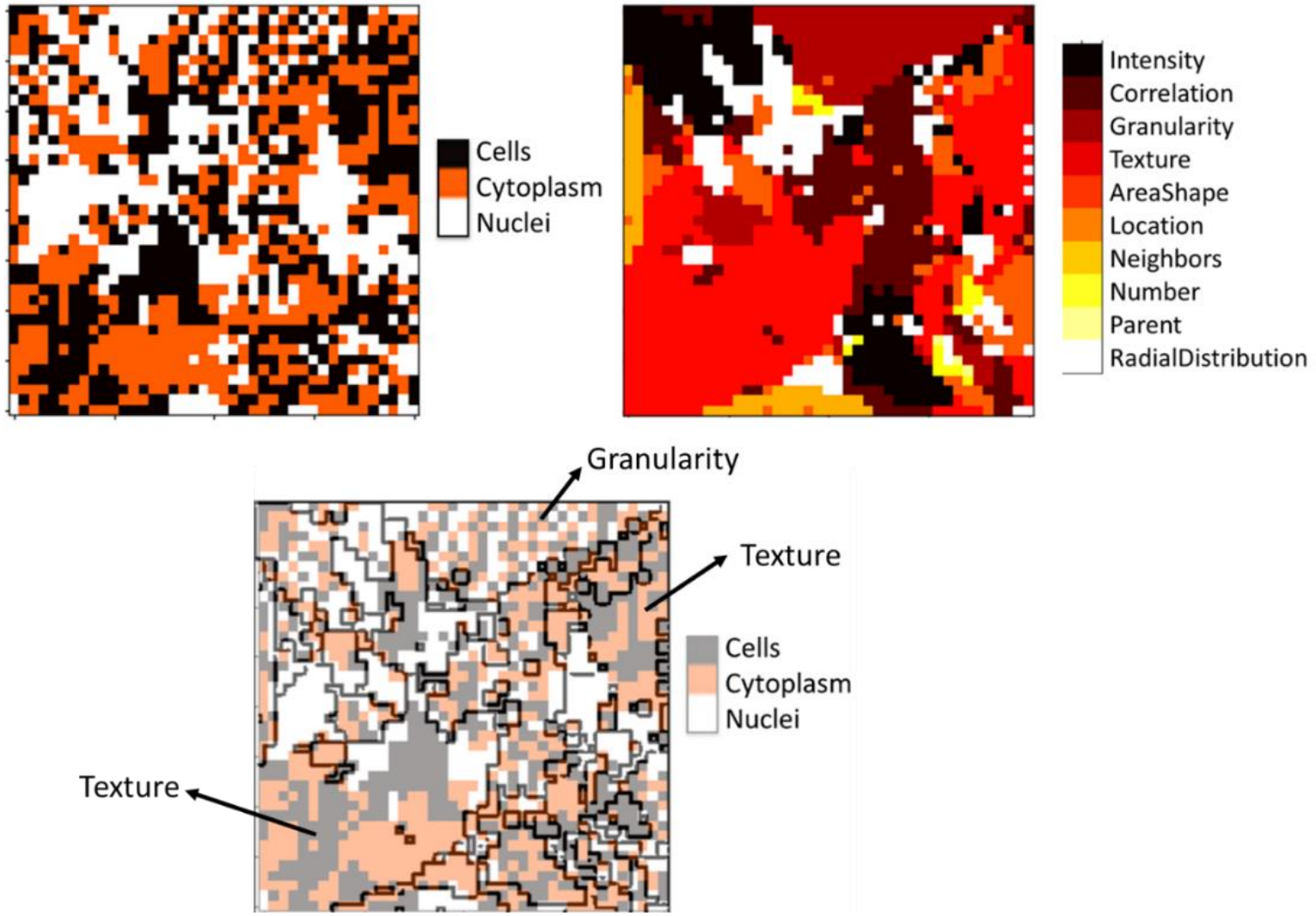
More important to model

Less contributing to model

Features are related by measurement type

- Majority of features are related by measurement function than by objects they were measured in (cells, cytoplasm, or nuclei)

- For example, granularity, features are clustered together from all compartments which means information on granularity was homogenous throughout the channels.

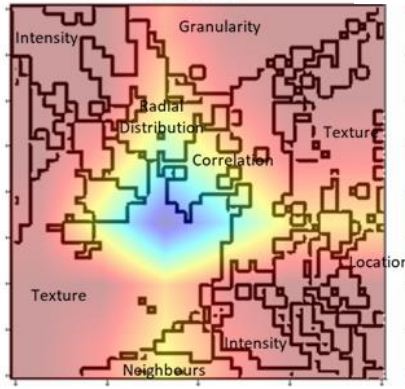


“The universe of toxic endpoints in cell painting feature space”

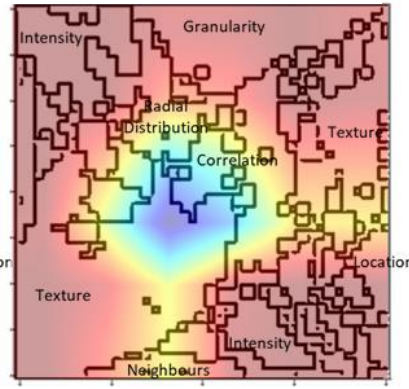
For models predicting proliferation decrease endpoint:

mitochondrial toxicity

Niclosamide

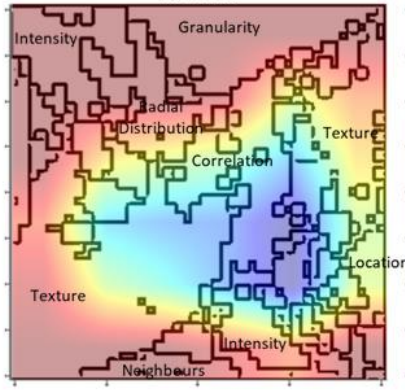


Colchicine

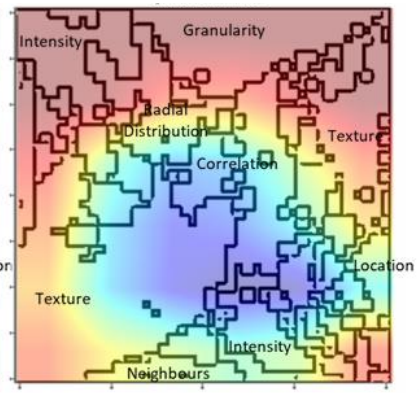


inhibits cell proliferation in the G(1) phase of the cell cycle

Lovastatin

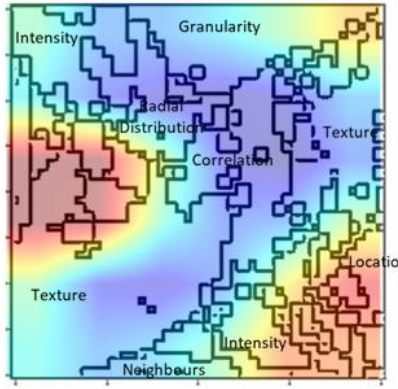


Cycloheximide

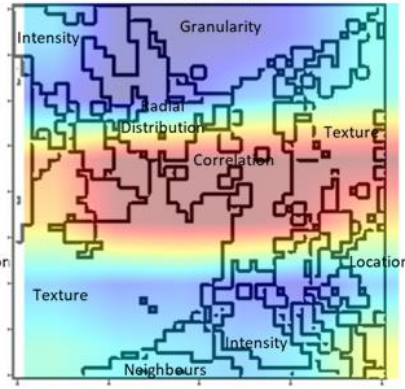


antiandrogens

Nilutamide

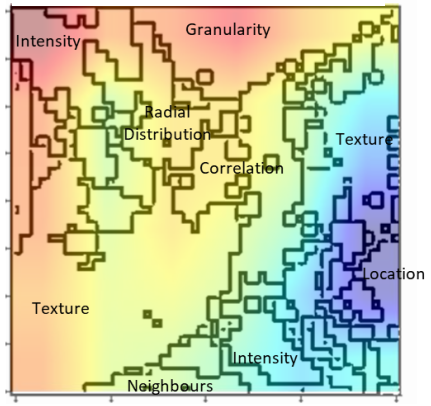


Mifepristone

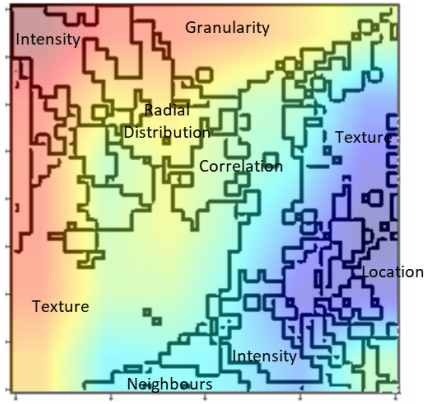


Microtubule disruptors and ER Stressors affect texture features

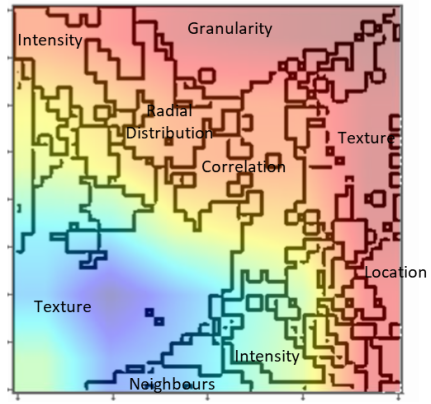
Cycloheximide



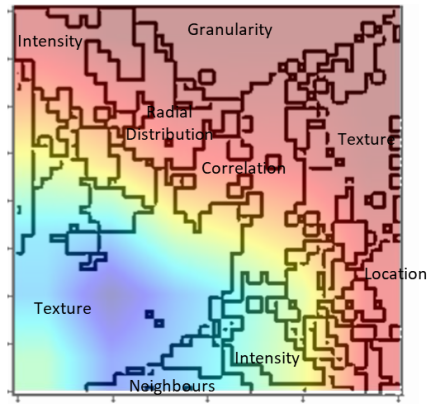
Daunorubicin



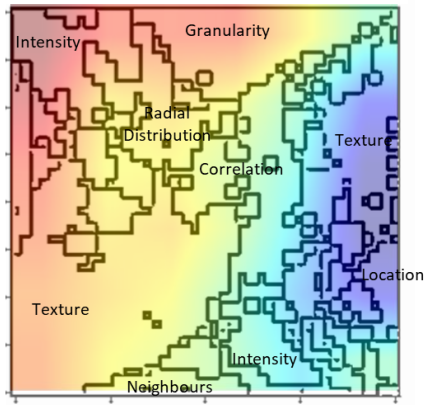
Niclosamide



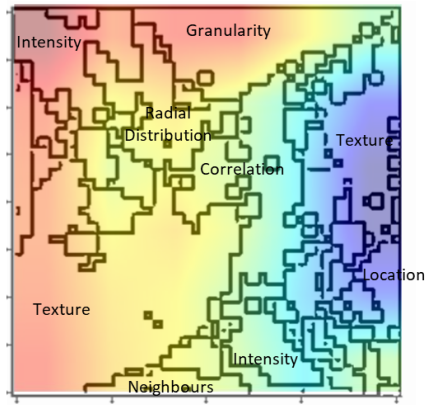
Pimozide



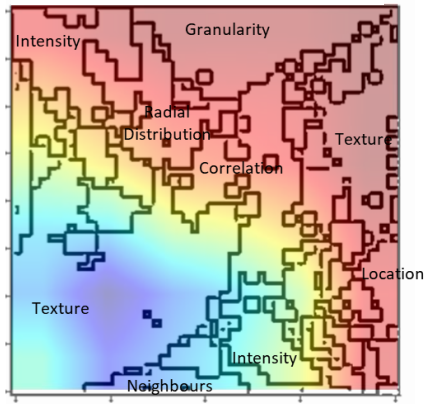
Paclitaxel



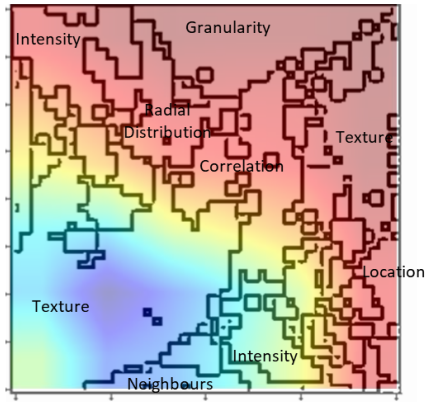
Parbendazole



Fluspirilene



2,5-Di-tert-butylhydroquinone



Microtubule disruptors causing apoptosis

Causing ER stress.

A few thoughts on –omics/cell morphology data for anticipating compound safety

- We mostly live in hypothesis-free, technology-push space – we should move to hypothesis-driven, science-pull space where we can
- ‘Sometimes you see something – but sometimes nothing, and sometimes far too much’. We often don’t know where we are when / don’t understand applicability domain of readouts
- We seem to be very good at detecting ‘the obvious’ (‘tubulin inhibitor’, ‘HDAC inhibitor’, etc.), but often not the finer details

- **To change this needs *real* consortia – including experimental design and prospective data generation, not just ‘sharing what we generated for entirely different reasons ages ago anyway’ (since this is often not what helps us now!)**

Conclusions

- We should analyze our data, absolutely!
- Life science data is difficult to label, and hence to model
- ‘Big data’ is good, but heterogeneous data makes quantitative decisions often difficult

To advance, we

- Likely need *forward-looking* consortia, for *generation* and *evaluation* of relevant data to predict *in vivo*-relevant endpoints
- Need to take care to understand applicability domain of readouts better
- Embedding *into process*, and building the *right model for decision making is key* (it’s not about ‘numbers’, outcome for the *real-world process* is what matters!)

bts driving excellence in toxicology

Lhasa Limited

Cambridge Academy of Therapeutic Sciences

RSC INTEREST GROUP CHEMICAL INFORMATION AND COMPUTER APPLICATIONS GROUP

3rd In Silico Toxicology Conference

On Zoom, free & open to all!

29 September 2022, 8.30-6pm UK/BST

*In Silico T*oxicology Consortia, Cell Painting, Gene Expression Data, Biomarkers, Interpreting Neural Networks, Drug-Induced Liver Injury/DILI, Skin Sensitization, Animal Histopathology Data, Species Concordance, *In Vivo* Pharmacokinetics (PK), Molecular Initiating Events (MIEs), Chemicals, Pharma, Food, Read-Across, … and more!

Join us and circulate!

http://www.DrugDiscovery.NET/tox2022

'In Silico modelling for dummies' session organized by the British Toxicology Society

- In November 2022
- 2 Hour session – Background, and seminar on 'how to build your own models'

- Mail me if you are interested and I will keep you posted: ab454@cam.ac.uk

Thank you for listening!
Any questions?

Contact: ab454@cam.ac.uk
Personal email: mail@andreasbender.de
Web: http://www.DrugDiscovery.NET
Twitter: @AndreasBenderUK