optibrium

Can We Really do Computer-Aided Drug *Design*?

ACS Spring National Meeting, March 29th 2012

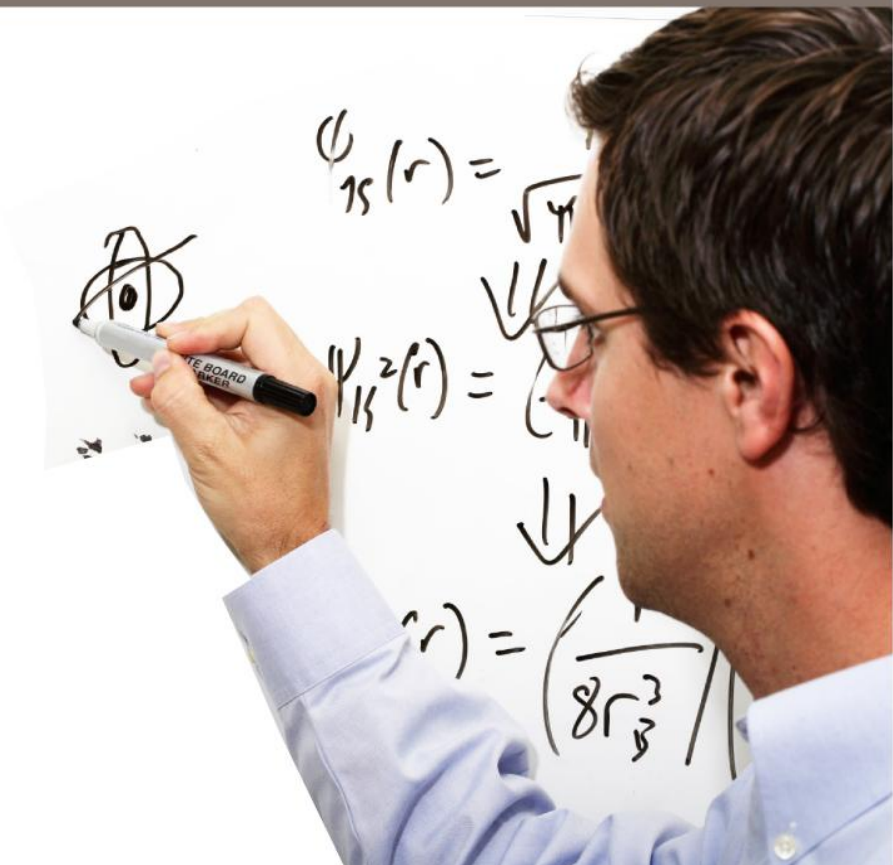
Matt Segall

Overview

- Design vs. Discovery
- Accuracy of predictive models in drug discovery
- How accurate do models need to be?
- Adding value with predictive models
- Moving toward drug *design*
- Conclusion

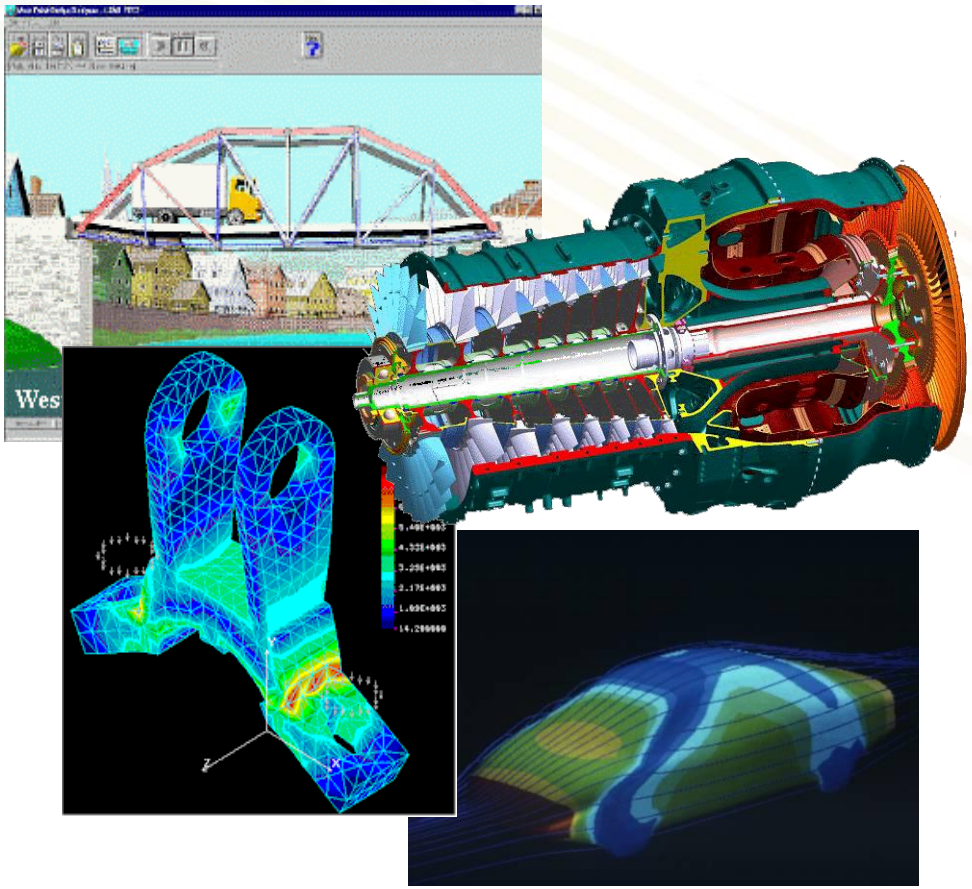
Design vs. Discovery

optibrium

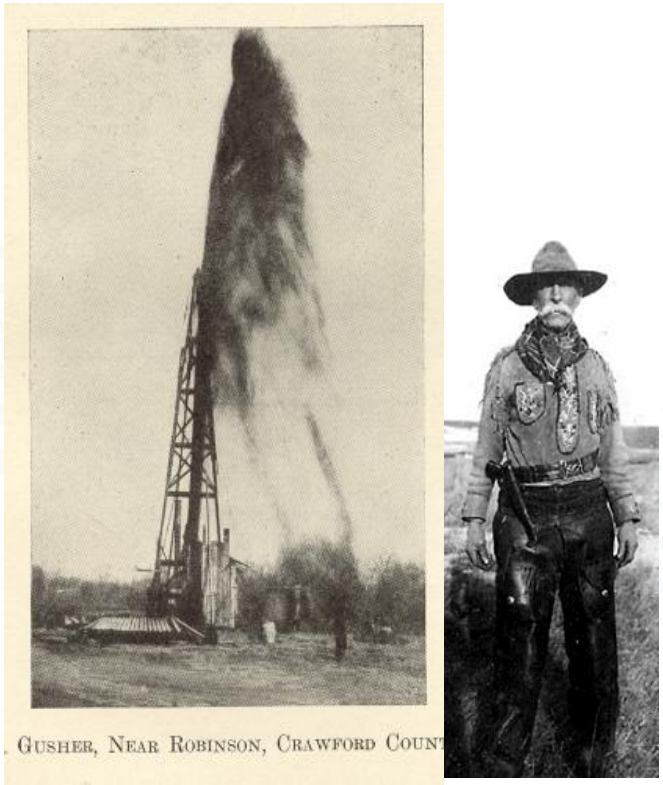


Design vs. Discovery

Design



Discovery



An Analogy of Drug Design
The Boeing 777*

* Selick *et al.* Drug Disc. Today, **7**, pp. 109-116 (2002)

- Designed entirely on computer
- Full-scale prototype built
- Successfully flown first time
- Compared with the “crash test” paradigm of drug discovery

Why Does this Analogy Break Down?
Complexity of Design Goals?

Airplane	Drug
• Cost	• Potency
• Efficiency	• Selectivity
• Range	• Absorption
• Capacity	• Metabolic Stability
• Safety	• Safety
• …	• ….

Why Does this Analogy Break Down?
Precision of Models

Airplane

$$\frac{D\rho}{Dt} + \rho\nabla\cdot\mathbf{v} = 0$$

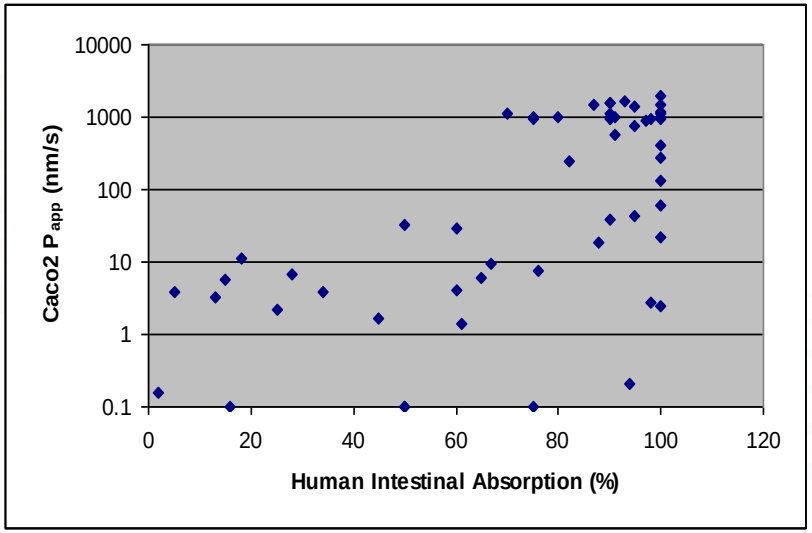
$$\rho\frac{D\mathbf{v}}{Dt} = -\nabla\rho + \nabla\cdot\mathbf{T}^* + \rho\mathbf{f}$$

$$\rho\frac{De}{Dt} = -\rho\nabla\cdot\mathbf{v} + \Phi - \nabla\cdot\mathbf{q} + \rho r$$

Drug

Caco2 vs. Human Intestinal Absorption*

$$y \approx a_1 x + b_2 x^2 + a_4 x^4 + \ldots$$

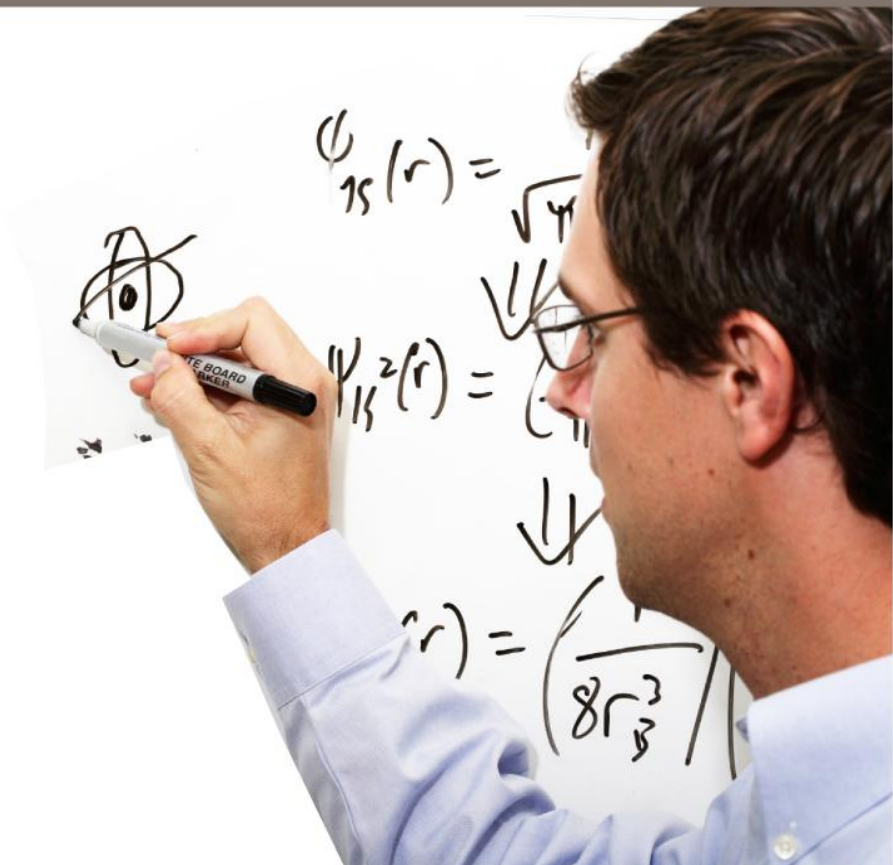


* Irvine *et al.* J. Pharm. Sci. 1999, 88, 28

R^2=0.81, RMSE=0.8 log units

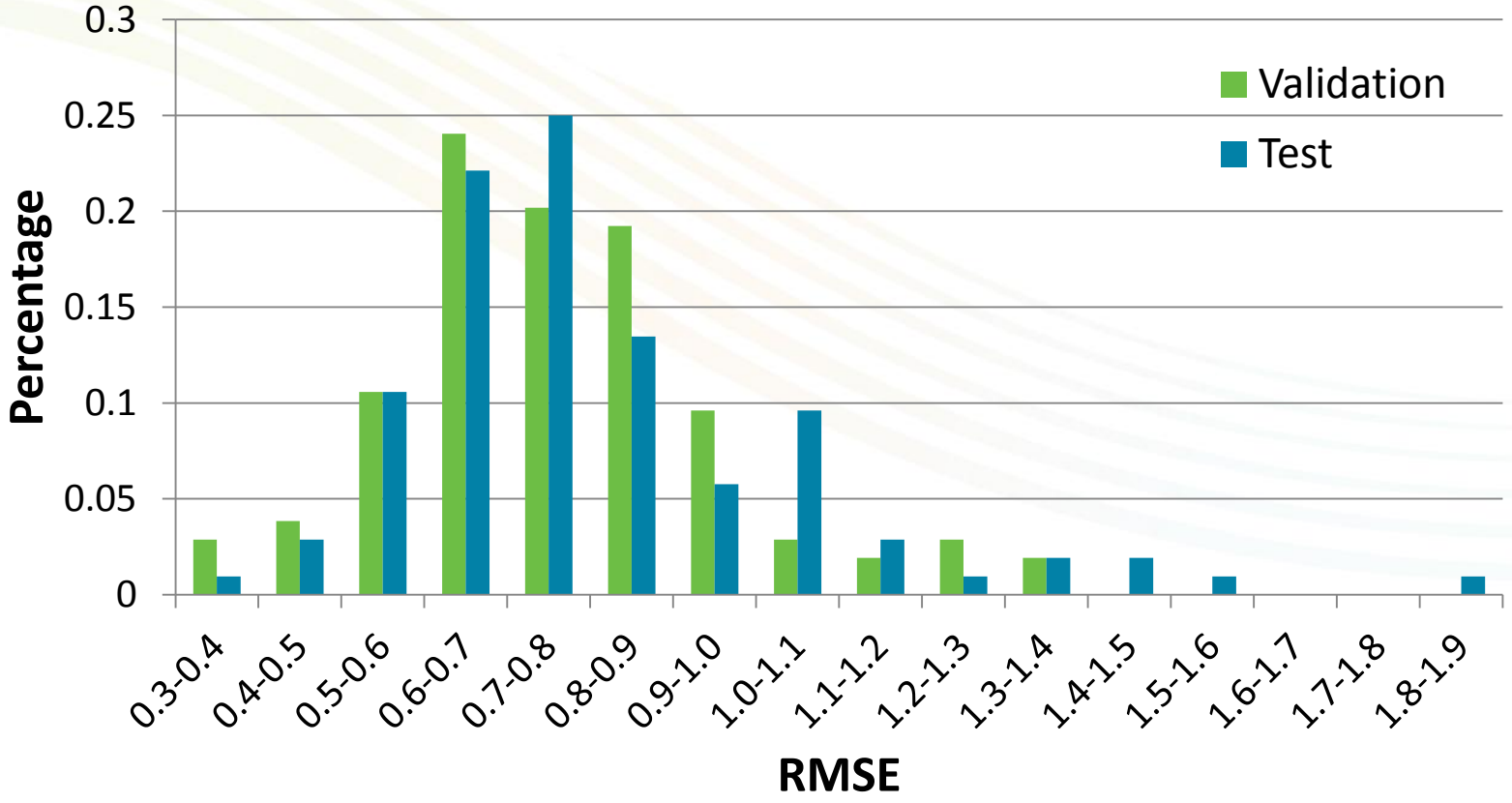
How Accurate are Predictive Models?

optibrium



2D QSAR Models of Target Potency*
Root Mean Square Error

- Average RMSE on validation set = 0.76 log units (factor of 5.8)
- Average RMSE on test set = 0.8 log units (factor of 6.3)



* Segall *et al.* ACS Spring National Meeting, 2012 COMP Thursday 2pm, Room 28E

Other Methods
Some examples

- 2D and 3D Similarity
  - Hit-rate of 20-30% among most similar compounds*
- Docking
  - Similar hit-rate, 20-30% †
- Structure-property relationships
  - Solubility models found to have RMSE of between 0.47 to 1.96 log units on 122 drugs‡

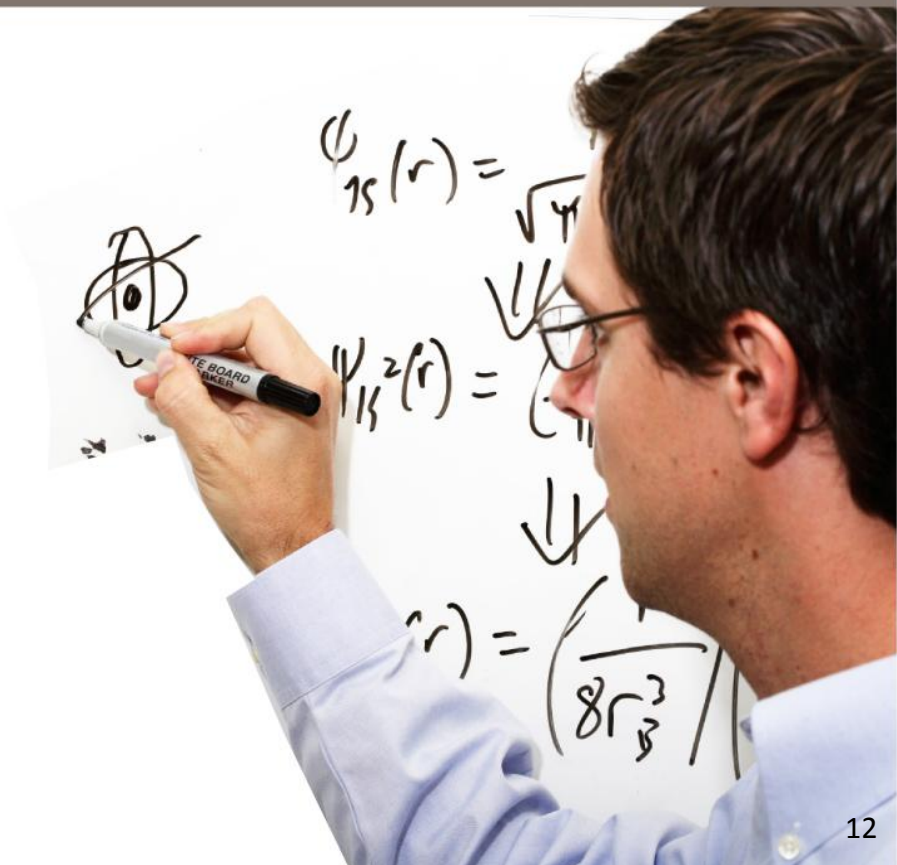
* Bender *et al.* (2005) J. Chem. Inf. Model. 45:1369-1375.

† Kroemer RT. (2007) Curr. Protein Pept. Sci. 8:312-328

‡ Dearden (2006) Expt. Opin. Drug Discov. 1:31-52.

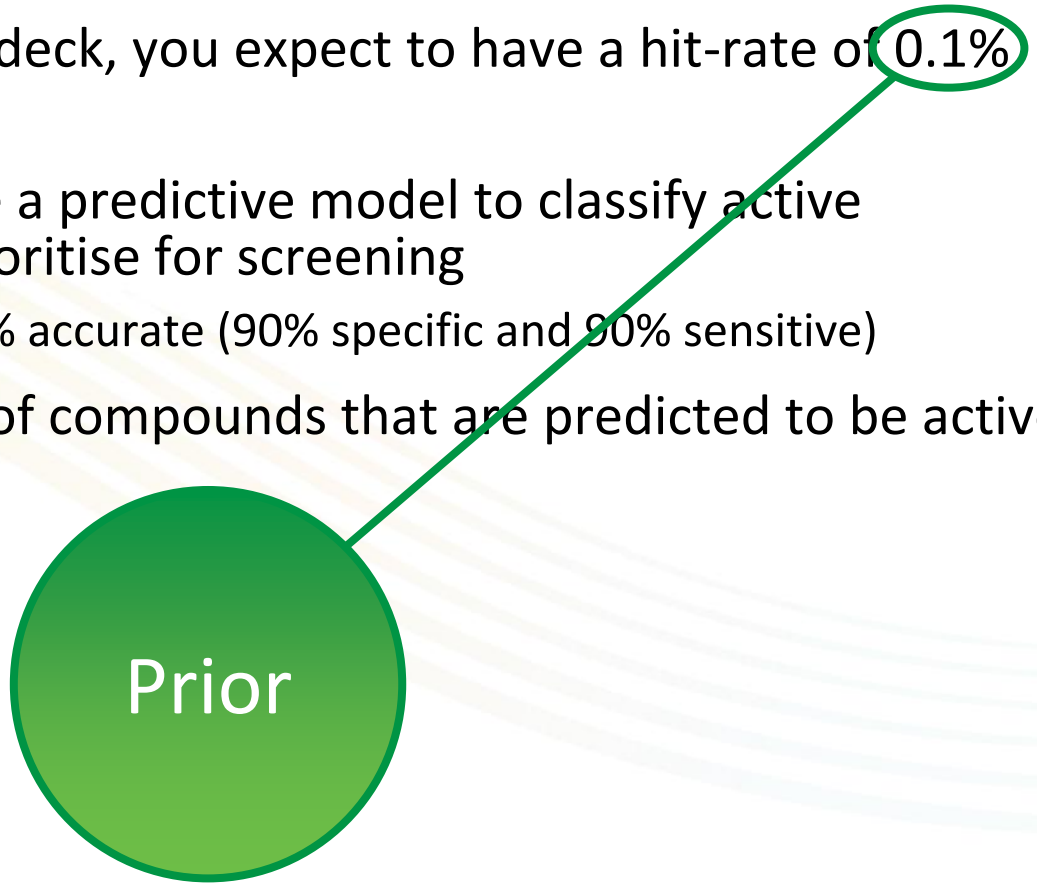
How Accurate Do Models Need to Be?

optibrium



How Well Does this Model Help Us to Identify Active Compounds?

- In your screening deck, you expect to have a hit-rate of 0.1% against a target
- You choose to use a predictive model to classify active compounds to prioritise for screening
  - The model is 90% accurate (90% specific and 90% sensitive)
- What proportion of compounds that are predicted to be active actually are?
  - a) about 0.1%
  - b) about 1%
  - c) about 10%
  - d) about 50%
  - e) about 90%
- Answer: **b)**
  - E.g. Of 10,000 compounds 9990 x 0.1 + 10 x 0.9 = 1008 would be predicted as active, of which only 9 really are.



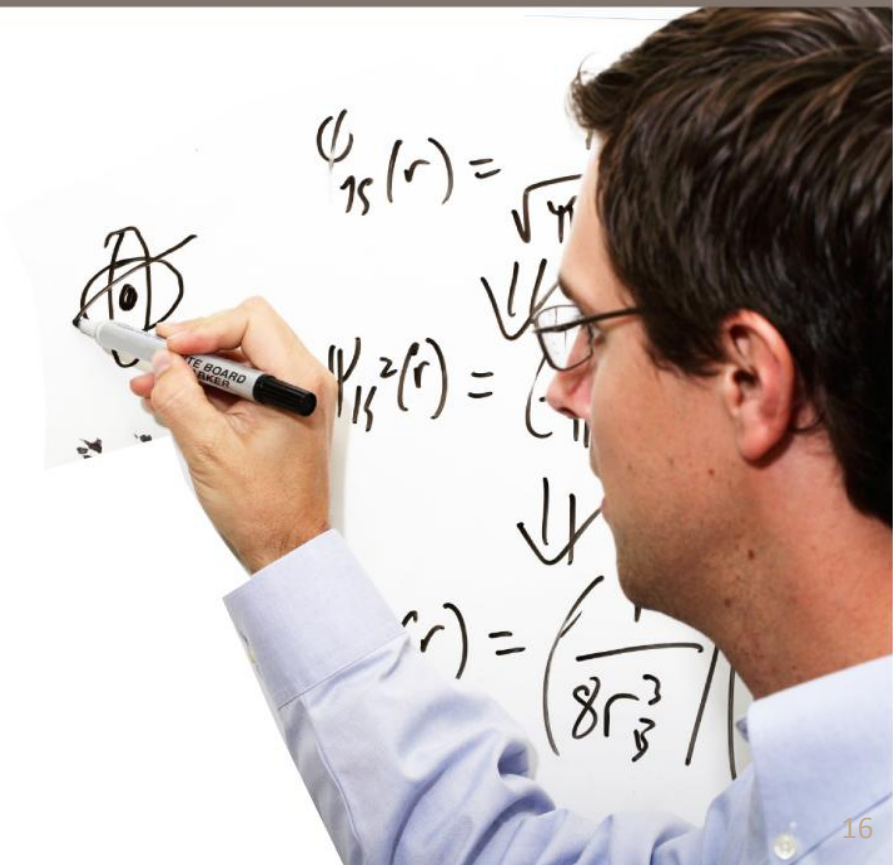
What Prior Probability Do We Need for a 90% Accurate Model to be Useful?

- Depends on what we mean by useful!
  - E.g. 1 in 10 compounds predicted to be active would be expected to be confirmed
- Answer: 1.2%
- Required accuracy depends on the prior probability
  - Until we know this, we don't know the accuracy we require

Sequential Filtering
Compounding errors

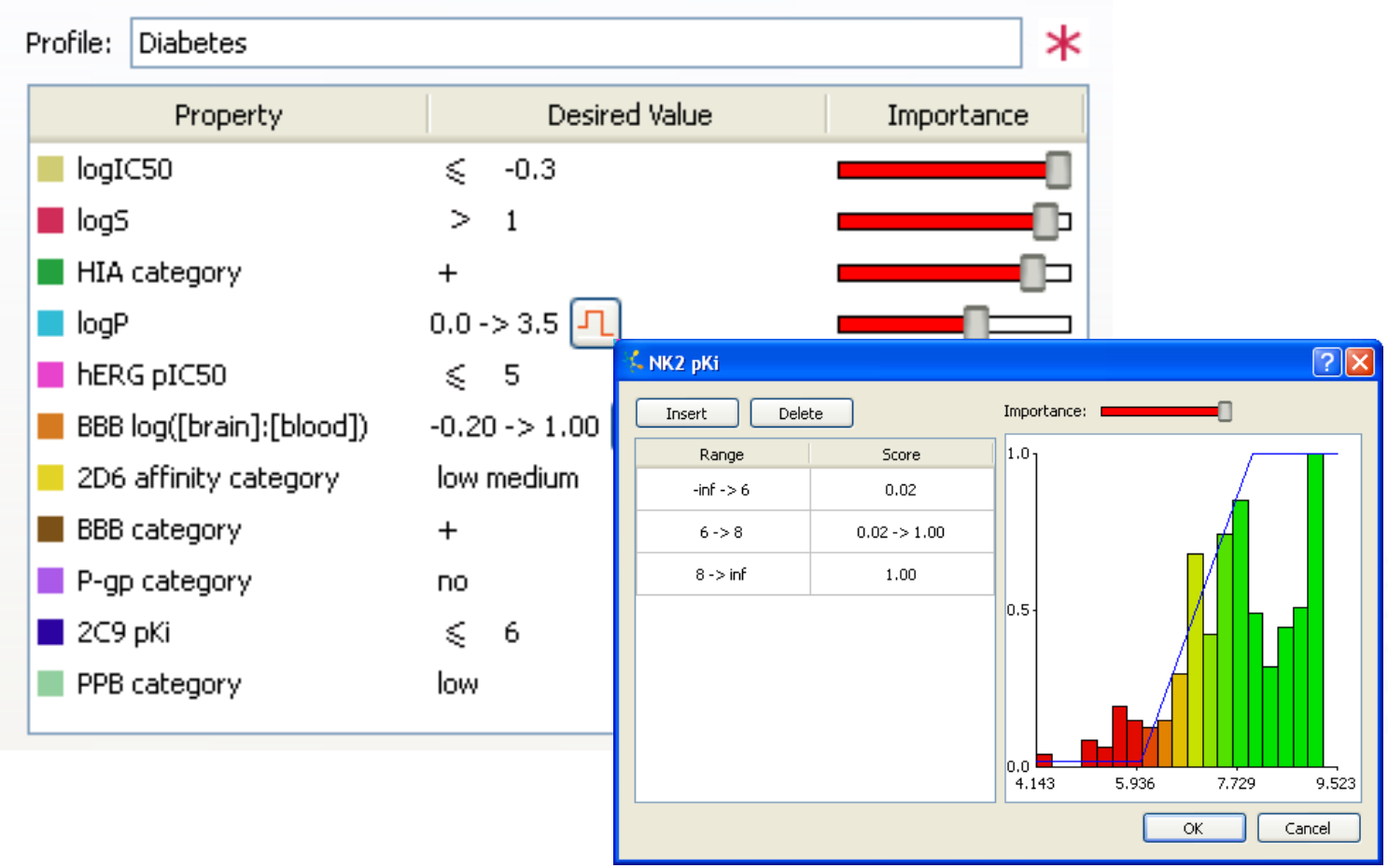
Adding Value With Predictive Models

optibrium



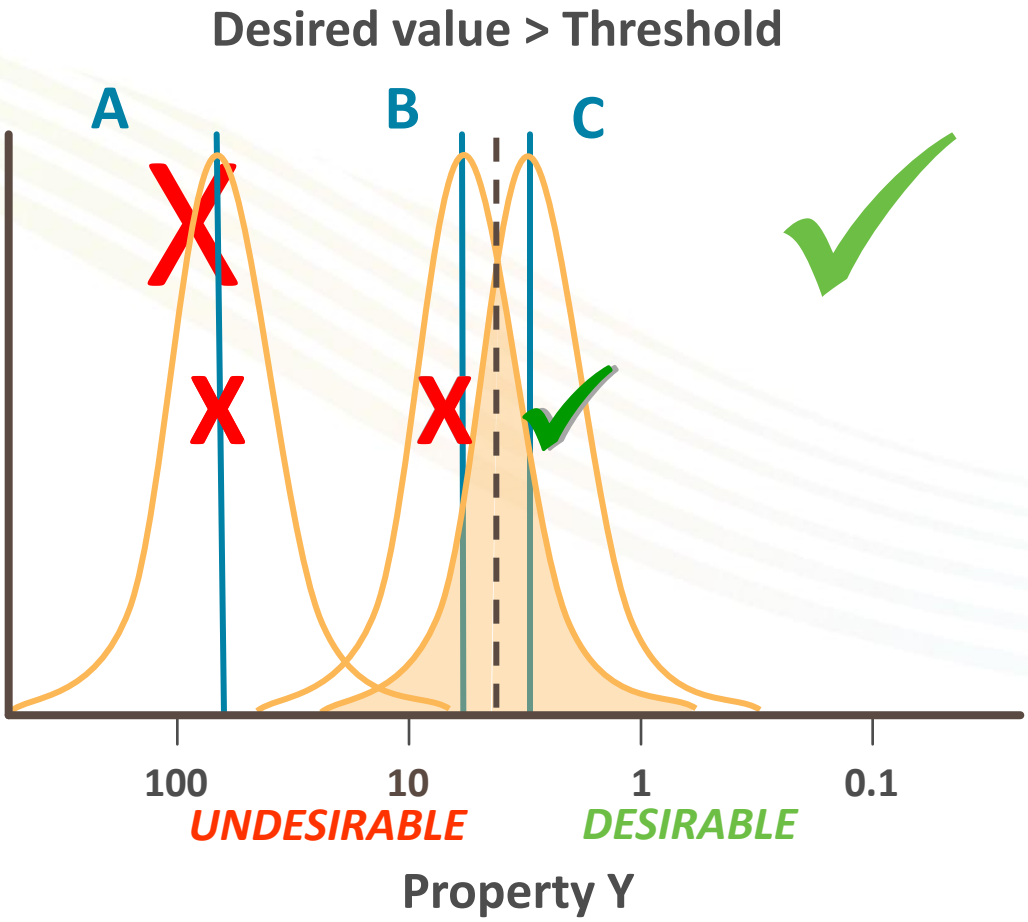
Probabilistic Scoring

StarDrop™5
Software that guides you to successful drug discovery



*Segall, Multi-Parameter Optimization..., Curr. Pharm. Des., **18**, 1292-1310(2012)

Importance of Uncertainty

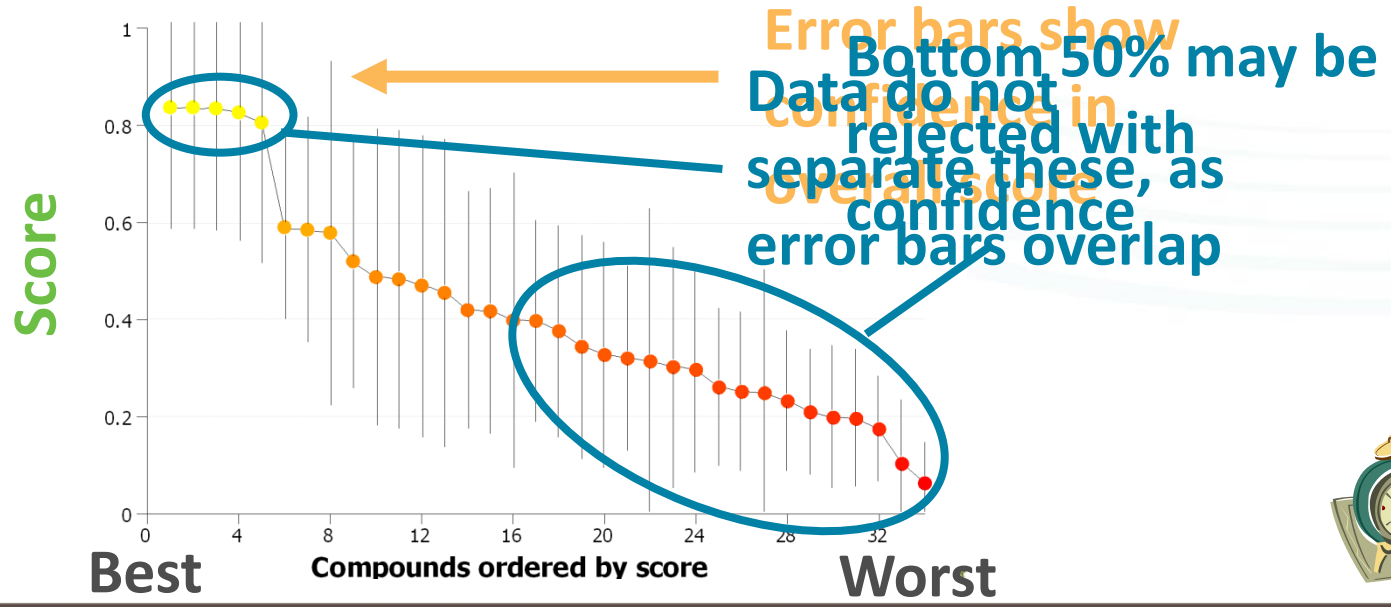


StarDrop Prioritisation
Probabilistic Scoring

StarDrop™5
Software that guides you to successful drug discovery

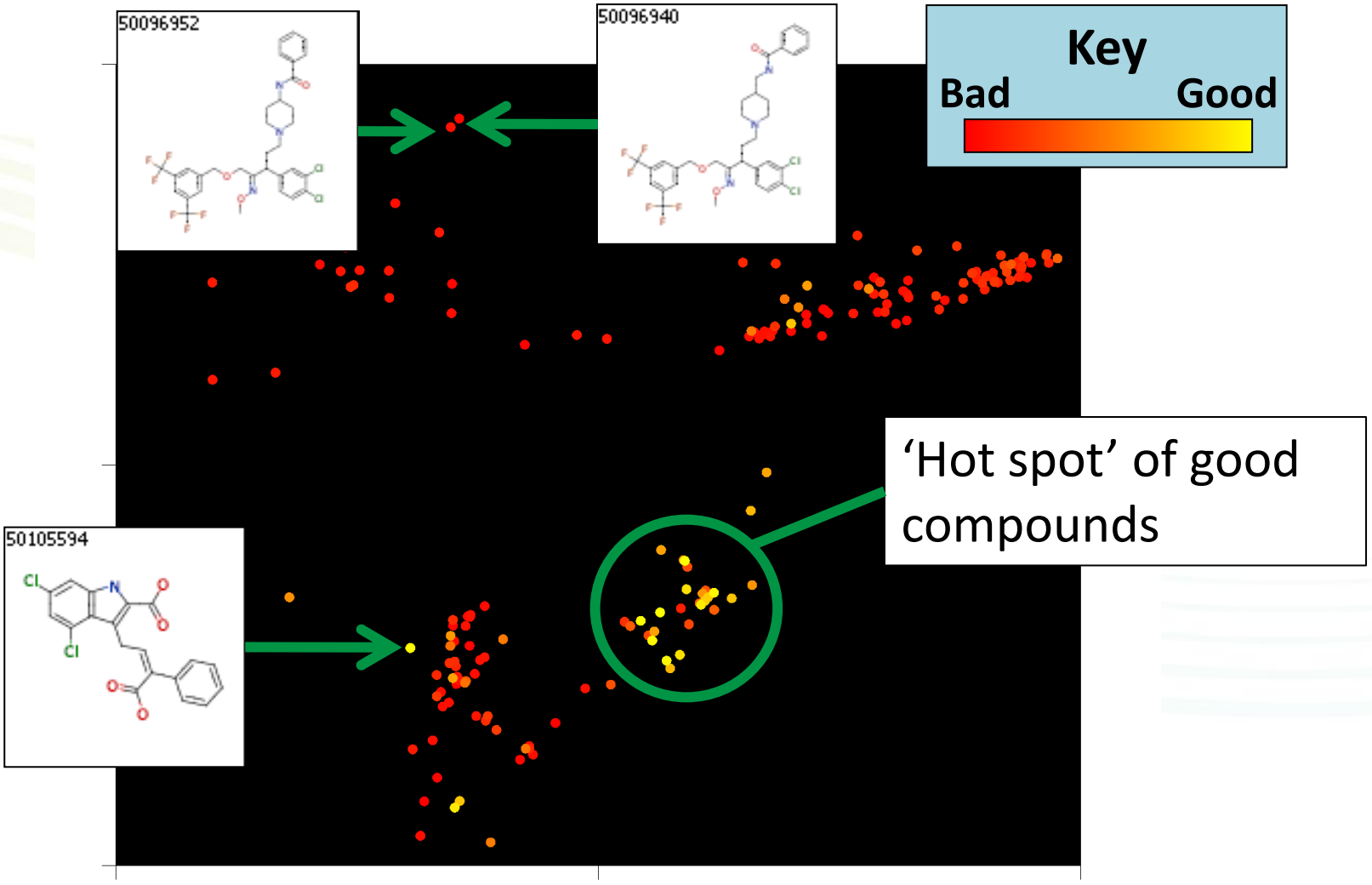
- **Property data**
  - Experimental or predicted
- **Criteria for success**
  - Relative importance
- **Uncertainties in data**
  - Experimental or statistical

- Score (Likelihood of Success)
- Confidence in score



Visualising 'Chemical Space'
Exploring trends in chemical diversity

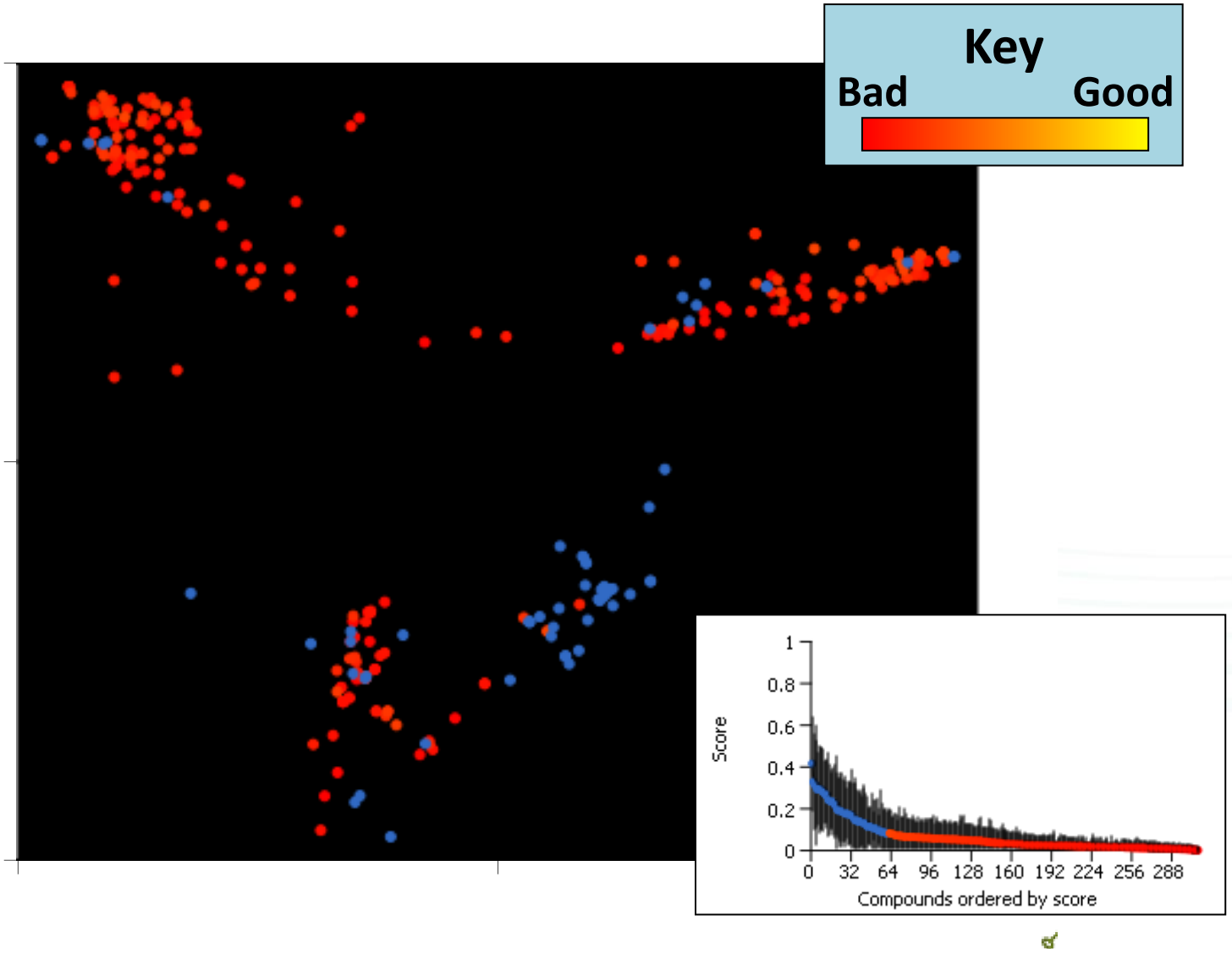
StarDrop™5
Software that guides you to successful drug discovery



Balance Quality Against Diversity

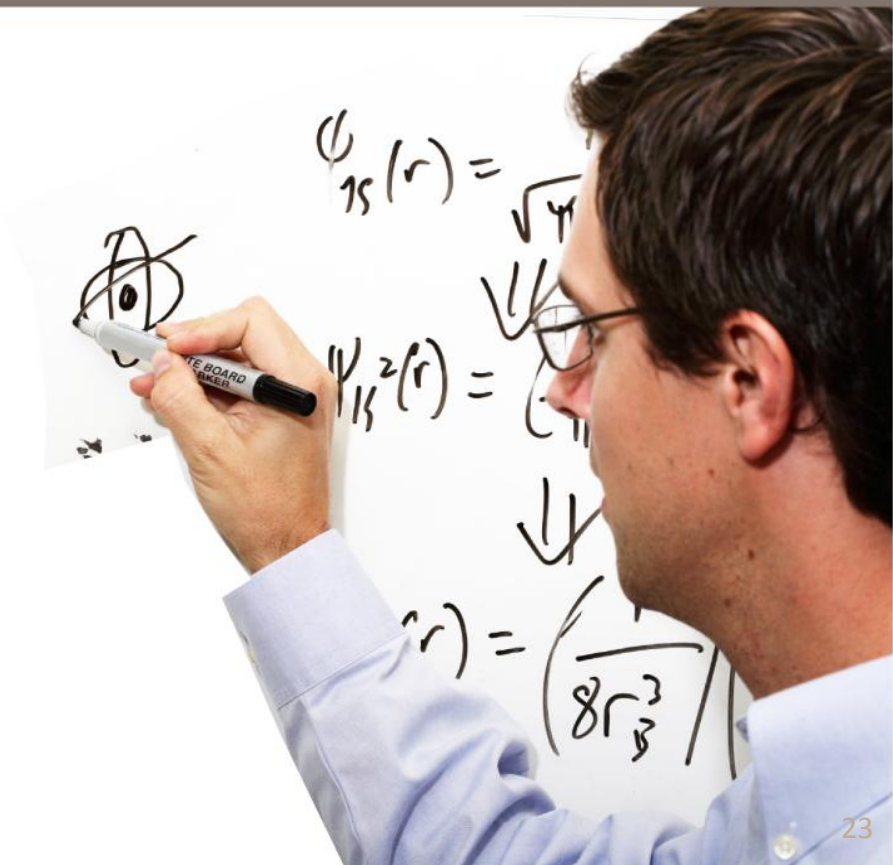
Mitigating risk

StarDrop™5
Software that guides you to successful drug discovery



Moving Towards Drug Design

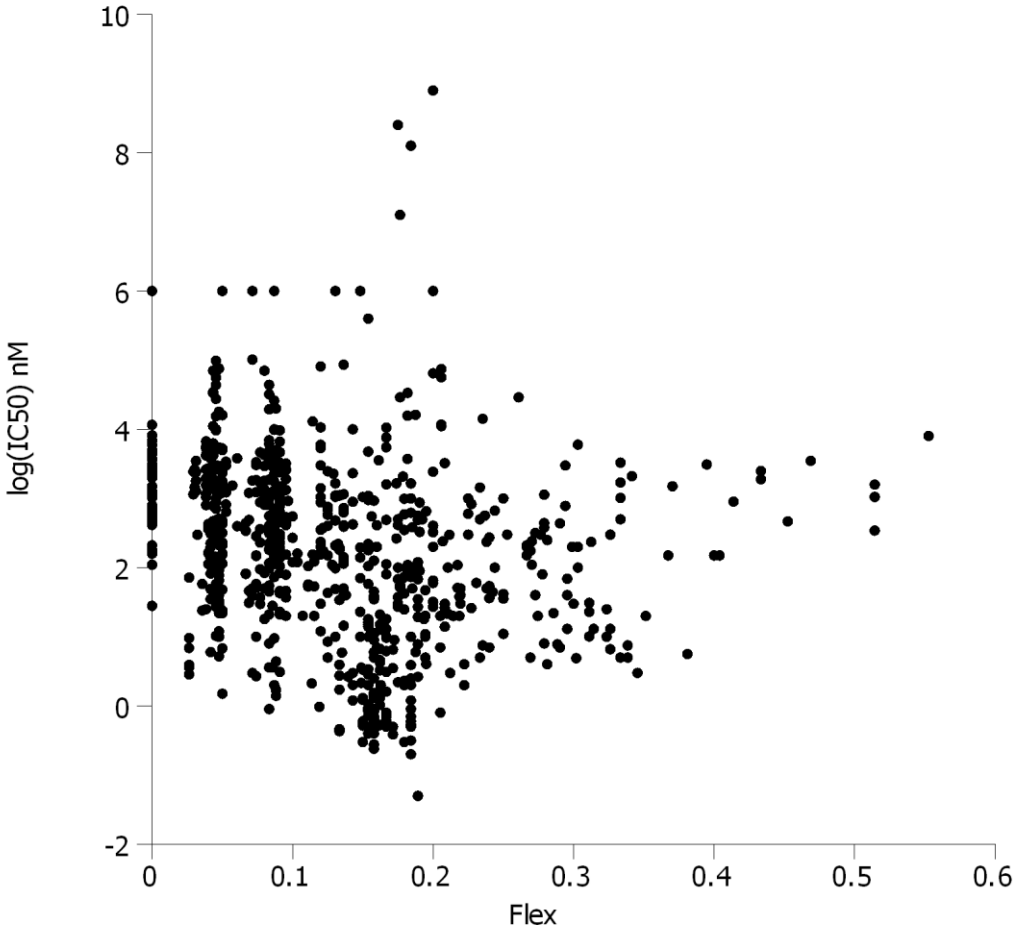
optibrium



Improve Accuracy of Prediction

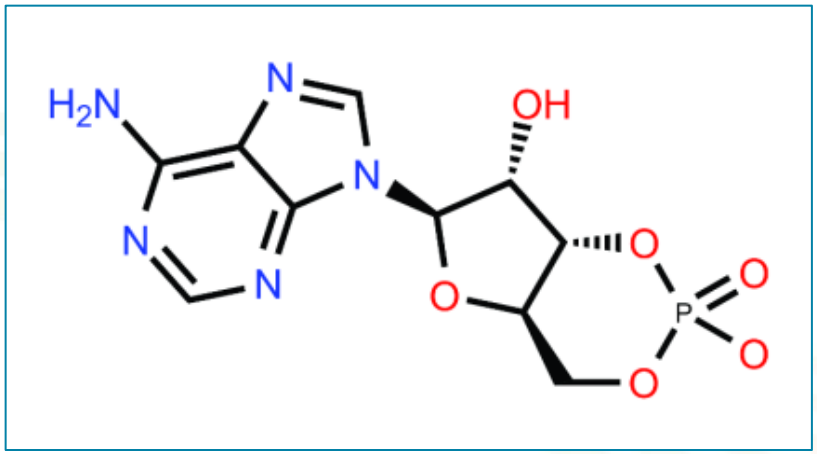
- Better modelling algorithms?
  - Advanced machine learning, e.g. random forests, Gaussian processes, support-vector machines...
- Better data?
  - Always welcome! But, lots more than data is available than ever before, e.g. PubChem, PDB, Chemble, Bindingdb...
- Better descriptors?

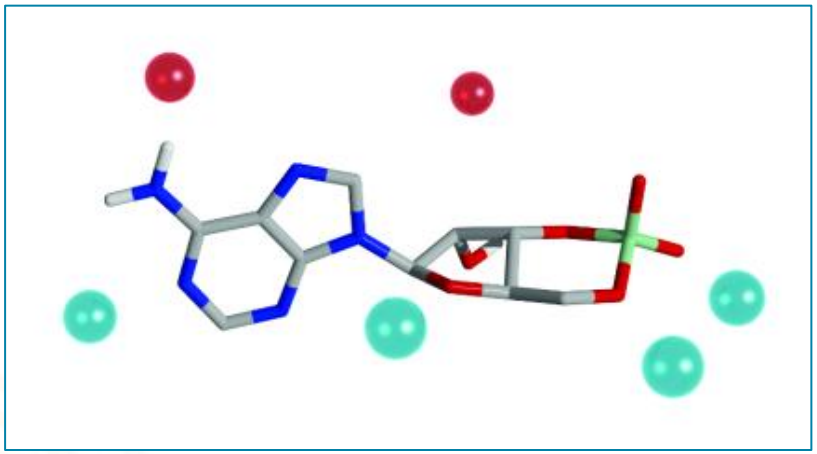
Structural Descriptors

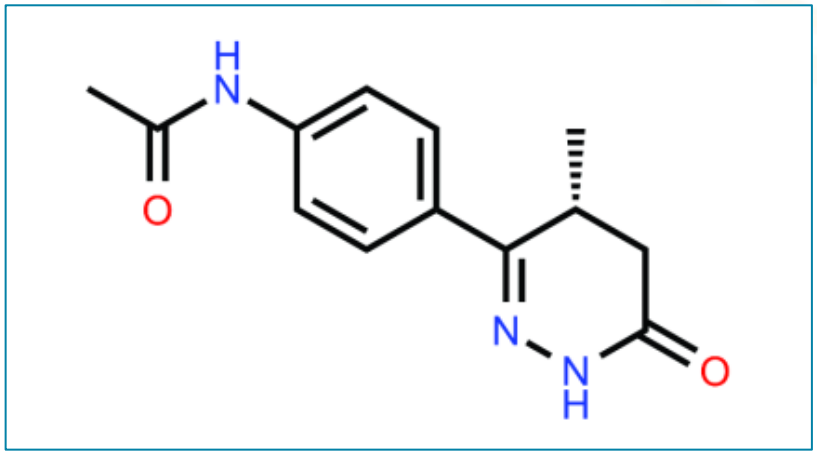


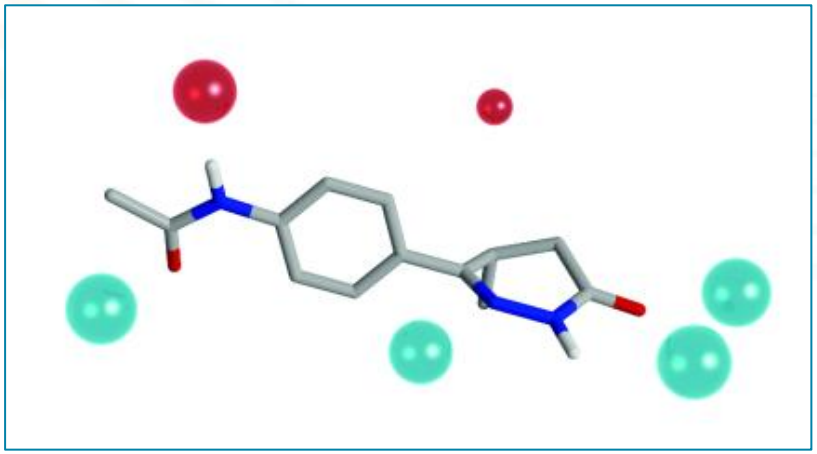
Better Description of Physics/Chemistry
E.g. Fields

cresset









Better Description of Physics/Chemistry
Quantum Mechanical Description

- Quantum mechanics captures electronic properties and energetics with a high degree of accuracy
  - Slow
  - But, becoming more accessible on a routine basis

- Examples:
  - Hydrogen bonding acidity
    - Kenny PW. (2009) J. Chem. Inf. Model. 49:1234-1244.
  - Lability to metabolism
    - Jones JP *et al.* (2002) Drug Metab. Dispos. 30:7-12.
  - Binding energies
    - Heady *et al.* (2006) J. Med. Chem. 49:5141-5153.
  - Classical MD parameterised using DFT
    - Bartok *et al.* (2010) Phys. Rev. Lett. 104:136403.

Conclusions

- Predictive models are not yet accurate enough to enable a true drug *design* paradigm
- However, models provide value by helping to reduce wasted effort and focus efforts on chemistries with the best chance of success
- QM approaches may offer one way to move towards true drug *design*
  - Still some way to go before these methods can be routinely used
- Of course, modelling also adds value by helping to understand and interpret SAR