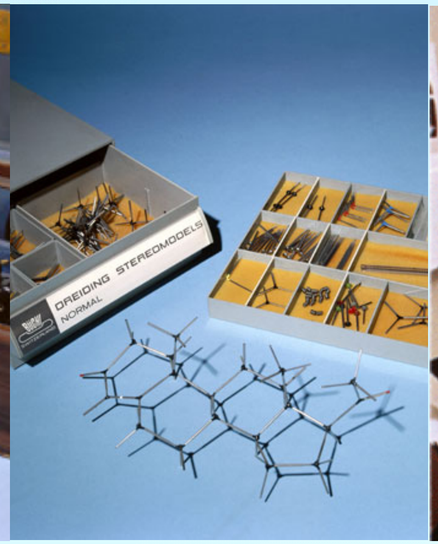
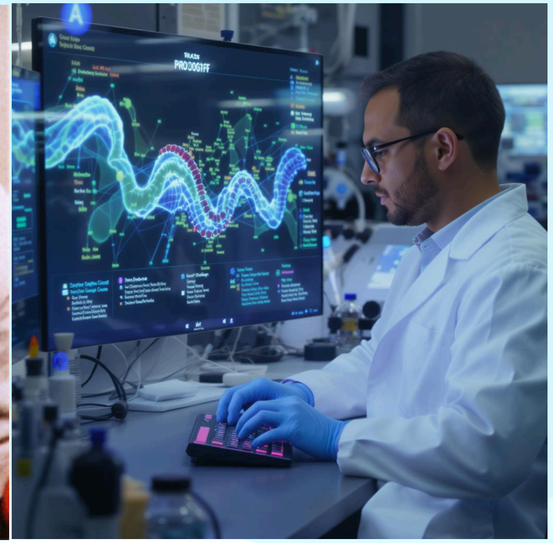
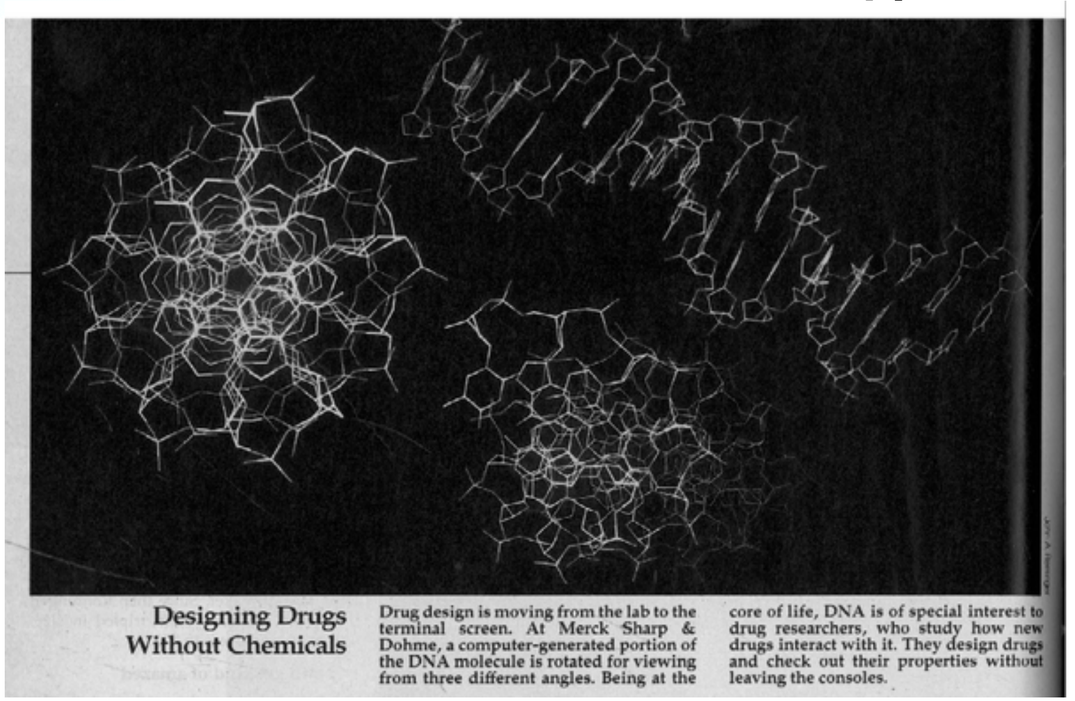




CADD: Where Have We Come From and Where Are We Going?

Learning from the mistakes of the last 37 years



Designing Drugs Without Chemicals

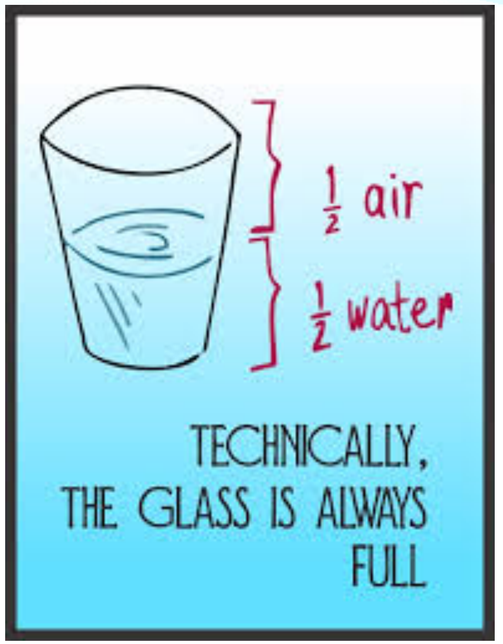
Drug design is moving from the lab to the terminal screen. At Merck Sharp & Dohme, a computer-generated portion of the DNA molecule is rotated for viewing from three different angles. Being at the core of life, DNA is of special interest to drug researchers, who study how new drugs interact with it. They design drugs and check out their properties without leaving the consoles.

Richard Lewis GRC 2025

Overview

Themes to be riffed on

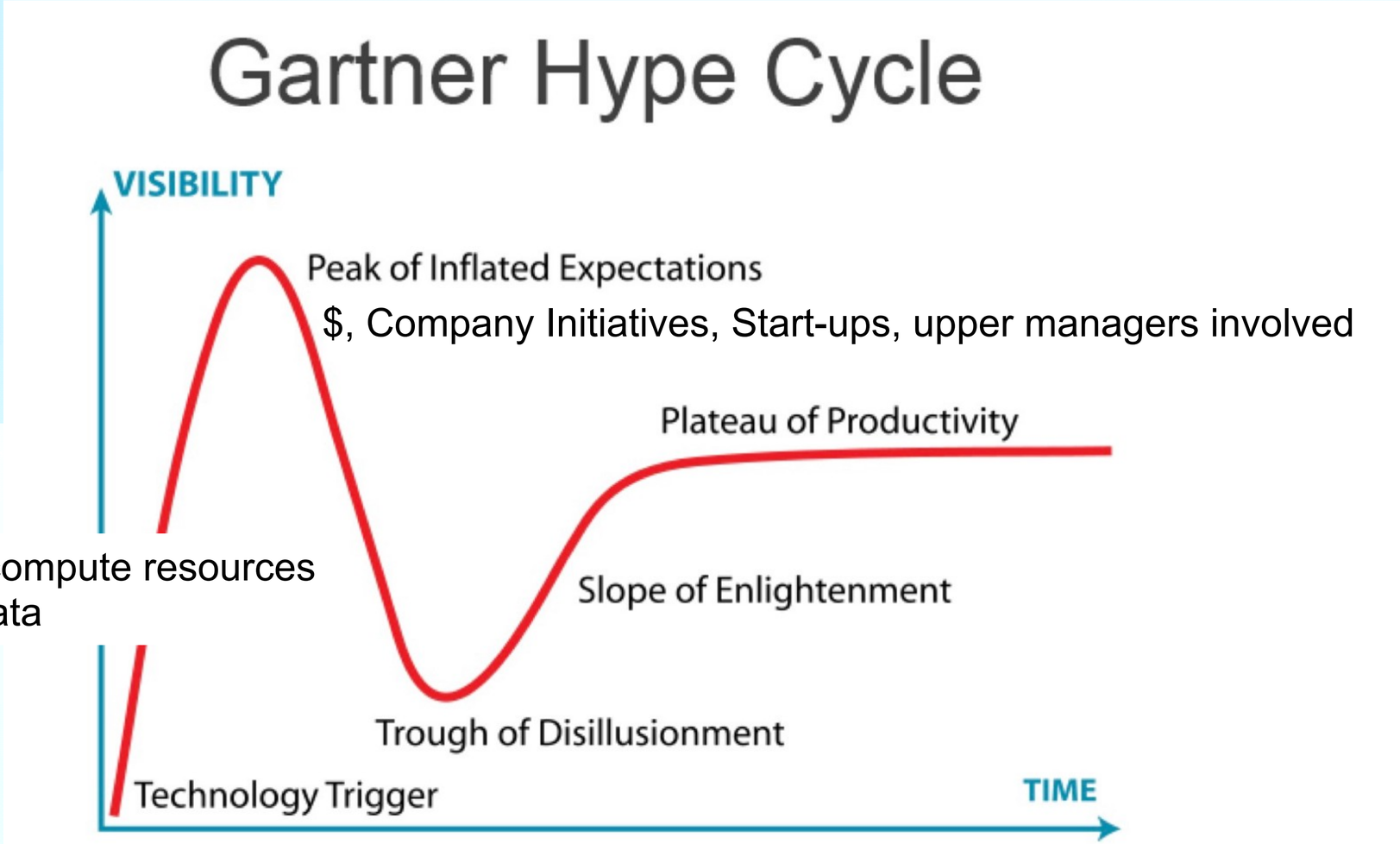
- Where did we go wrong?
- Have we got smarter or is it all down to improved compute resources?
- Do/Can we validate our work and our data sufficiently?
- How can we best share our work?
- Challenges for the future



Only a fool learns from his own mistakes. The wise man learns from the mistakes of others. Bismarck

Richard Lewis

The enemy of the good



Better compute resources
More data

What is a model

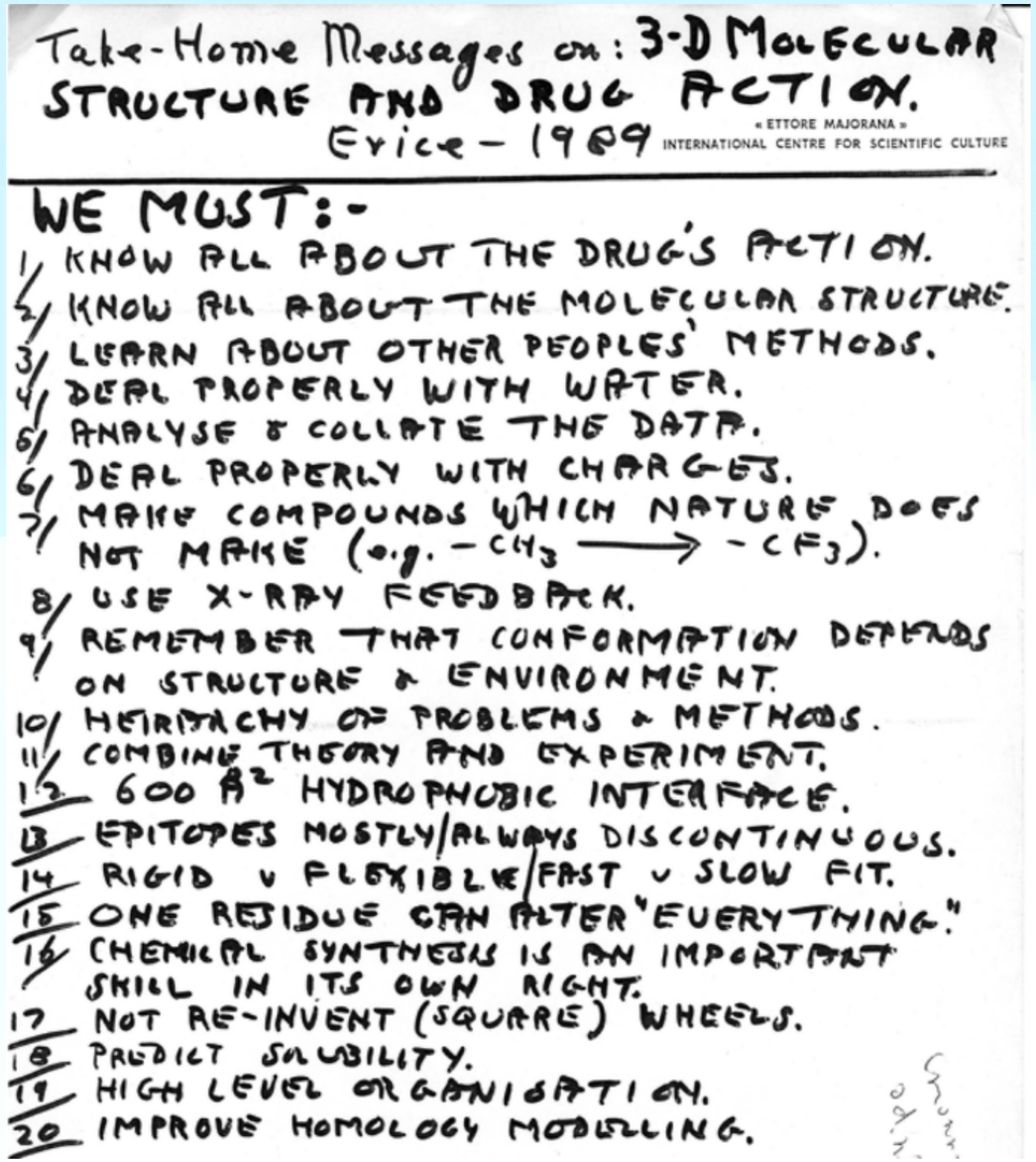
- Any explanation built on prior observations
- It should also explain future observations or be modified/abandoned
  - Good tests kill flawed theories; we remain alive to guess again. K. Popper
    - Those who promise us paradise on earth never produced anything but a hell. K. Popper, miserable so-and-so.

Where we were

- Very little data (and certainly not FAIR!)
  - PDB had 365 entries in 1989
  - The ClogP starlist was the largest at 10k points
    - www.daylight.com/meetings/mug98/Leo/clogp_history.html
- Very little computing power (compared to today)
  - SGI Indigo 4*32-bit cpu @ 10 MHz, 96MB RAM
  - QM was AM1 in small doses, MD was a few ns.
- Few CADD groups in Pharma

In the beginning there was Peter Goodford…

Erice 1989



Take-Home Messages on: 3-D MOLECULAR STRUCTURE AND DRUG ACTION.
Erice – 1989
« ETTORE MAJORANA »
INTERNATIONAL CENTRE FOR SCIENTIFIC CULTURE

WE MUST:-

1/ KNOW ALL ABOUT THE DRUG'S ACTION.
2/ KNOW ALL ABOUT THE MOLECULAR STRUCTURE.
3/ LEARN ABOUT OTHER PEOPLES' METHODS.
4/ DEAL PROPERLY WITH WATER.
5/ ANALYSE & COLLATE THE DATA.
6/ DEAL PROPERLY WITH CHARGES.
7/ MAKE COMPOUNDS WHICH NATURE DOES NOT MAKE (e.g. $-CH_3 \longrightarrow -CF_3$).
8/ USE X-RAY FEEDBACK.
9/ REMEMBER THAT CONFORMATION DEPENDS ON STRUCTURE & ENVIRONMENT.
10/ HEIRARCHY OF PROBLEMS & METHODS.
11/ COMBINE THEORY AND EXPERIMENT.
12/ 600 $Å^2$ HYDROPHOBIC INTERFACE.
13/ EPITOPES MOSTLY/ALWAYS DISCONTINUOUS.
14/ RIGID v FLEXIBLE/FAST v SLOW FIT.
15/ ONE RESIDUE CAN ALTER "EVERYTHING."
16/ CHEMICAL SYNTHESIS IS AN IMPORTANT SKILL IN ITS OWN RIGHT.
17/ NOT RE-INVENT (SQUARE) WHEELS.
18/ PREDICT SOLUBILITY.
19/ HIGH LEVEL ORGANISATION.
20/ IMPROVE HOMOLOGY MODELLING.

What have we solved: homology models thanks to the large amount of high quality curated structures. 365 in 1989 to 238622 this year.

Cruciani, G., Martin, Y., Vinter, A. et al. How computational chemistry develops: a tribute to Peter Goodford. J Comput Aided Mol Des 33, 699–703 (2019).

Richard Lewis

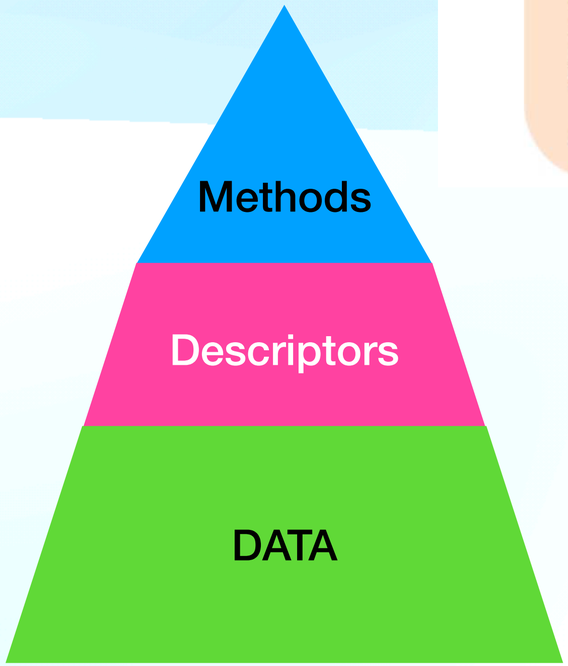
Data - Good data _should_ be the basis of everything

Apologies to Greg Landrum

Vas Narasimhan, CEO of Novartis AG

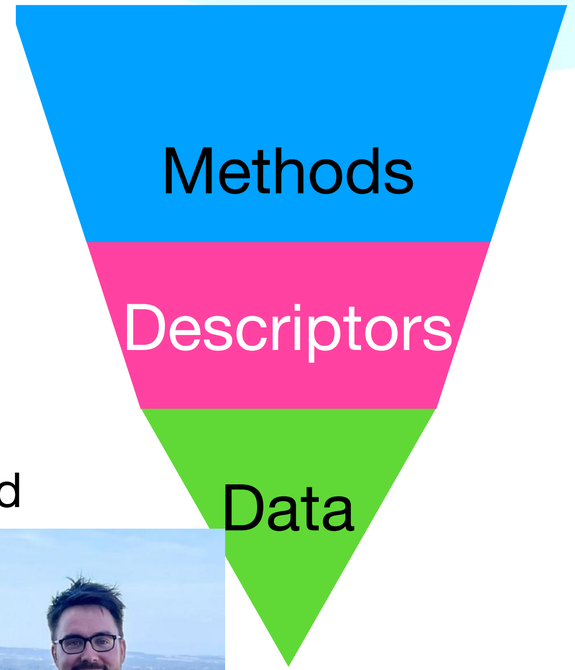
"The first thing we've learned is the importance of having outstanding data to actually base your ML on. In our own shop, we've been working on a few big projects, and we've had to spend most of the time just cleaning the data sets before you can even run the algorithm. That's taken us years just to clean the datasets. I think people underestimate how little clean data there is out there, and how hard it is to clean and link the data."

Forbes, Jan 2019



This is not just for ML, but also for FEP, docking,MD
ML amplifies the issue

- Examples to illustrate these points will be drawn from "Machine Learning for Fast, Quantum Mechanics-Based Approximation of Drug Lipophilicity" Isert et al. ACS Omega 2023, 8, 2, 2046–2056



Richard Lewis

Where does data come from?

The foundation of any model (and any process that uses a model…)

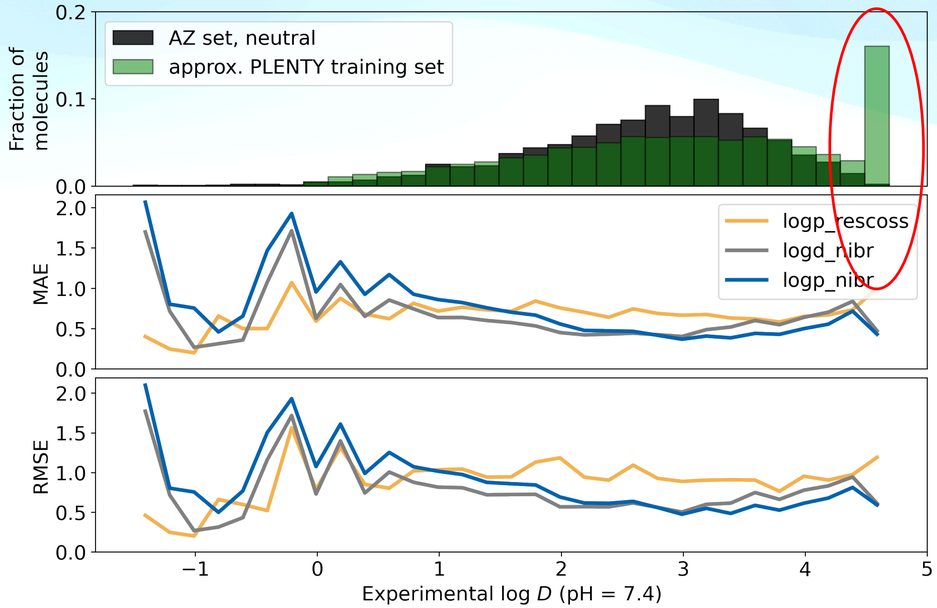
- The literature via ChEMBL
  - Can we trust this data? Could we ever?
    - Not really - Publication bias, no real link to the assay meta-data,…
      - This is not a pop at the ChEMBL folk - they are bravely fighting many dragons
      - How noisy can it be? Very!
        - See Landrum and Riniker, JCIM, 2024, 64, 1560-1567
        - Fine to use chembl if that's all there is. It was certainly great when you had no access to the data like BBB but take real care!
- Curated data
  - PDB, CCDC - but not necessarily with assay data
  - Individual curated datasets - Pat's Rant on their errors - I AGREE!!!
- Pharma datasets
  - Generally high quality measurements and meta-data
    - Access to the people who set up and ran the experiment
    - Some biases due to cost and throughput
  - Hard to publish due to IP concerns

What do we need to build models successfully?

- Domain Knowledge
  - What is the data and what does it mean?
- Data Science and statistics
  - What is the quality of the data and models built from it
- Software engineering
  - How do we share the models
- When we only have two (or even 1) of these 3 skills in the team, model building will fail (but not obviously!)

Domain Knowledge

- Do we understand the source of the observed data?
- What is being measured?
- What are the limits on the measurements?
- What is the experimental error?
- What is the dynamic range?
- Beware of stripes in your data!
    - What to do with qualified data?

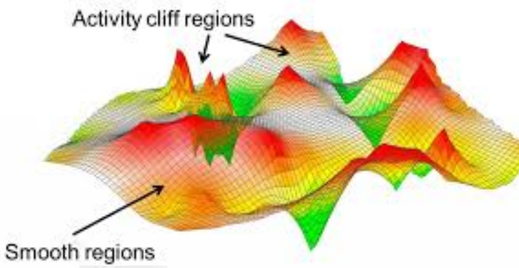


OK, I get it, I need data. Now can I build a model?

Thoughts from Gerry Maggiora from pre 2006…

For many years it has been assumed that similar molecules tend to have similar activities, leading to activity landscapes comparable to the gently rolling hills found on the Kansas prairie. Mounting evidence suggests, however, that this picture is not as universal as once thought but is in many cases rather more like the rugged landscapes of Utah's Bryce Canyon. This new topographical metaphor clearly implies that very similar molecules may in some cases possess very different activities leading to what can be called *activity cliffs*–an activity cliff is defined by the ratio of the difference in activity of two compounds to their “distance” of separation in a given chemical space. The existence of such activity cliffs is not entirely surprising since molecular recognition plays a crucial role in determining activity. For example, a change as “small” as that obtained by replacing an ether oxygen by a secondary amine can have a significant effect on activity.

Moreover, while new computational methodologies such as support vector machines may ameliorate some difficulties and produce better, more predictive QSAR models, they are ultimately constrained by the quality of the data, the number and nature of the compounds in the sample, and, importantly, by the underlying characteristics of the molecular representation that define the chemical space in which the compounds lie.



Gerald. M. Maggiora On Outliers and Activity Cliffs–Why QSAR Often Disappoints. J. Chem. Inf. Model. 2006, 46, 4, 1535

…and confessions from John Dearden (>1000 QSAR papers?)

Dearden, J. C., Cronin, M. T. D., & Kaiser, K. L. E. (2009). How not to develop a quantitative structure–activity or structure–property relationship (QSAR/QSPR). *SAR and QSAR in Environmental Research*, *20*(3–4), 241–266.

OECD principles

1. a defined endpoint;
2. an unambiguous algorithm;
3. a defined domain of applicability;
4. appropriate measures of goodness of fit, robustness and predictivity;
5. a mechanistic interpretation, if possible.

Type of error	OECD principle
Failure to take account of data heterogeneity	1
Use of inappropriate endpoint data	1
Use of collinear descriptors	2 4 5
Use of incomprehensible descriptors	2 5
Error in descriptor values	2
Poor transferability of QSAR/QSPR	3
Inadequate/undefined applicability domain	3
Unacknowledged omission of data points	3
Use of inadequate data	3
Replication of compounds in dataset	3
Too narrow a range of endpoint values	4
Over-fitting of data	4
Use of excessive numbers of descriptors in a QSAR/	4
Lack of/inadequate statistics	4
Incorrect calculation	4
Lack of descriptor auto-scaling	4
Misuse/misinterpretation of statistics	4
No consideration of distribution of residuals	4
Inadequate training/test set selection	4
Failure to validate a QSAR/QSPR correctly	4
Lack of mechanistic interpretation	5

The mistakes we made

There is nothing new under the sun.

- We didn’t really curate the data before starting
  - Noise, reproducibility, quality
  - Cleaning of structures for false positives
- We didn’t check on the roughness of the landscape
- Qualified data? Meh. Data was just accurate.
- We built a model and used it
    - q^2 was a novelty at the time
- Are we sure we are not making the same mistakes more subtly?

An experiment

- Everyone stand up
- Sit down if you have never used Rule-of-5
- Sit down if you have used Rule-of-5 but never read the paper
- Sit down if you think that all compounds that fall inside Rule-of-5 are bioavailable
- Sit down if you think the rules still apply
- And here's the paper (over 31k citations)
    - Experimental and computational approaches to estimate solubility and permeability in dru and development settings. CA Lipinski, F Lombardo, BW Dominy, PJ Feeney Adv. Drug D (1-3), 3

Lipinski's 90% rule of 5 for oral bioavailability based on compounds with USAN numbers in 1997

- Easy to understand, easy to encode, easy to screw up.
  - Does not apply to natural products, compounds with Internal H-bonds
    - So no use for cyclic peptides, RLTs, PROTACs
  - cLogP values are not accurate like MW!
    - And we should be using cLogD anyway
  - Is the interaction of a carbonyl with solvent the same as an ether, or pyridine the same as piperidine?
- It was a great step forward in the days of Comb Chem
- Implicitly believed although the model does not revalidate
    - Soares et al. https://doi.org/10.1002/minf.202300115
    - Peter Kenny: http://fbdd-lit.blogspot.com/2024/07/a-nobel-for-property-based-drug-design.html
- " The bias against Ro5 non-compliant compounds in most HTS screening libraries is likely one factor in the failure of almost all HTS campaigns to discovery viable PPI inhibitor leads" – Lipinski, 2016, Adv Drug Del Rev., 101, 34

A model is a property of the data set

No universal truth

- What are the best practices?
  - See the Dearden list
  - Use MoleculeACE to check for cliffs (or a best practices toolkit?)
    - Francesca Grisoni et al., JCIM, 2022, 62, 5938
    - Wognum et al. A call for an industry-led initiative to critically assess machine learning for real-world drug discovery. Nat Mach Intell 6, 1120–1121 (2024).
- Build some simple models based on sensible descriptors
  - These can act as baselines
  - Look at learning curves to check for coverage
- What does success look like?
  - Docking < 2.0A?
    - Erickson J.A., Jalaie M., Robertson D.H., Lewis R.A. & Vieth M. (2004) Lessons in Molecular Recognition: The Effects of Ligand and Protein Flexibility on Molecular Docking Accuracy. J.Med.Chem., 47(1), 45-55.
    - RMSD is easy but possibly wrong
    - A docking should really be assessed on fit to electron density

Thoughts on literature data from an ex E-i-C

Some of the dragons ChEMBL is fighting

www.ebi.ac.uk/training/online/courses/chembl-quick-tour/what-is-chembl/how-is-chembl-data-curated/

- Rewards come to those who publish
  - Do *just* enough to satisfy the reviewers
    - Publication bias for positive data
    - Negative data or inability to replicate is not published
  - Corrections/updates are very rarely done and if so have a new doi that is not forward-linked from the original
  - Data is not checked (e.g. reaction yields, IC50's). Just assumed to be right.
  - Requirements for data discourage Pharma publishing models (loss of IP)
  - Inappropriate safeguards (e.g. PAINS)
  - Publishing model is broken - all for profit

there is a growing mountain of research. But there is increased evidence that we are being bogged down today as specialization extends. The investigator is staggered by the findings and conclusions of thousands of other workers - Vannevar Bush

Faux News? Another source of error in activities

- Docking models have been shown to have an enhanced hit rate for finding actives
- Therefore docking scores are accurate ranking methods and a score can be quoted as a binding energy and converted to an IC_{50}
  - Sars-COV2 is a bad thing
  - “we docked a large number of compounds from some source or other against mPro, took the top 100, applied some ADME filters, maybe MMGBSA, and the best scoring molecule has an affinity of -11.7356 kcal/mol or IC50 of 2.445 nM and is worth testing in the clinic”
- Just in 2024 I received 331 papers, rejected 314 papers with only 17 sent for external review. Some got through in other journals and have polluted the knowledge base
- Sadly not everything in the literature is ‘true’ or consistent

How can we do better?

- By all means publish, but create a Github repo too
    - The paper is a fixed reviewed reference but code/data can be checked and updated by the authors and community
    - Models can be rebuilt on updated/local data or used as is.
    - Upticks for good repos
    - But we all have to do some work to help maintain/improve the repos
- When it works, you get something for the whole community

Sharing without revealing

How to improve models without IP issues

- Pharma have some of the ‘best’ data sets. They also need to protect IP.
- How can we share them?
    - Between consenting groups, it is just a matter of legal agreements!
        - Encrypting - MELLODY (Ceulemans et al. J Chem Inf Model 2024 Apr 8;64(7):2331-2344.)
            - AISB - John Karanicolas
        - Honest broker - SimPlus pKa model (Fraczkiewicz et al.10.1002/minf.202400088
        - Anonymising - MMPA (Griffen et al. doi:10.1016/j.drudis.2018.03.011)
- But none of this helps academics
    - Journals should cut Pharma some slack
    - Let Pharma publish models without revealing in-house data, perhaps use Chembl for validation, as long as the model is in a repo
    - If you are using a model for GenChem, you get a ‘better’ model

Validation - We need blinded competitions!

Not just for docking and ML, but for all methods. polarishub.io

- This idea arose directly from the last GRC
    - Kramer et al. The Need for continuing blinded Pose- and Activity Prediction Benchmarks J. Chem. Inf. Model. 2025 65 (5), 2180-2190
- We cannot really assess any method without more blinded data
    - Docking, ML, FEP, conformers in solution, GenChem
    - Benchmark of how well we are doing
    - Afterwards, unblinding helps to find areas for improvement
- The buck stops with Pharma to provide the data
    - It is not easy but it can be done.

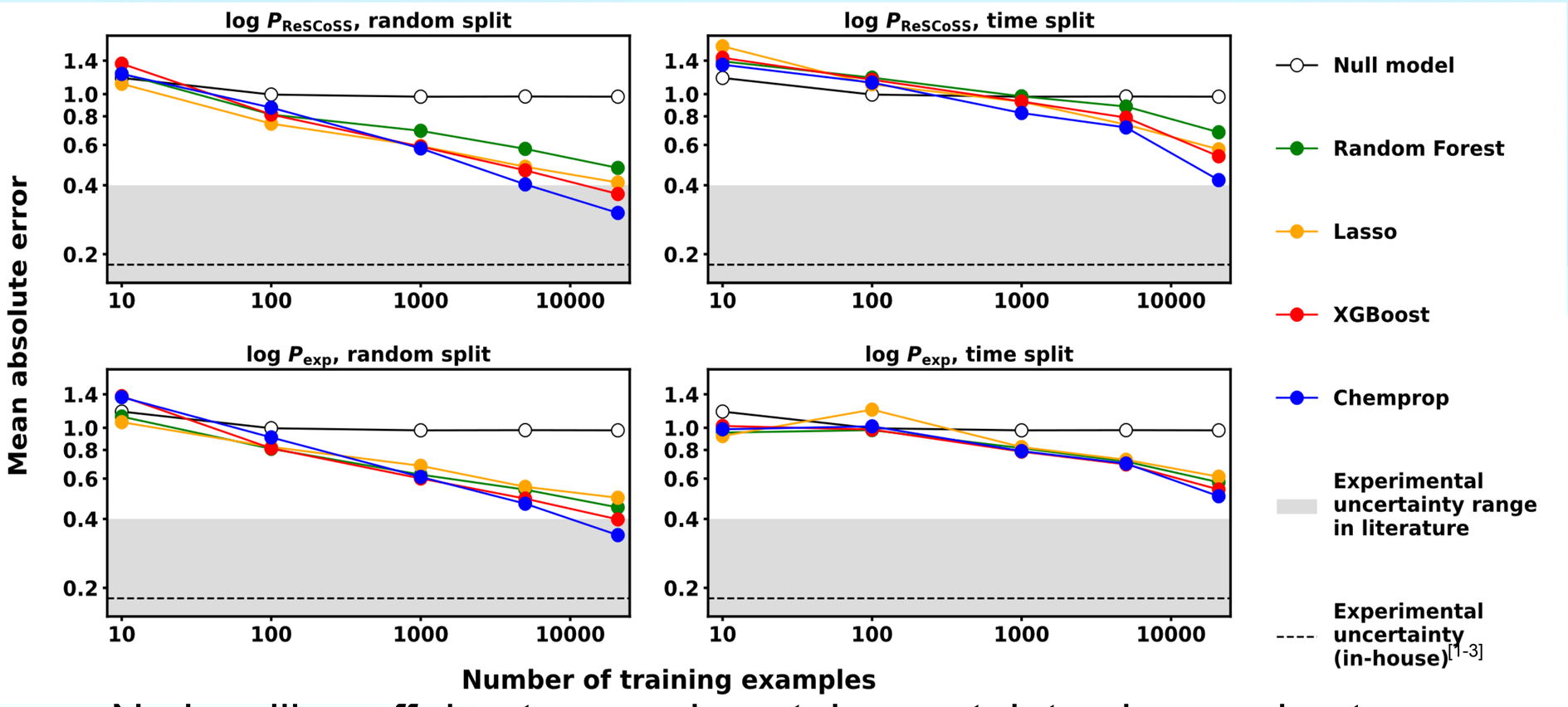
        “With orders signed in triplicate, sent in, sent back, queried, lost, found, subjected to public inquiry, lost again, and finally buried in soft peat for three months and recycled as firelighters.” Thanks D. Adams
    - It does require giving up IP
    - Without this, are we really doing science?

A modest proposal (great minds think alike)

Not as good as a blinded competition but gets over IP

- Using Pharma data sets as validation material
  - Create a web-based register of datasets with meta-data about the assay
  - Create a best-practise toolkit for assessing a model's performance
  - Users can select the 'best' assay to validate their model
    - User supplies their model (and data?) as a docker image plus instructions and submits it
      - It must run without further customisation effort by the supplier using standardised i/o formats
    - Supplier runs image against their data in a sandbox
  - Supplier returns only aggregated statistics
    - Average similarity to train set, performance stats, learning curves, time splits…
  - User gets referenced validation statistics that can be used to support publication
  - Supplier gets to keep the model and use it (no licenses or purchase needed if it is subsequently commercialised)

Are we there yet? Learning curves



- No levelling off due to experimental uncertainty observed yet

⟹ **Larger training set is expected to decrease prediction error further**

Richard Lewis

[1] I. Aliagas, A. Gobbi, M. Ling, L. Benjamin, *J. Comp. Aided. Mol. Des.* **2022**, doi:10.1007/s10822-022-00450-9.
[2] N. Ulrich, K. U. Goss, A. Eberg, *Commun. Chem.* **2021**, *4*, 1-10.
[3] C. C. Bannan, G. Calabro, D. Y. Kyu, D. L. Mobley, *J. Chem. Theory Comput.* **2016**, *12*, 4015-4024.

So I have built my model, so I can just publish it?

- In the old days, that is just what happened (see confessions of John Dearden)
  - Docking models with no protein preparation
- We need a best-practices toolkit to validate models
  - Is the data sound?
  - Is the bleed from validation into training data?
  - Is the model robust to different splits?
  - Has the model overfitted?
  - And many others…
  - It is all there but not in one place

GenChem/Ultralarge databases

Model destroyers

- The size of chemical space that can be explored will amplify any flaw in the scoring function
- Are the structures that are generated/highly scored sensible in terms of the scoring function?
  - Not just the function in the code, but a human-interpretable one.
  - Quantification of uncertainty and confidence
    - Synergy conformal prediction applied to large-scale bioactivity datasets and in federated learning. U Norinder, O Spjuth, F Svensson, Journal of Cheminformatics 13 (1), 77
    - Unique: A framework for uncertainty quantification benchmarking. J Lanini, MTD Huynh, G Scebba, N Schneider, R Rodríguez-Pérez, JCIM 64 (22), 8379-8386
    - Poster by Stefanie Kickinger

And a recent post by Ash Jogalekar

1. Did the team test the model prospectively - i.e., on brand-new compounds that were then made and assayed? Validating a model on real-world molecules is the only proof the model isn't just memorizing history.
2. How was the data split, and are the test molecules chemically different from the training ones?
3. Was data leakage ruled out?
4. Did they compare against the current standard methods on the exact same data and metrics?
5. Which evaluation metric was used, and does it match the decision you'll make?
6. Do they provide an uncertainty estimate for each prediction?
7. Is the model's domain of applicability clearly defined?
8. Can they show why the model makes a prediction (interpretability)?
9. Is the data origin transparent?
10. Is the whole workflow reproducible—code, random seeds, parameter files, and (when IP allows) the data set?

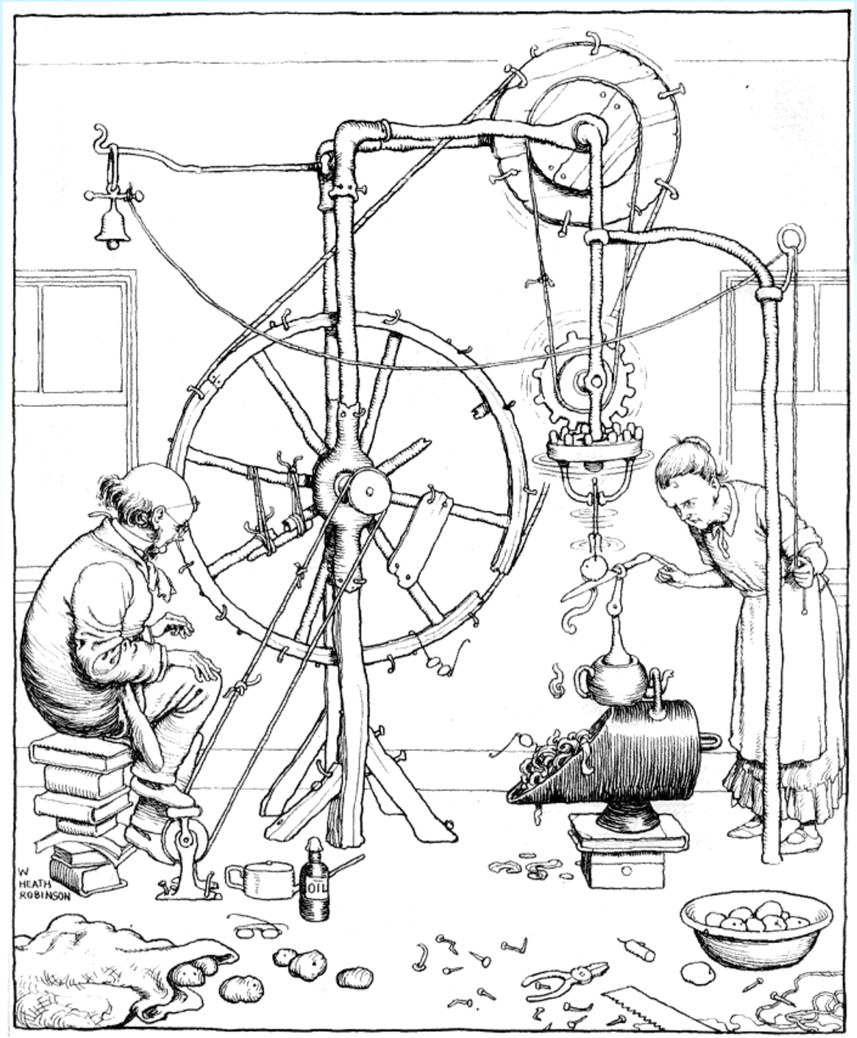
https://www.linkedin.com/feed/update/urn:li:activity:7337632298718195712/

Sharing your results - Software Engineering

- Can the model be simply deployed?
- To whatever platform consumes it
- Is it scalable (to millions, billions,...)
  - Ultra large databases (Enamine)
  - Output from Generative Chemistry runs
  - Is the model versionable?
    - When new data is available
  - Is the code base well-documented, robust and easily maintainable?
  - Is the interface and results what the users need and like?
    - Well defined endpoints that can be easily consumed
- It is easy to get a forest of poorly documented models, given the ease of construction

Applying models

When the rubber hits the road



The Professor's invention for peeling potatoes.

Richard Lewis

De Novo design and AI

A grand failure (by me)

- Molecules were built in 3D but…
- Poor scoring functions
  - Too driven by making interactions (hydrophobicity, h-bonds)
    - Role of water ignored
    - Do we still have an unconscious bias towards DH?
  - For every atom/group added you had to retype its neighbours
  - Synthetic nightmares, ugly compounds
- I may have even contributed to the first AI winter with this paper - Sorry!
  - Drug design by machine learning: The use of inductive logic programming to model the structure-activity relationships of trimethoprim analogues binding to dihydrofolate reductase RD King, S Muggleton, RA Lewis, MJ Sternberg 1992 PNAS 89 (23), 11322-11326

MPOs

- MPOs reflect all the conflicting desires that control selection or generation
- Combinatorial chemistry was ideal for this
  - Andrew C. Good and Richard A. Lewis New Methodology for Profiling Combinatorial Libraries and Screening Sets: Cleaning Up the Design Process with HARPick JMC 1997 40 (24), 3926-3936
- What I didn’t do well
  - Coded everything in F77 and ChemX
  - Did not use a Pareto front to balance desires
- I think this is still an area for development
  - Immediate value (high score) vs future value (improve model)
- We should be proposing sets of molecules to be tested, not just the ‘best’

Automated Iterative Design (or GenChem Exploit)

- Idea: modify your lead using SMIRKS then score
  - Lewis R.A. (2005) A General Method for Exploiting QSAR Models in Lead Optimization. J. Med. Chem., 48(5), 1638-1648.
- The pitfalls:
  - What happens if the transforms introduce substructures outside of the model?
  - What happens if the model can be exploited to improve the score but do the wrong thing
    - example change every Me to Cl because cLogP is a +ve contributor
  - Need to explicitly encode guardrails for applicability; this should be implicit

GenChem now

- The molecule generation algorithms are so much better than previous approaches
  - Guacamol to check the generators
  - More ‘drug-like’ structures
    - But there are still biases from the data

      Understanding Our Love Affair with p-Chlorophenyl: Present Day Implications from Historical Biases of Reagent Selection. D. Brown, M. Gagnon and J. Boström J. Med. Chem. 2015, 58, 5, 2390-2405
    - Need a good objective function/target property profile with a defined domain of applicability

      Weighting of terms

      Understand when the objective is ‘frustrated’ - activity vs solubility etc

      Are we making the same mistakes as before but faster?
- What is missing are really good links between ML and ‘physics-based’ modelling
  - Generation is easy, scoring is hard - Attributed Chris Murray
- Synthetic accessibility
  - Cheap Cheat - use similar compounds from Enamine to test ideas
  - SPARROW from Jenna Fromer

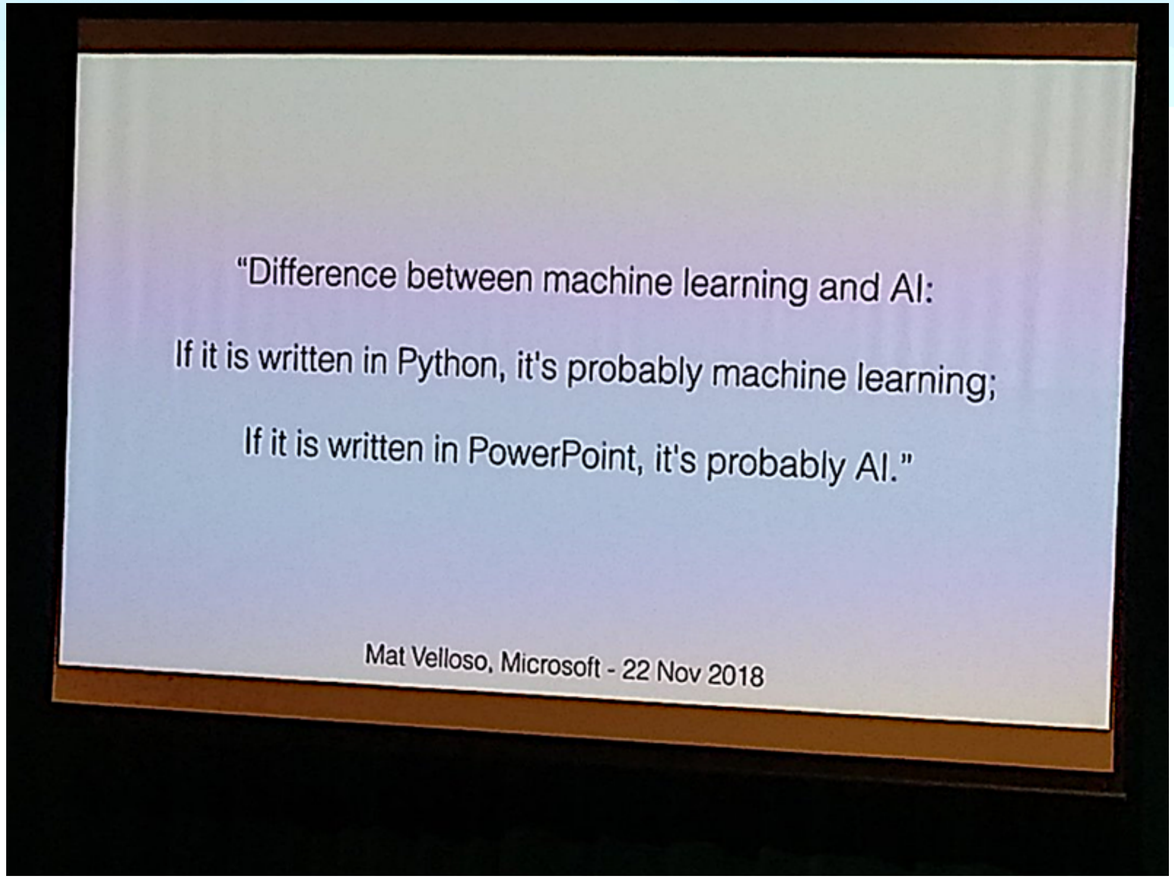
Turning to physics

- New modalities are coming in
- The assumption that properties are independent of conformation breaks down
  - Macrocycles, degraders, RLTs, anything with high cLogP…
    - Traditional additive-based methods don’t work!
  - QM is the only route to handle these - Christopher Tautermann
  - Can I get a quick but good Boltzmann population of conformers in a solvent of choice?
    - Can I use this to predict cell penetration?
    - What is the charge on an RLT?
    - Can I compute properties/spectra reliably and quickly
    - Can I build useful ML models of the QM data?

Building Models - Data Science

- Or AI/ML as it is called in management presentations and grant applications

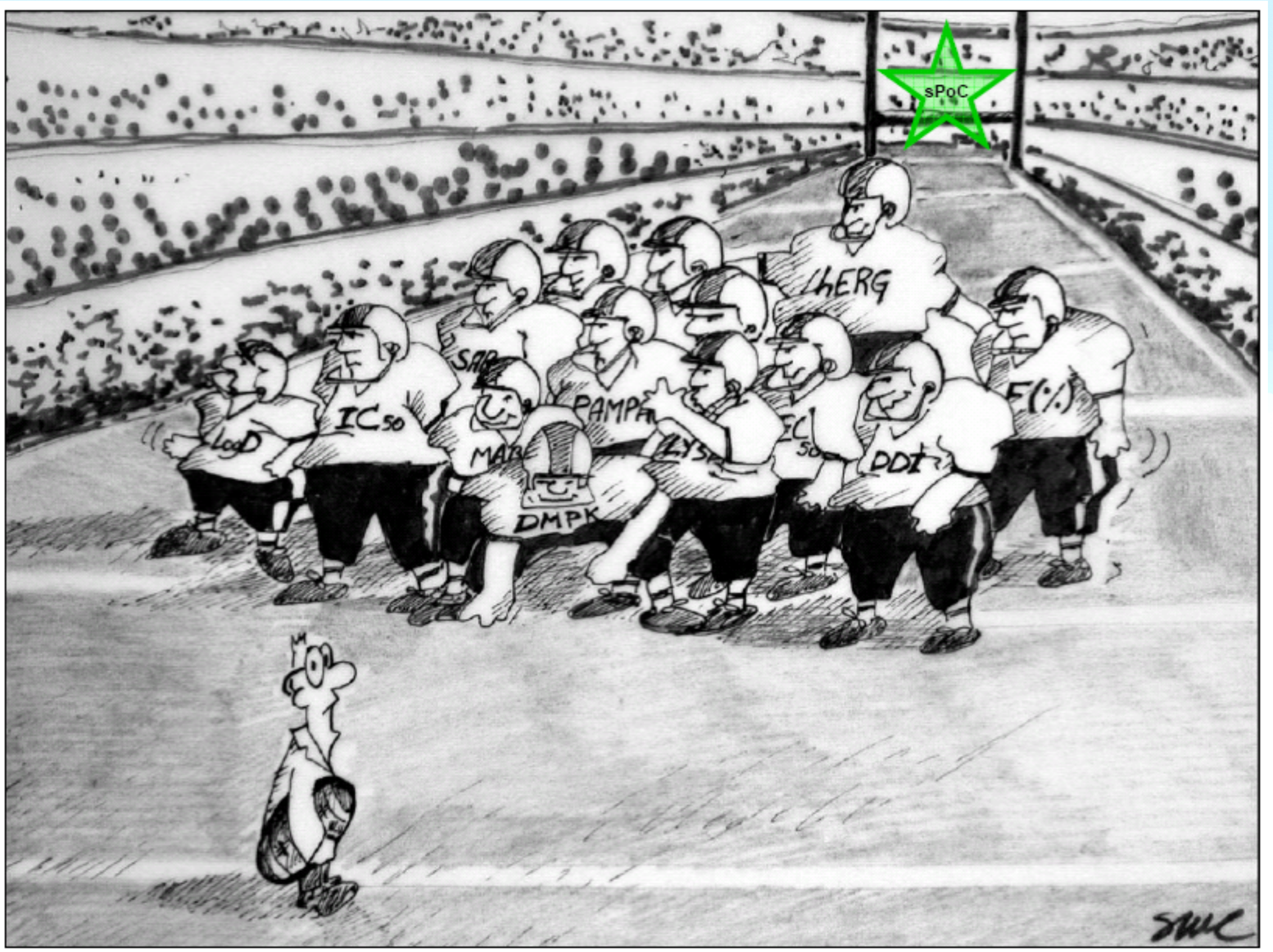
- The mechanics of building models are fairly trivial
- Jupyter notebooks exist to do just this
  - patwalters.github.io/practicalcheminformatics/jupyter/ml/interpretability/2021/06/03/interpretable.html
  - Others are available!
    - Even for chemprop (https://monkey-mind.blog/2025/03/15/chemprop-v2/)



The pitfalls in CADD are well known but we are still falling into them

- Better resources mean that we have more opportunities to fall in too!
  - FEP runs overnight
  - Docking of billions in a week
  - GenChem exploration
- We have to ensure the foundations are secure.
- We need to make models more interpretable
  - For sanity checking by humans, if possible
- We need to quote results to the correct number of significant figures
- We need to solve the issues that cause ligands to fail to progress as drugs

Richard Lewis



Richard Lewis

Take-away messages for the future

- Can I have an open toolkit for checking data and models?
    - Activity cliffs, quality checks, behaviour under noise, correct statistics, learning curves, test/train split methods, data bleeding

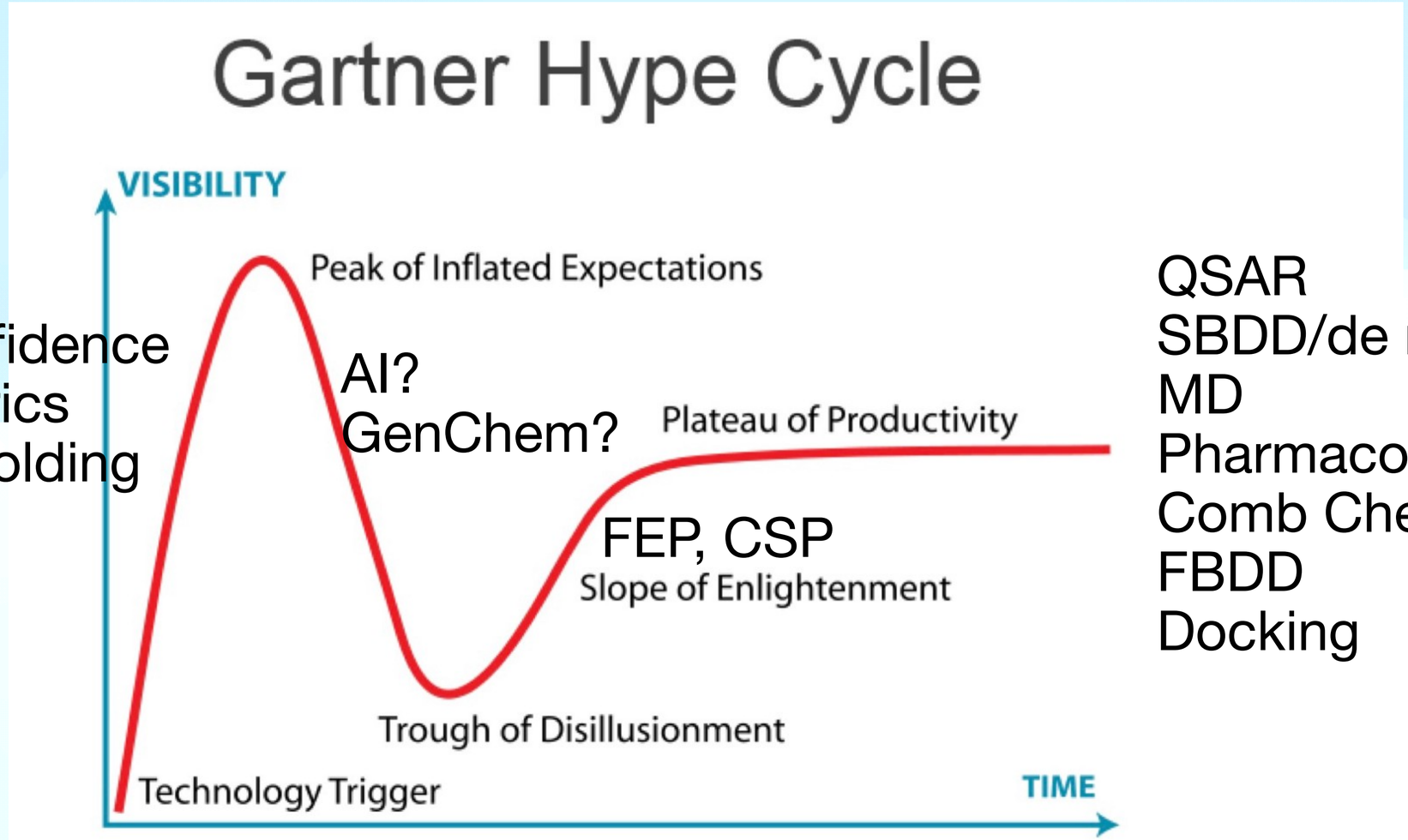
        Much of this is already available, it just needs to be assembled into a single place
- Can we have talks about validation of methods by independent teams (See Eva's talk)
- Can we have confidence and uncertainty measures for any prediction?
- Can we have good prediction of PChem, PK & Tox properties from molecular structure?
    - everything from solubility to bioavailability to idiosyncratic in vivo toxicity
    - And for larger structures where additive methods break down

        RLT, macrocycles, PROTACs
- Can we have methods that focus more on the key data, not just all of it?
    - pKa between 5.5 and 8.5
    - The 'active' molecules
- Can we have robust methods for interpreting ALL models?
- How can we share more?
    - **How do we share models without necessarily sharing data?**
    - **How do we share data so that it can be trusted?**

And more…

- Can we get accurate Boltzmann population in the relevant environment (quickly and reliably)
  - With confidence metrics for sampling and energies
  - And for complex systems (RLT, macrocycles, PROTACs)
- Can we understand more about when Cofolding, BFE by ML, ABFE and RBFE will work and when not (quickly and reliably)?
  - What are the limitations? When do they kick in?
- How can we fuse ML, QM, 3D-QSAR and RBFE?
  - QuanSA, Cresset,…
- Can we model reaction mechanisms reliably (covalent, Cyps, enzymes)
- Can we do crystal structure prediction (not as slowly)
- Water and entropy - Do we have an unconscious bias to enthalpy through docking?
- When is a cryptic pocket really a cryptic pocket? 😆

The enemy of the good



Confidence
Metrics
CoFolding

AI?
GenChem?

FEP, CSP
Slope of Enlightenment

QSAR
SBDD/de novo
MD
Pharmacophores
Comb Chem
FBDD
Docking

Acknowledgments over my career

There are many more, these are just some of people I have published with

- RPR
  - Jon Mason, Iain McLay, Andy Good, David Clark, Stephen Pickett, Paul Bamborough
- Eli Lilly
  - Jim Wikel, Ian Watson, Howard Broughton, Andrea Zaliani, Michael Vieth, Mike Bodkin, Jon Erickson
- Novartis
  - Pascal Furet, Peter Ertl, Nik Stiefl, Rainer Wilcken, Greg Landrum, Peter Gedeck, Ansgar Schuffenhauer, Nadine Schneider, Anna Vulpetti, Janneke Jansen, Eric Martin, Arkadij Kummer, Eric Ma, Christian Kramer, Niko Fechner, Jimmy Kromann, Grahame Woollam, Peter Hunt, Jessica Lanini, Gianluca Santarossa, Finton Sirockin, Bernard Pirard, Rajeshri Karki, Rishi Gupta, Stephen Chan, Rachel Rodrigues-Perez, Gregori Gerebtzoff, Jessica Lanini, Hagen Muenkler, Nathan Ricke
- Academia/Collaborators
  - Philip Dean, Tack Kuntz, Mike Sternberg, Peter Gillett, Val Gillet, Brain Shoichet, Andrew Leach, Jon Hirst, Clemens Isert, Nathan Brown, Elaine Meng, Jon Essex, Mike Gilson, Gisbert Schneider, Markus Neumann

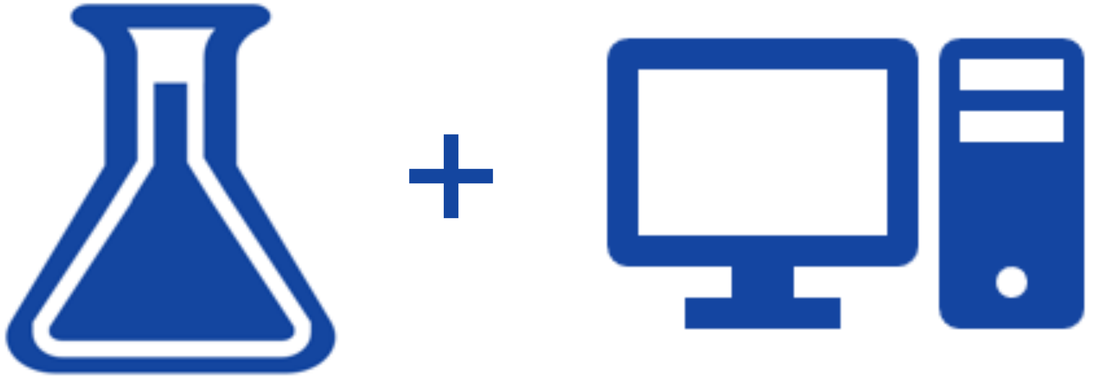
Three things cannot be long hidden: the sun, the moon, and the truth

Attr. Buddha

- St Augustin of Hippo:
  - Hope has two beautiful daughters: Their names are anger and courage. Anger at the way things are, and courage to see that they do not remain the way they are
- Desmond Tutu
  - If you are neutral in situations of injustice, you have chosen the side of the oppressor
- What can we do to help American science?
  - Build on the GRC CADD community

EFMC goes Computational

EFMC²



Working Group

MedChem CompChem

EFMC2 Tandem Talks (> ~200 attendees)

EFMC² Tandem Talks

3rd EFMC² Tandem Talks

Virtual Event May 28, 2024

- https://www.efmctandemtalks.org/

So long and thanks for all the fish!

Douglas Adams