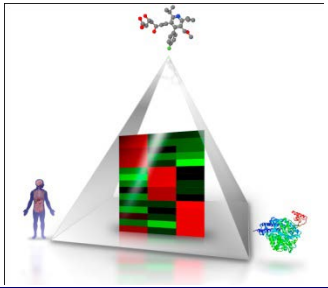
Using Chemical, Biological, and *In Vivo* Data for NAMs: Which data do we have, And what can we do with it?

Andreas Bender, PhD
Natural Philosopher for Molecular Informatics
Centre for Molecular Science Informatics
University of Cambridge
Fellow of King's College, Cambridge
Associate Director, Clinical Pharmacology & Safety Sciences
AstraZeneca, Cambridge

UNIVERSITY OF CAMBRIDGE



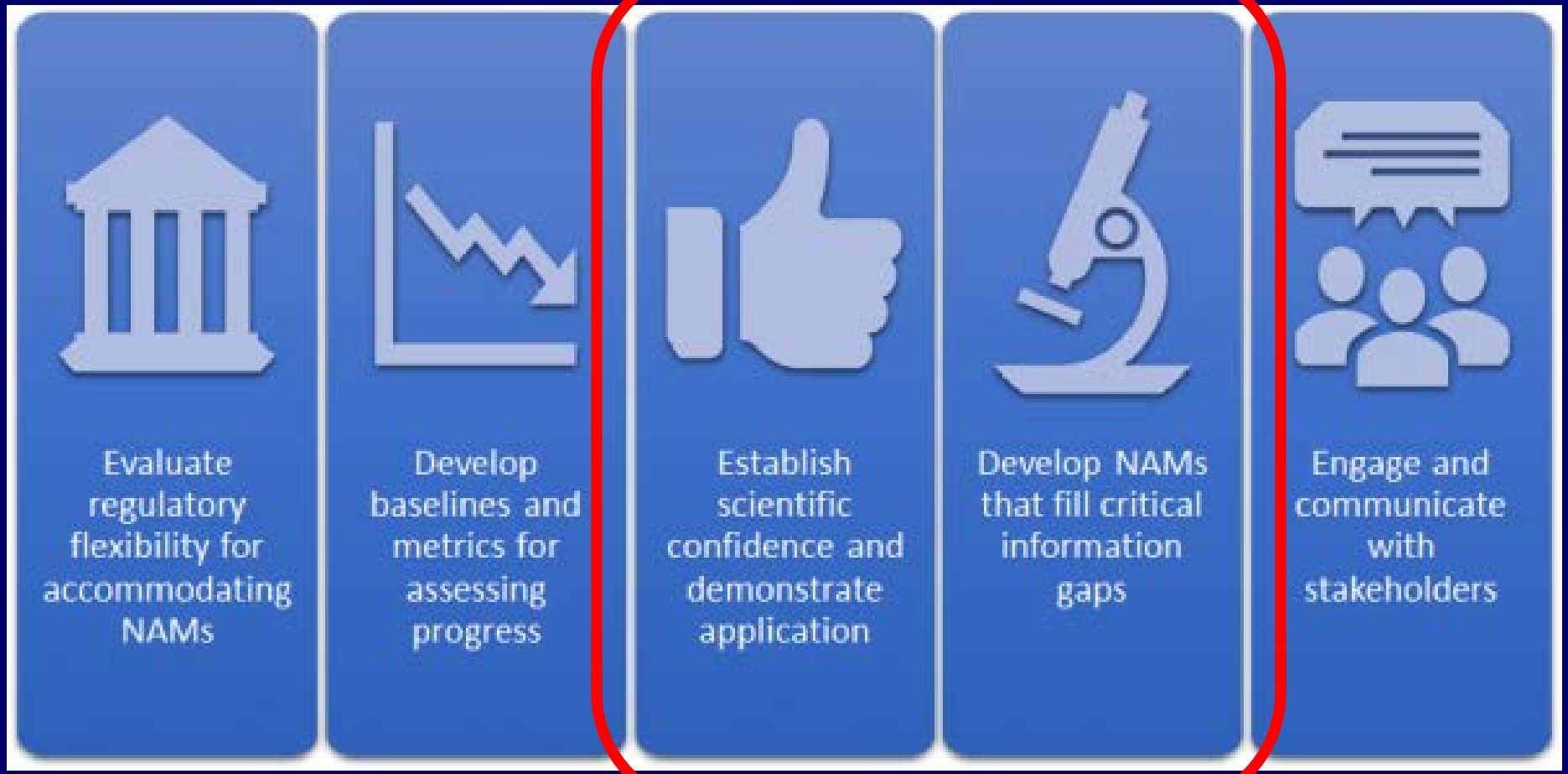
Outline

- Chemical, biological and phenotypic data – some key characteristics

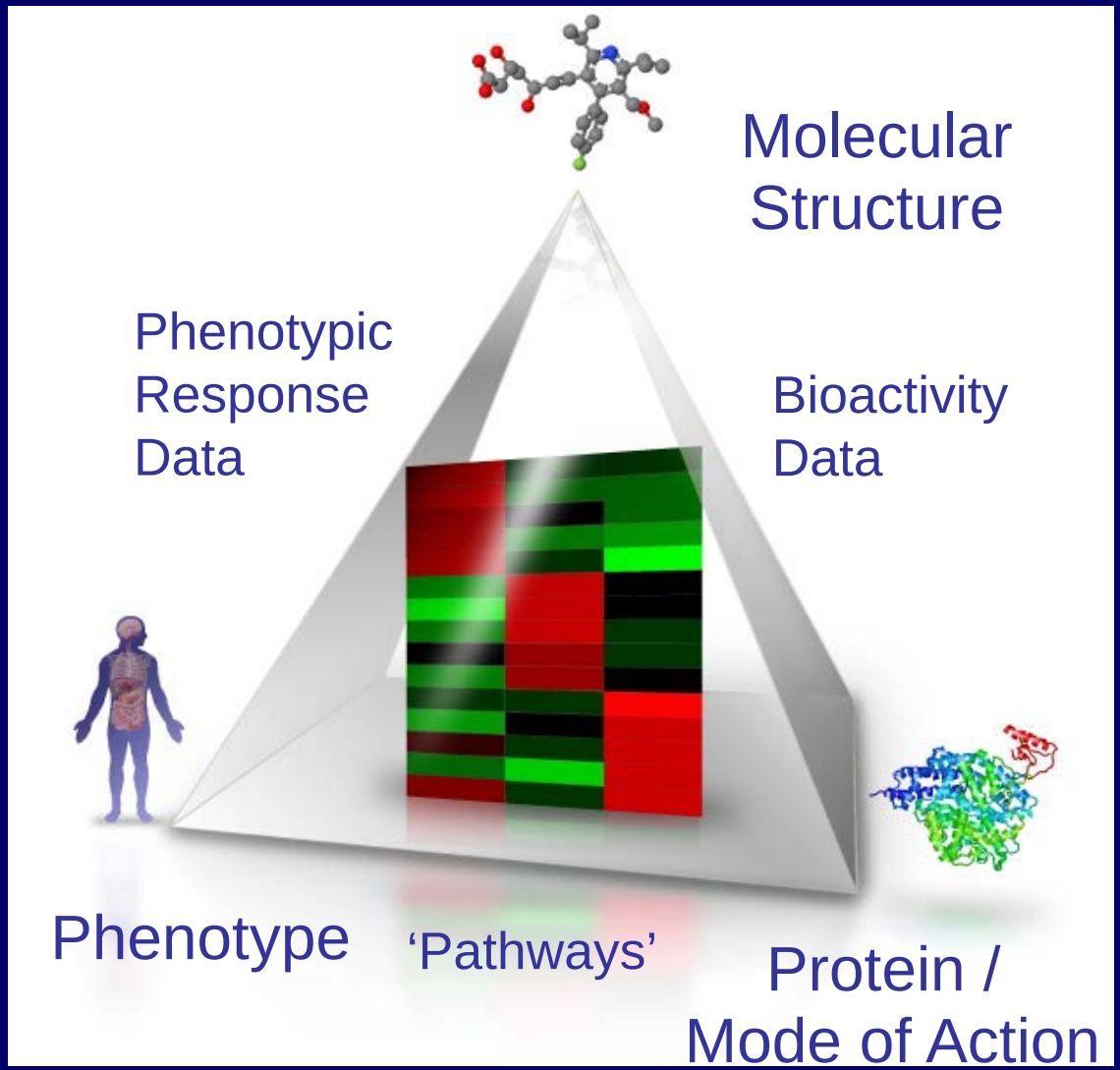
- From compound to biological readout to endpoint: What do we see, and understand?
  - Linking chemistry/assay to adverse endpoint
  - Is –omics data the solution?
  - Learning from existing animal data: Making use of what we know already

- Conclusions

Where does our work fit into NAMs?



Core Data Considered: Chemistry, Phenotype, Targets / Mode of Action



Data/'AI' in early discovery vs safety

Early discovery/hazard detection

- 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

Quantitative safety/risk prediction

- 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: *Recording* vs data *suitable for mining* – eg animal data tricky, even within single company

Problem setting in early discovery vs safety

Early discovery/hazard detection

- 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

Quantitative safety/risk prediction

- Need to predict for *this particular data point*
- *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!)
- Even ‘protecting’ based on evidence of absence

Brief case studies

- Chemistry to endpoint
  - Data-driven derivation of AOPs
  - Predictivity of assays for adverse outcomes
- Gene expression/cell morphology data to endpoint
  - What do we *really see* in gene expression data?
  - Stress pathways: What do we *define as relevant*?
  - Case study: Markers for Drug-Induced Vascular Injury
- Learning from existing animal data

Using data to derive AOPs for Structural Cardiotoxicity

- Fredrik Svensson, Azedine Zoufir, Samar Mahmoud, Avid Afzal, Ines Smit, Kathryn Giblin (University of Cambridge)
- Peter Clements, Jim Harvey, Jordi Munez-Muriedas (GSK)
- Jay Mettetal, Amy Pointon, Nigel Greene (AZ)
- Richard Williams (Lhasa Ltd.)

Information-Derived Mechanistic Hypotheses for Structural Cardiotoxicity

Fredrik Svensson,[*,†] Azedine Zoufir,[†] Samar Mahmoud,[†] Avid M. Afzal,[†] Ines Smit,[†] Kathryn A. Giblin,[†] Peter J. Clements,[‡] Jerome T. Mettetal,[§] Amy Pointon,[∥] James S. Harvey,[‡] Nigel Greene,[⊥] Richard V. Williams,[#] and Andreas Bender[*,†]

[†]Centre for Molecular Informatics, Department of Chemistry, University of Cambridge, Lensfield Road, Cambridge CB2 1EW, United Kingdom

[‡]GlaxoSmithKline R&D Ltd, Park Road, Ware, Hertfordshire SG12 0DP, United Kingdom

[§]Drug Safety and Metabolism, AstraZeneca, 35 Gatehouse Drive, Waltham, Massachusetts 02451, United States

[∥]Safety and ADME Translational Sciences, AstraZeneca, Cambridge Science Park, Milton Road, Cambridge CB4 0WG, United Kingdom

[⊥]Predictive Compound Safety and ADME, AstraZeneca, 35 Gatehouse Drive, Waltham, Massachusetts 02451, United States

[#]Lhasa Limited, Granary Wharf House, 2 Canal Wharf, Leeds LS11 5PS, United Kingdom

- Svensson et al., Chem. Res. Tox 2018

Project motivation and setup

- Mechanisms behind structural cardiotoxicity are poorly understood
- Working group with GSK, AZ, Lhasa
- Establish prototype workflow to use data from different sources (ChEMBL, *ToxCast*, *FAERS*, …) to derive Adverse Outcome Pathways

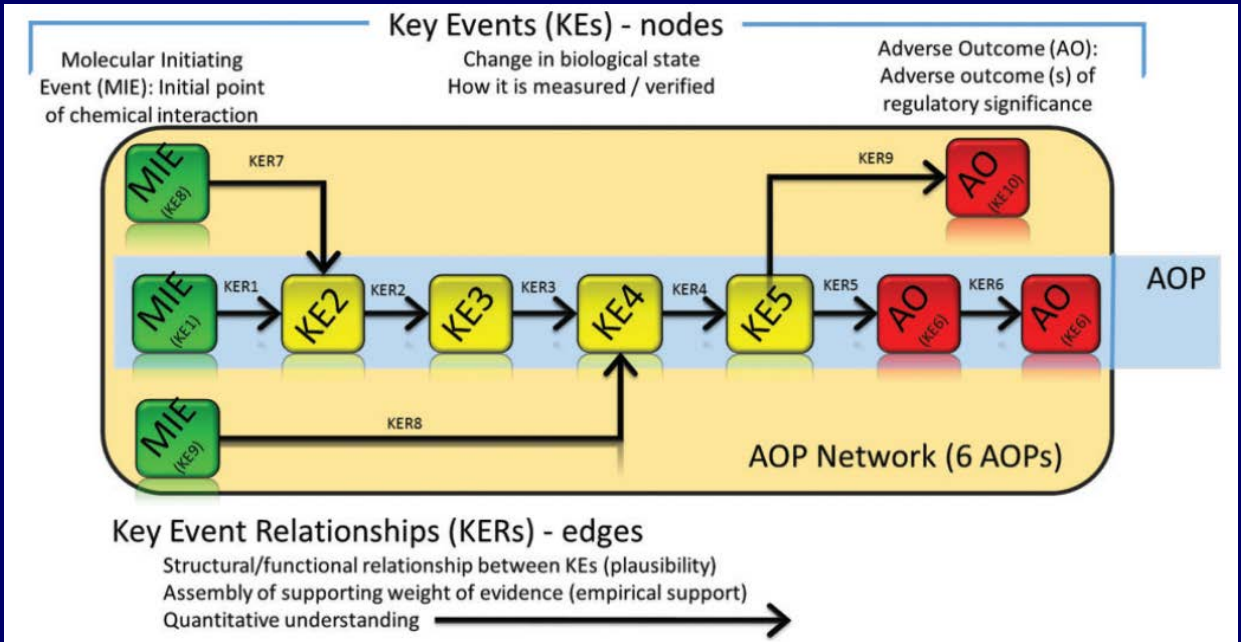


FIG. 1. Conceptual representation of the adverse outcome pathway (AOP) framework including modular representation as Key Events (KEs) and Key Event Relationships (KERs), 2 specialized types of KEs, molecular initiating events (MIEs) and adverse outcomes (AOs), that serve as upstream and downstream anchors in an AOP, and assembly of multiple AOPs sharing common KEs and/or KERs into an AOP network.

Wittwehr 2017

What is needed for meaningful Key Event Relationships?

- OECD recommendation: Capture general (functional) form of relationship, range of uncertainty, known sources of uncertainty, approximate time scale, …

- Modelling: AOPs with more/less knowledge can be modelled in different ways (quantifying uncertainty)

- *But:* In practice we know *few* AOPs, and they are often not sufficiently quantitative for practical use (at this point in time)

Workflow of project – ToxCast and FAERS Data

1. Correlation between ToxCast data and FAERS events
2. Identification of AE-ToxCast (KE) relationships
3. *Manual* construction of putative AOPs from KEs

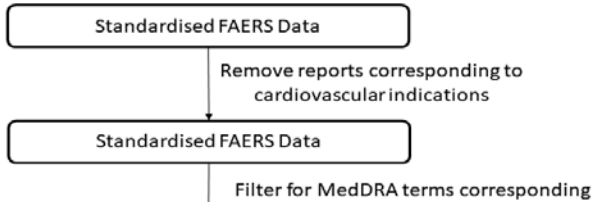


Table 1. The 22 associations derived through the mutual information analysis. Rows including assay readouts that could be linked to structural cardiotoxicity in the literature survey are highlighted in grey.

Adverse Event	ToxCast Assay*	Assay Target/Readout	Mutual Information	Pearson Correlation
Cardiac failure congestive	BSK_BE3C_PAI1_down	PAI-1	0.120	0.369
Cardiac failure congestive	BSK_4H_Eotaxin3_down	CCL26	0.116	0.366
Left atrial dilatation	ATG_PXRE_CIS_up	NR1I2	0.110	0.332
Left atrial dilatation	NVS_NR_cAR	AR	0.105	0.337
Mitral valve incompetence	NVS_MP_rPBR	TSPO	0.104	0.338
Mitral valve prolapse	NVS_NR_bER	ERα	0.096	0.322
Cardiac failure congestive	Tox21_PPARg_BLA_antagonist_ratio	PPARγ	0.096	0.314
Left atrial dilatation	ATG_VDRE_CIS_up	VDR	0.096	0.323
Mitral valve incompetence	NVS_NR_cAR	AR	0.084	0.320
Mitral valve prolapse	NVS_NR_hER	ERα	0.082	0.299
Mitral valve incompetence	BSK_BE3C_IP10_down	CXCL10	0.081	0.309
	BSK_CASM3C_MCSF_do	CSF1		0.282

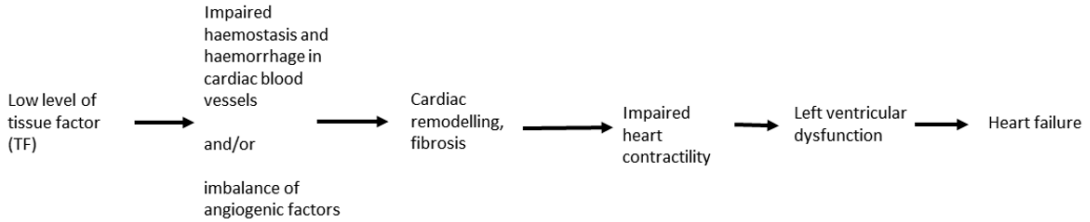


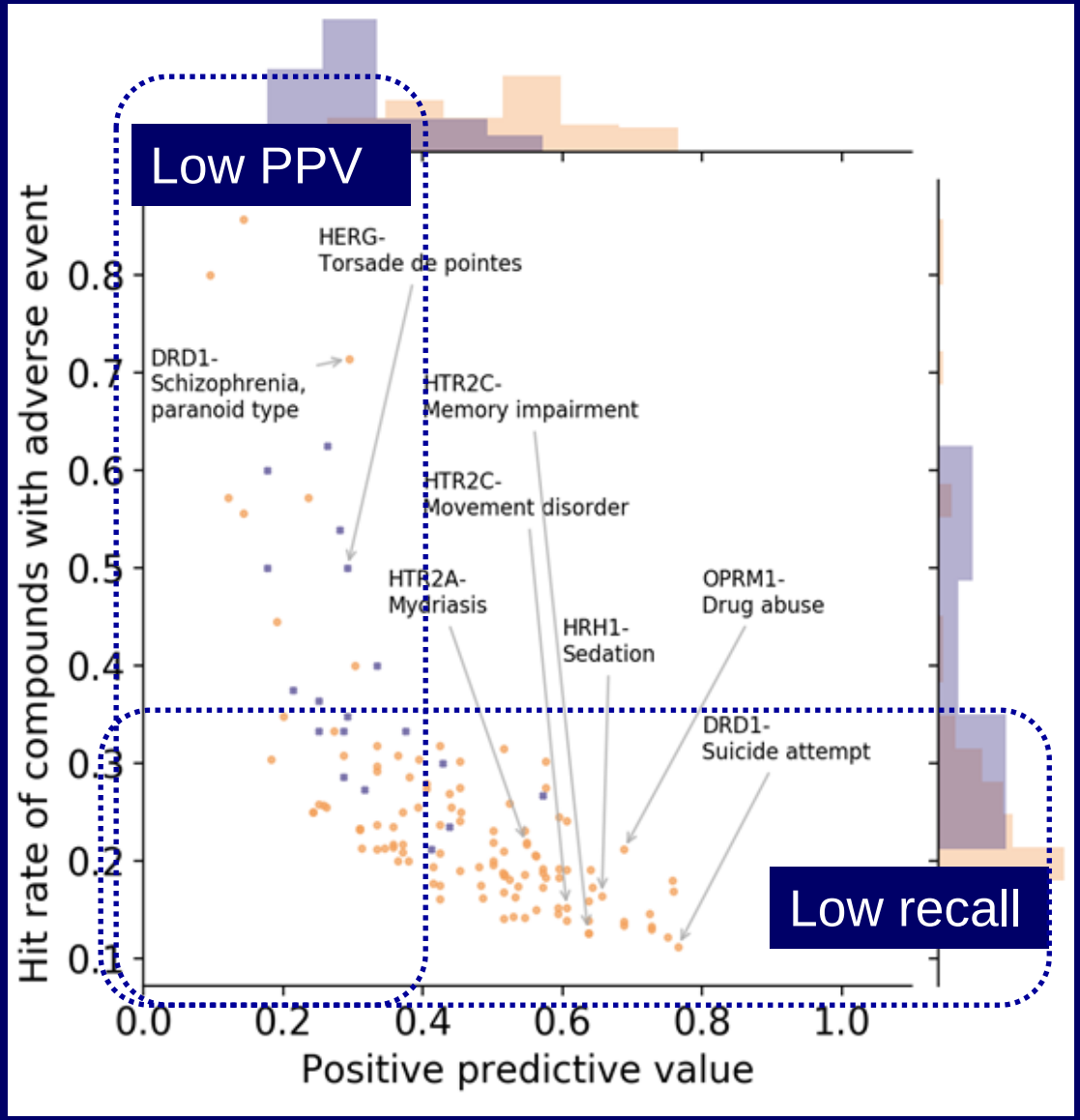
Figure 2. Putative AOP connecting TF to heart failure *via* left ventricular dysfunction.

Protein target/assay to adverse outcome links are tricky (low recall)

Safety profiling/ adverse event (FAERS) links

Smit *et al.* BioRxiv 2020

https://doi.org/10.1101/2020.06.12.135939



Target-based compound profiling: Some learnings

- Compound overlap between sources (very) low – leads to lack of associations, biases, *etc.*

- Data sources where compounds structures overlap across key events is crucial

- Probably *joint data generation* (not just sharing) between companies in safety area needed

- Chemical and biological coverage problem

- Interplay of data and manual review crucial (!)

- Key events from ‘omics data? (Qualified ‘yes’)

So does gene expression / cell painting data help?

There is hope for better coverage of biology with less experimental burden…

Evaluation of Gene Expression Data in Lead Optimization: QSTAR Project

- Eight drug discovery projects in ‘big pharma’
- Led Johnson&Johnson/Janssen, many partners
- *Prospective* study of (then) ‘live’ projects

- ***Sometimes* hardly any response in GE space, *sometimes* ‘too much response’ of system, *sometimes* interpretable… applicability domain of gene expression data??**

- Decisions only possible with *clear rationale for interpretation*

- Verbist et al., in *Drug Discovery Today* 2015

So what is the reference frame for gene expression data?

- Define *what matters*, establish *what we can see in a readout*

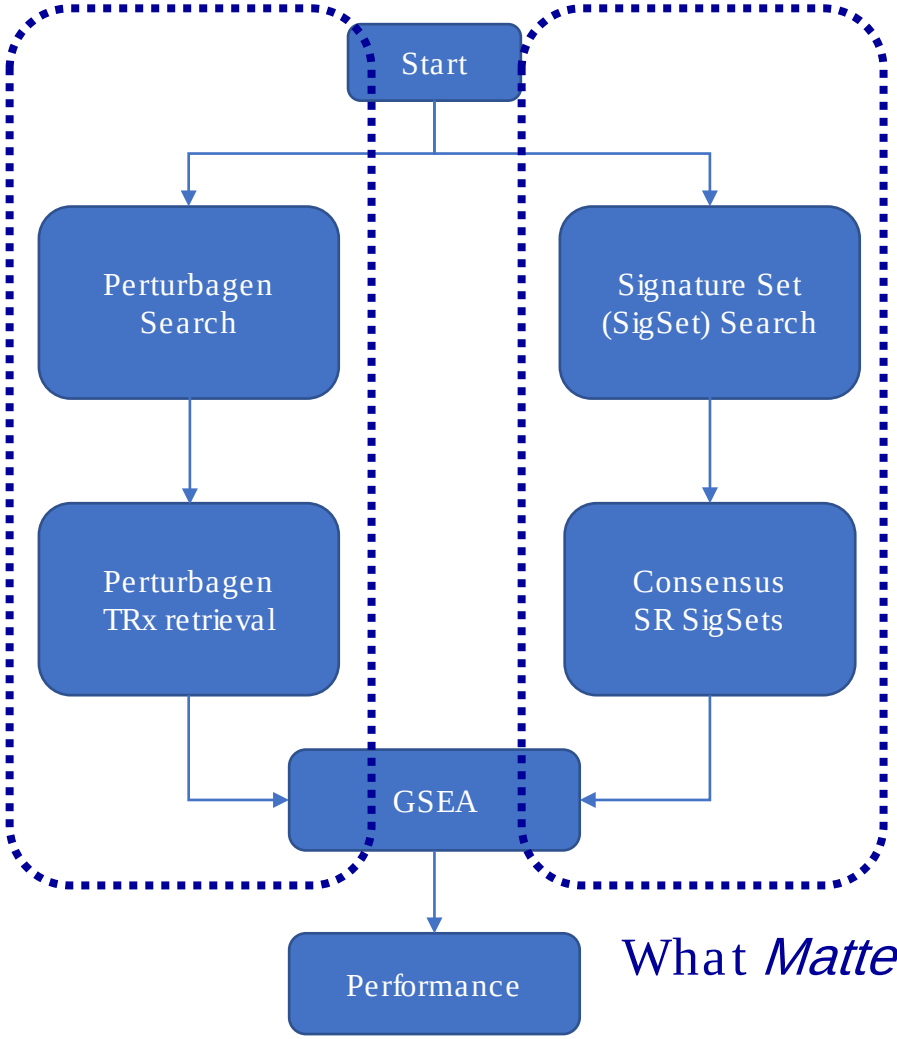
- Eg ‘Stress Pathways’ (Oxidative stress, Heat shock response etc.; Simmons 2009)
- ***BUT*: Need to define setup (dose/time point/cell line); method; parameters; thresholds… how to use in practice!**

- Work with Imran Shah, Bryant Chambers (EPA); Danilo Basili (Cambridge); Alistair Middleton (Unilever)

- Large number of compounds tested dose/time-resolved manner at EPA, once shared will be a rich data resource!

Aligning what we can see in the data to what matters to us: Approach to evaluate stress response pathway activity

1. Find reference chemicals
2. Reference chemical transcriptomics dataset (TRx)
3. Construct consensus SR signature sets
4. Evaluate performance of signature sets for characterizing reference chemicals

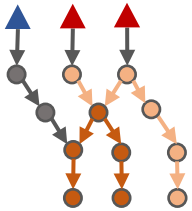


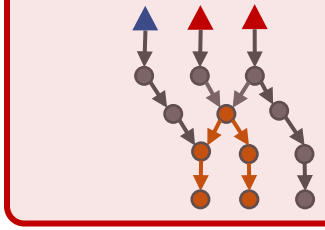
What *Can We See* In The Data?

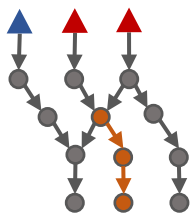
What *Matters*?

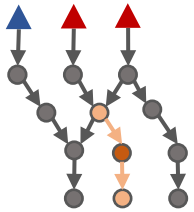
Slide adjusted from Bryan Chambers/EPA

A spectrum of ways exists to look at 'gene expression data'…. (slide/work by Anika Liu, with Jordi Munoz-Muriedas, Deidre Dalmas Wilk/GSK)









•	• Most likely downstream processes where AOP converges	• Distinguishes phenotype	• Distinguishes phenotype + large effect size
•	• Misses upstream regulation which is likely compound-specific	• Might miss key parts of AOP • Importance != Specificity	• Little insight into mechanism • Importance != Effect size

Mechanism ← **Question** → **Biomarker**

High coverage ← **Detected biological entity** → **Strong evidence**

Compound
- No adverse event
- Adverse event

Biological entity (Protein/pathway/..)
- "Mechanism"
- Maybe "mechanism" (depending on evidence)
- Not "mechanism"

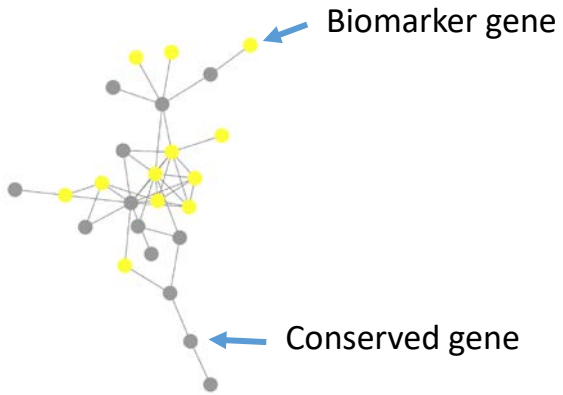
UNIVERSITY OF CAMBRIDGE gsk

How are the identified genes potentially linked with mechanisms in DIVI pathogenesis?

1) **Consistency** across conditions with DIVI
2) **Specificity** for conditions with DIVI
3) **Dose-dependency** of expression change for compounds with DIVI
4) **Large (measurable) effect** across conditions with DIVI

Conserved genes across DIVI conditions are highly interlinked on protein level

Endothelium
(50 conserved genes)



p-value: 5.6e-12

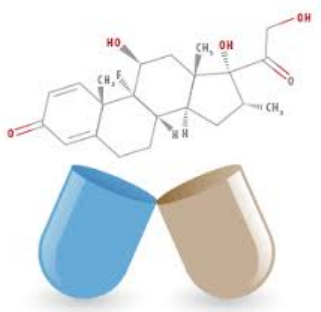
So did/does gene expression (and high-dimensional biology) data help?

- Yes, *but*…
- Experimental
  - How to pick cell line, dose, time point, setup (to ensure relevance for *in vivo* setting)? Approximating exposure (?)
  - You also don't always 'see something' in gene expression data, large chemical space/pathway bias (NHRs vs other targets etc.)!
  - Danger of too generic/hypothesis-free readouts (both for gene expression and cell morphology etc. screens)

- Data analysis:
  - Where do we want to be on the high coverage vs specificity spectrum?
  - Results *hugely* depend on experimental setup *and* data analysis, needs to be standardized in some form

Analysis of historical animal data: eToxSys database (work of Peter Wright; funded by NC3Rs; with Lhasa Ltd.)

Database of repeat-dose *in-vivo* studies



- Dataset donated by eTOX partners including large pharmaceutical companies (e.g. GSK, AstraZeneca), universities, and SMEs
- Dataset was obtained from Lhasa (non-profit honest broker)
- 1210 unique compounds (after standardization [3])
- 5026 unique studies
- Histopathological findings across all conceivable organ types

Four levels of study covariates

- Species (10)
- Strain (1-16 unique per species)
- Sex ('Male', 'Female', or 'Both')
- Administration route (1-20 unique per species)

3. https://github.com/*flatkinson*/*standardiser*

Three main aims of analyzing historical animal data (*ongoing work*)

- **Background rates** (to have better estimates of treatment-related effects)

- **Inter-species concordance** ('do we need a second species'?):
  - Presence of pathology in one species is *sometimes* (but *rarely*) diagnostic of presence in another
  - Absence of pathology in one species almost never diagnostic of absence in another

- **Are early timepoints predictive for late time points** ?
  - Yes, but only some timepoints
  - Yes, but only (very) few organs (eg liver)

Some comments on animal data

- Still real lack of annotations, data, standardization (SEND helps to an extent… but the problem that remains is the terminology inside the domains)

- Databases such as eToxSys are still underanalyzed

- … as long as we don’t even understand existing animal data it will be difficult to compare new methods to them and replace them!

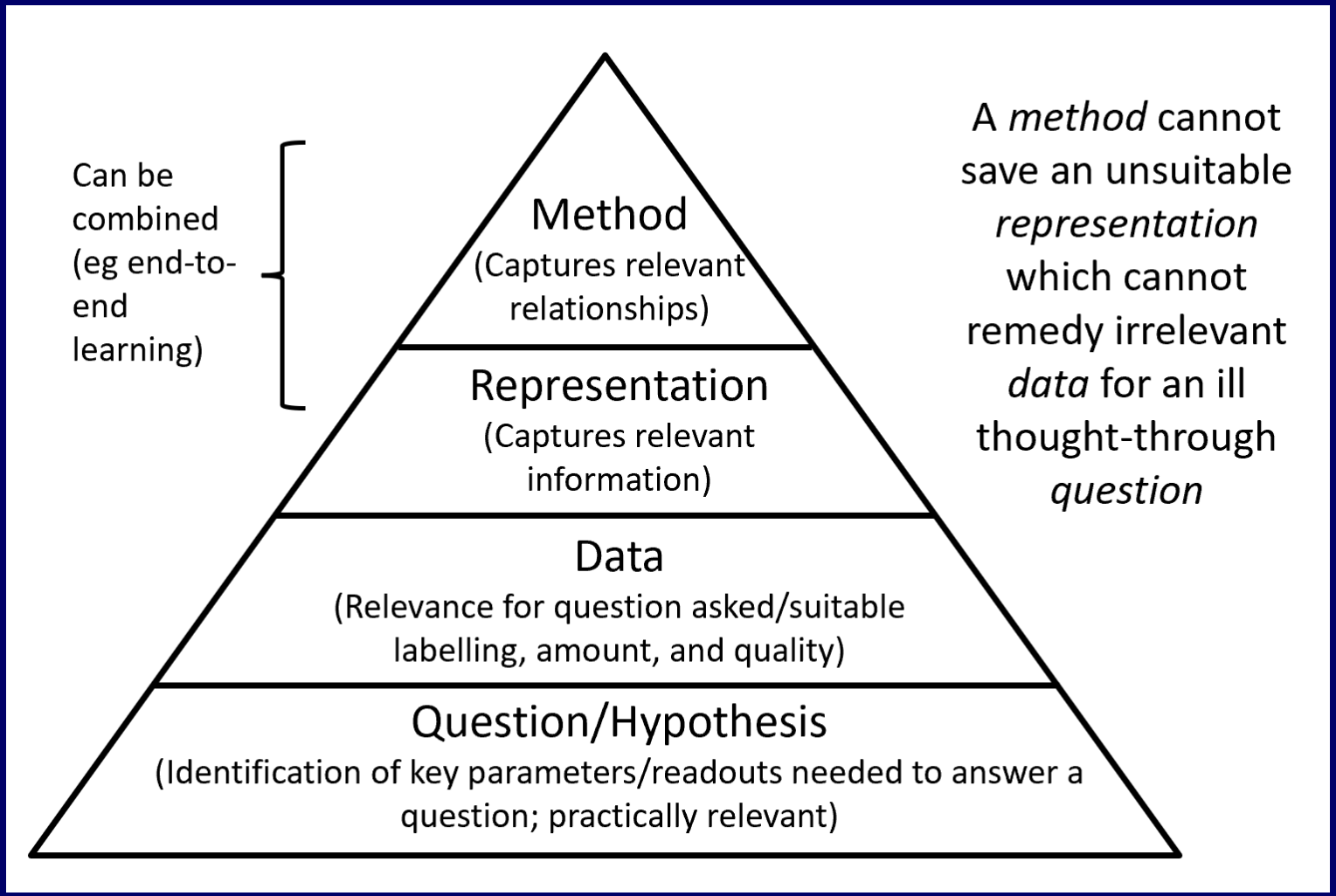
Key Learnings

- Computational models are only as good as the underlying data
- ‘Data’ is not ‘data’ – it needs to be useful for the problem at hand

- *There is no shortcut* – we will always run into coverage problems (chemically, mechanistically etc.)
- We *do* need consortia (not only data sharing, but data generation!)

- Data needs to be *relevant to the question at hand* and be stored in proper ontologies

Article on computational model validation: http://Drugdiscovery.net/HowToLie (also upcoming Drug Discovery Today article)



In Silico Toxicology Network Meeting: 15 September 2021, on Zoom, open to all

- 2020 meeting on 30 September 2020, with 320 registered participants, contributions from pharma, consumer goods, agrochemicals, etc.

- **2021 meeting on 15 September 2021**

- Please let me know if you are interested in presenting
- Registration for participants opens 1 July 2021

- **http://drugdiscovery.net/tox2021/**

Thank you for listening

Contact: Andreas Bender, ab454@cam.ac.uk