optibrium

Visual analyses for guiding compound selection and design

Ed Champness – March 25th 2010

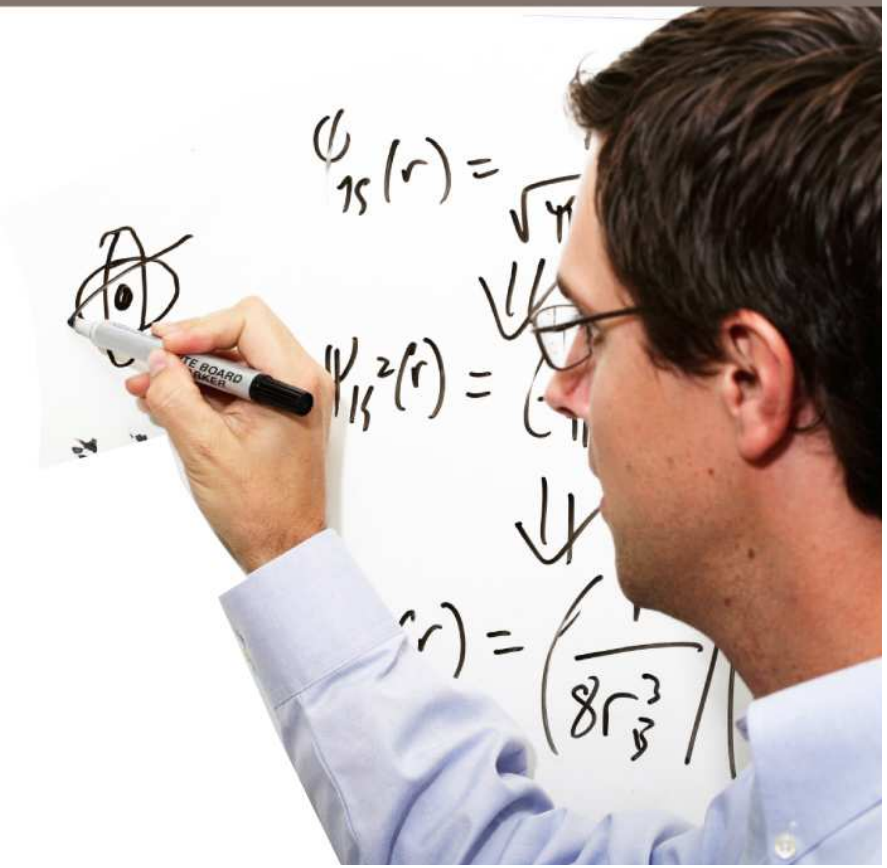
ACS Division of Chemical Information, Pub # 049

Overview

- Challenges of decision making in Drug Discovery
- What questions do we ask?
- A workflow
  - Prioritisation
  - Selection
  - Design/redesign
- Conclusions

Challenges

optibrium



Challenges

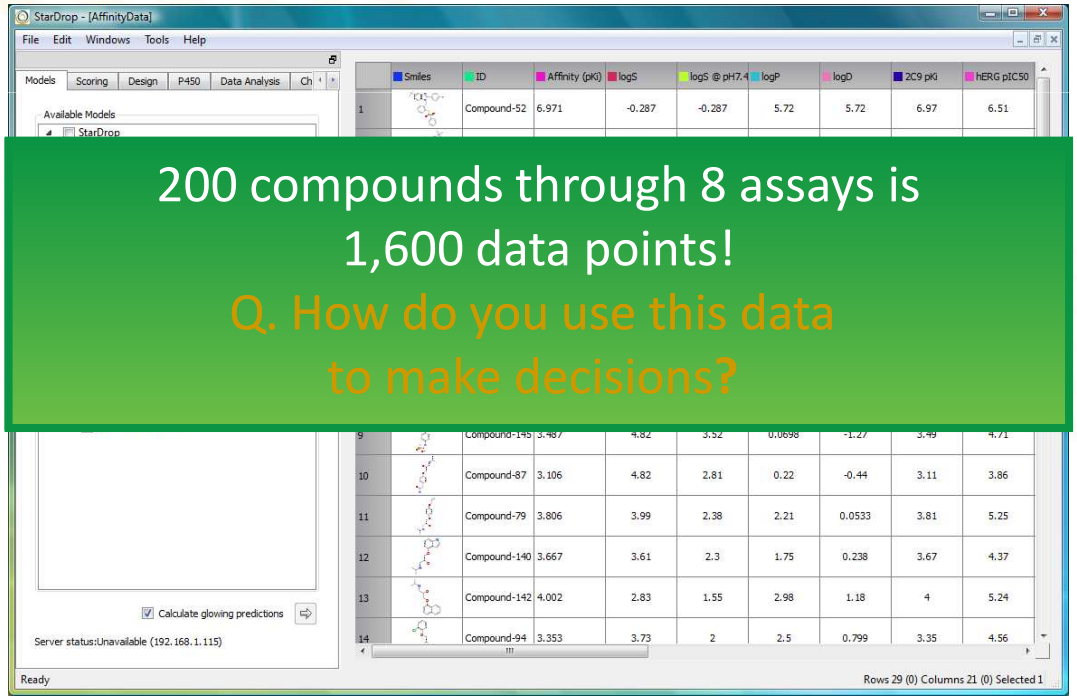
Decision-making in Drug Discovery involves:

- Potentially large volumes of information, multiple parameters, many sources
- Uncertain, sparse data
- Cross-discipline coordination/agreement

Challenges

Decision-making in Drug Discovery involves:

- Potentially large volumes of information, multiple parameters, many sources
- Uncertain, sparse data
- Cross-discipline coordination/agreement



200 compounds through 8 assays is 1,600 data points!

Q. How do you use this data to make decisions?

Challenges

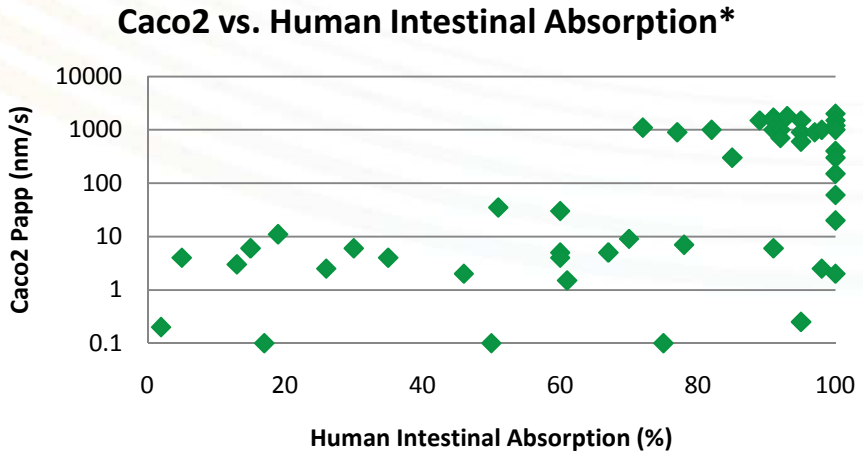
Decision-making in Drug Discovery involves:

- Potentially large volumes of information, multiple parameters, many sources
- Uncertain, sparse data
- Cross-discipline coordination/agreement

In silico, in vitro, in vivo – they're all models, but...

We don't have this:

Instead, something more like this?

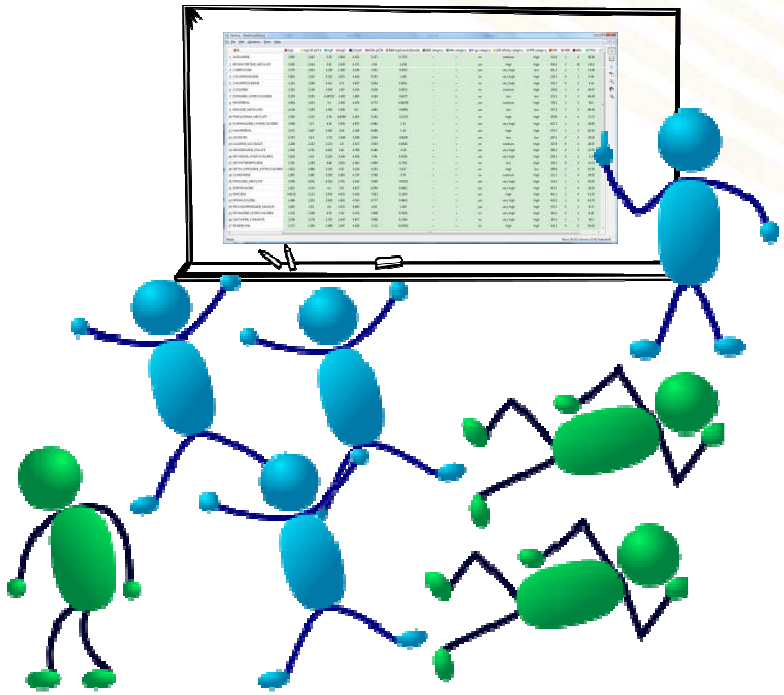


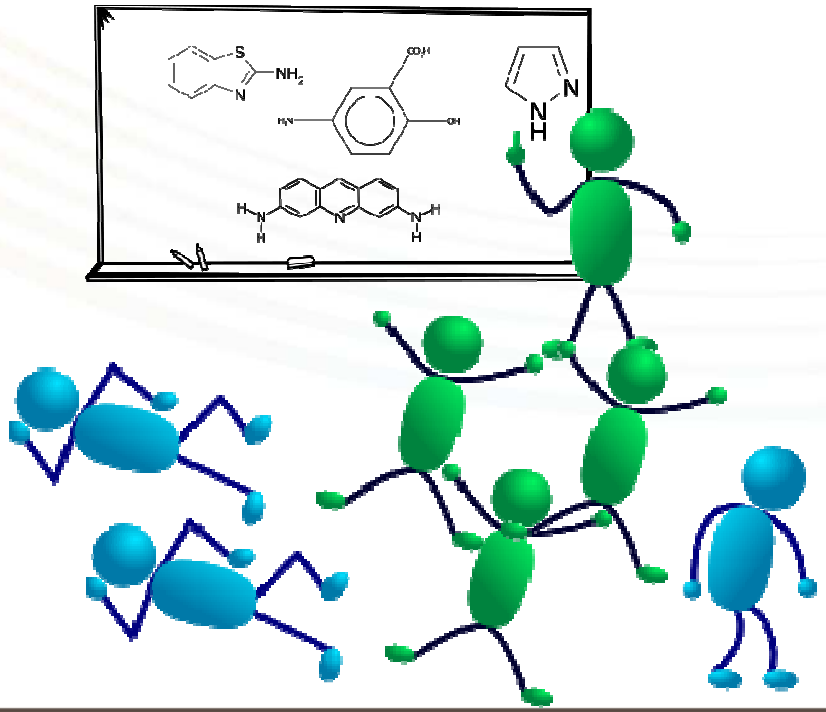
* Irvine, *et al.*, J. Pharm. Sci. 1999, 88, 28

Challenges

Decision-making in Drug Discovery involves:

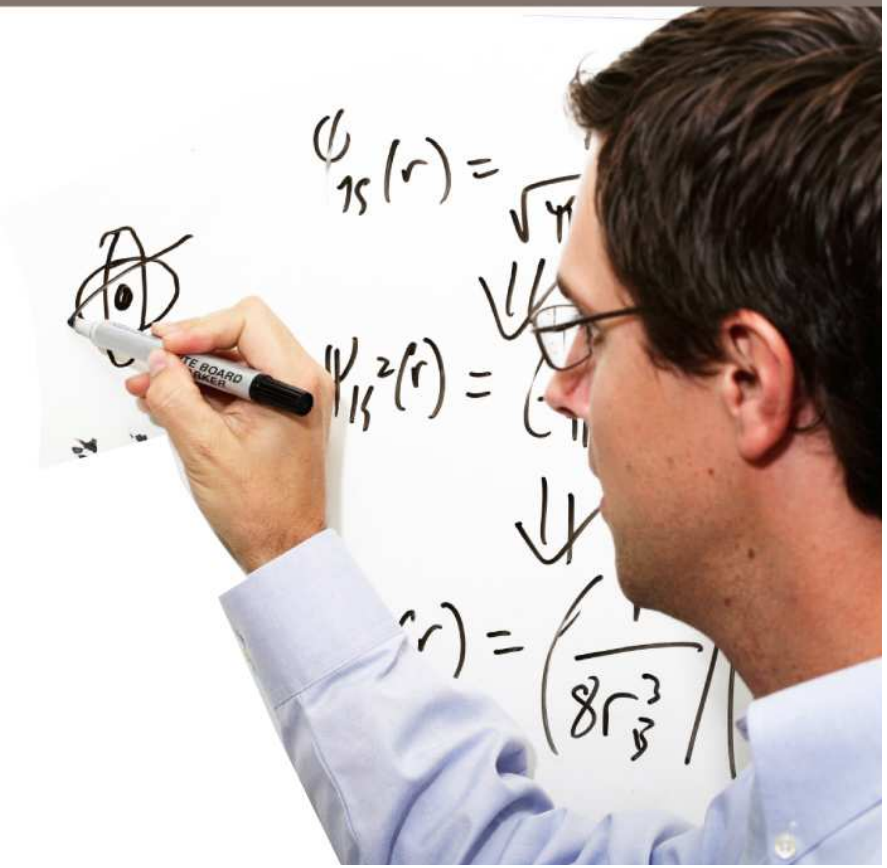
- Potentially large volumes of information, multiple parameters, many sources
- Uncertain, sparse data
- Cross-discipline coordination/agreement





What questions do we ask?

optibrium



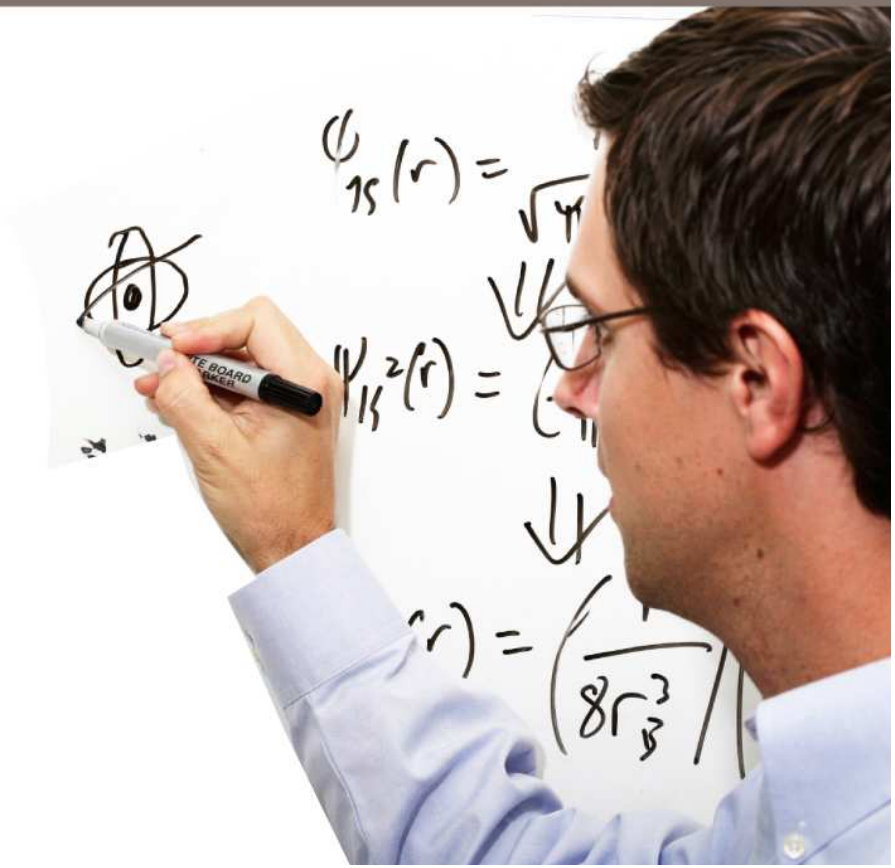
What questions do we ask?

Decision-making in Drug Discovery involves:

- Potentially large volumes of information, multiple parameters, many sources
  - "How can I get a good high level view of my data?"
  - "How do I get to the detail?"
- Uncertain, sparse data
  - "Which compound has the best overall balance of properties?"
  - "How much confidence can I have in my decisions?"
- Cross-discipline coordination/agreement
  - "Why does this structure have that property value?"
  - "What should I do to my molecule to change a property?"

Workflow

optibrium



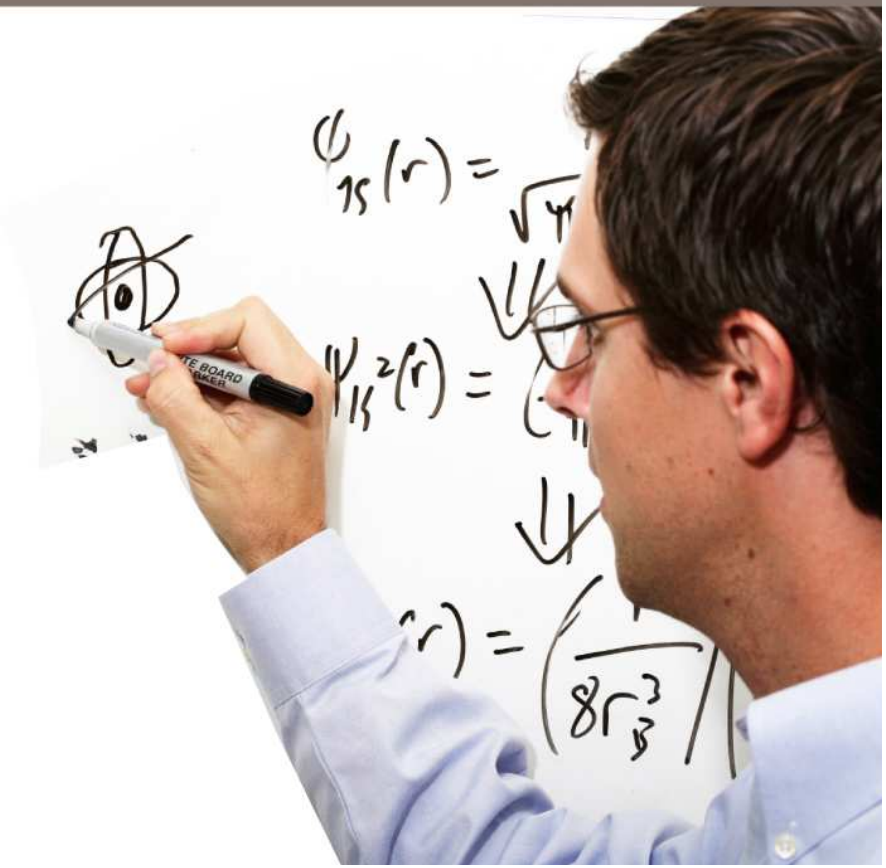
Workflow

Applying data to guide decisions

In silico	Importance	Quality
In vitro	Uncertainty	Diversity
In vivo		

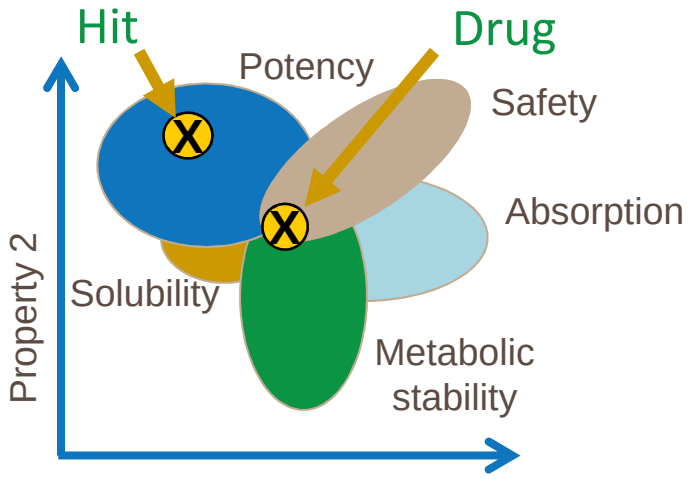
Prioritisation

optibrium

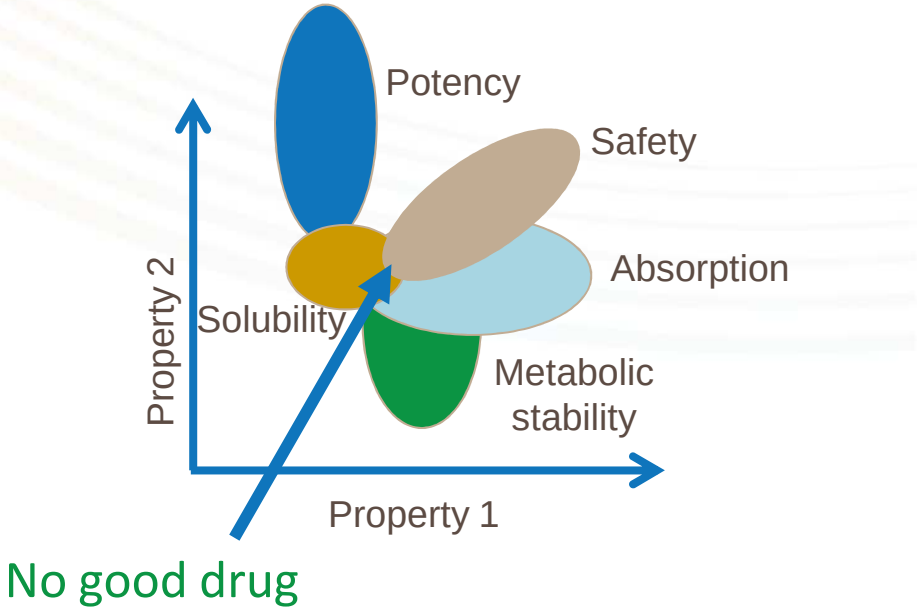


The Objectives

- Identify chemistries with an optimal **balance** of properties



- Quickly identify situations when such a balance is not possible
  - –Fail fast, fail cheap
  - –Only when **confident**



StarDrop Prioritisation:
Probabilistic Scoring

"Which compound has the best overall balance of properties?"
"How much confidence can I have in my decisions?"

Integrated assessment of data against project criteria

Uniquely accounts for the uncertainties in all compound-related data (experimental or calculated)

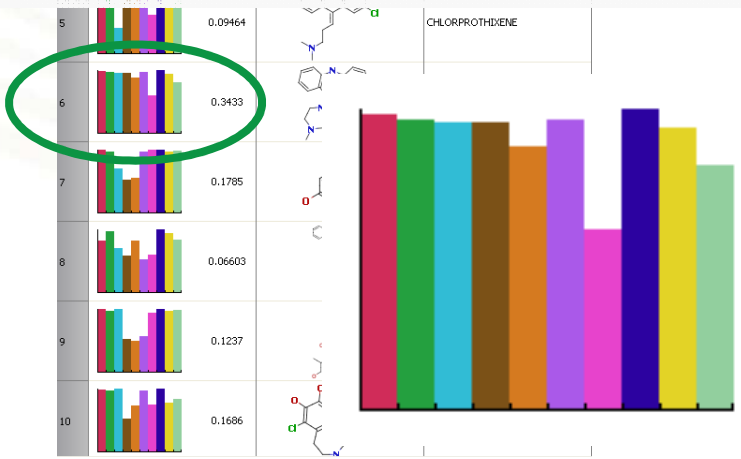
User-defined scoring profile

Profile: Oral CNS Scoring Profile

Property	Desired Value	Importance
logS	> 1	
HIA category	+	
logP	0.0 -> 3.5	
BBB category	+	
BBB log([brain]:[blood])	-0.20 -> 1.00	
P-gp category	no	
hERG pIC50	⩽ 5	
2C9 pKi	⩽ 6	
2D6 affinity category	low medium	
PPB category	low	

Compounds ranked

Histograms for quick visual guide to compound properties



StarDrop Prioritisation:
Probabilistic Scoring

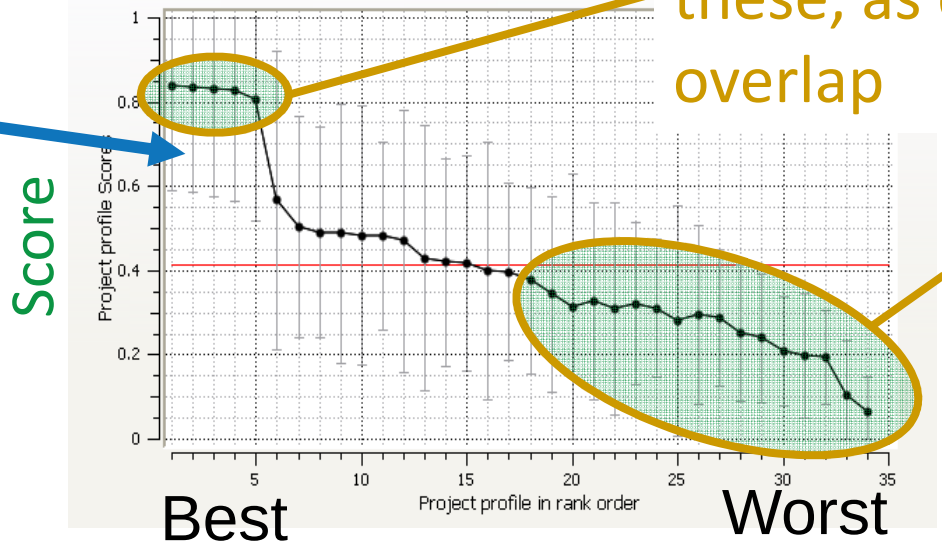
"Which compound has the best overall balance of properties?"
"How much confidence can I have in my decisions?"

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

- **Score (Likelihood of Success)**
- **Confidence in score**

Data do not separate these, as error bars overlap

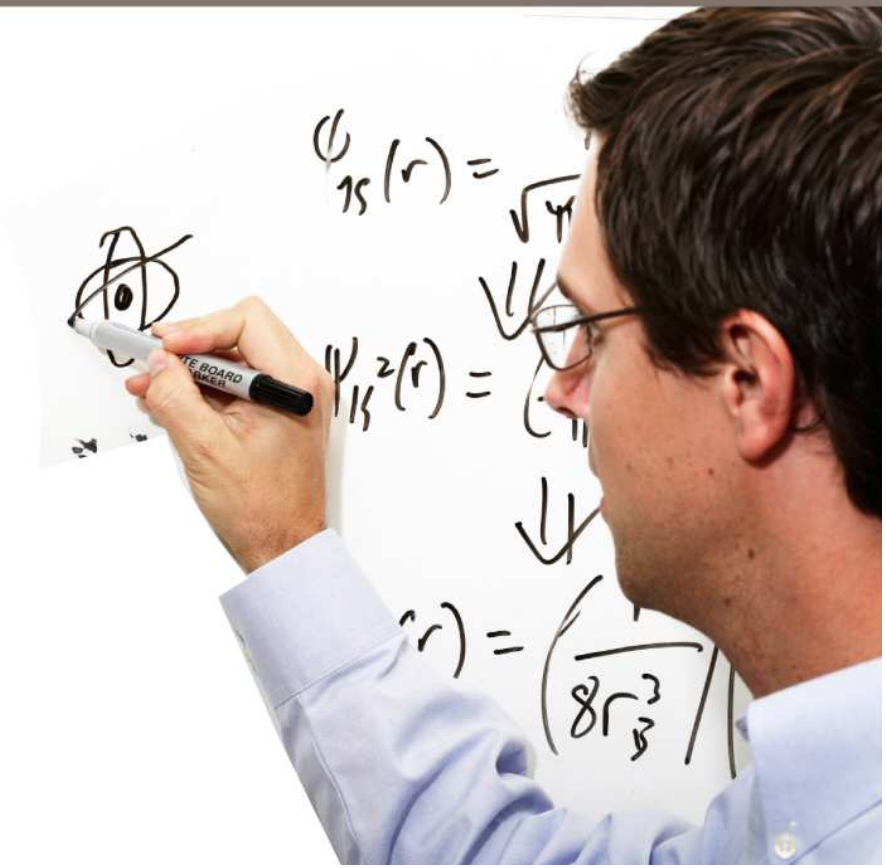
Error bars show confidence in overall score



Bottom 50% may be rejected with confidence

Selection

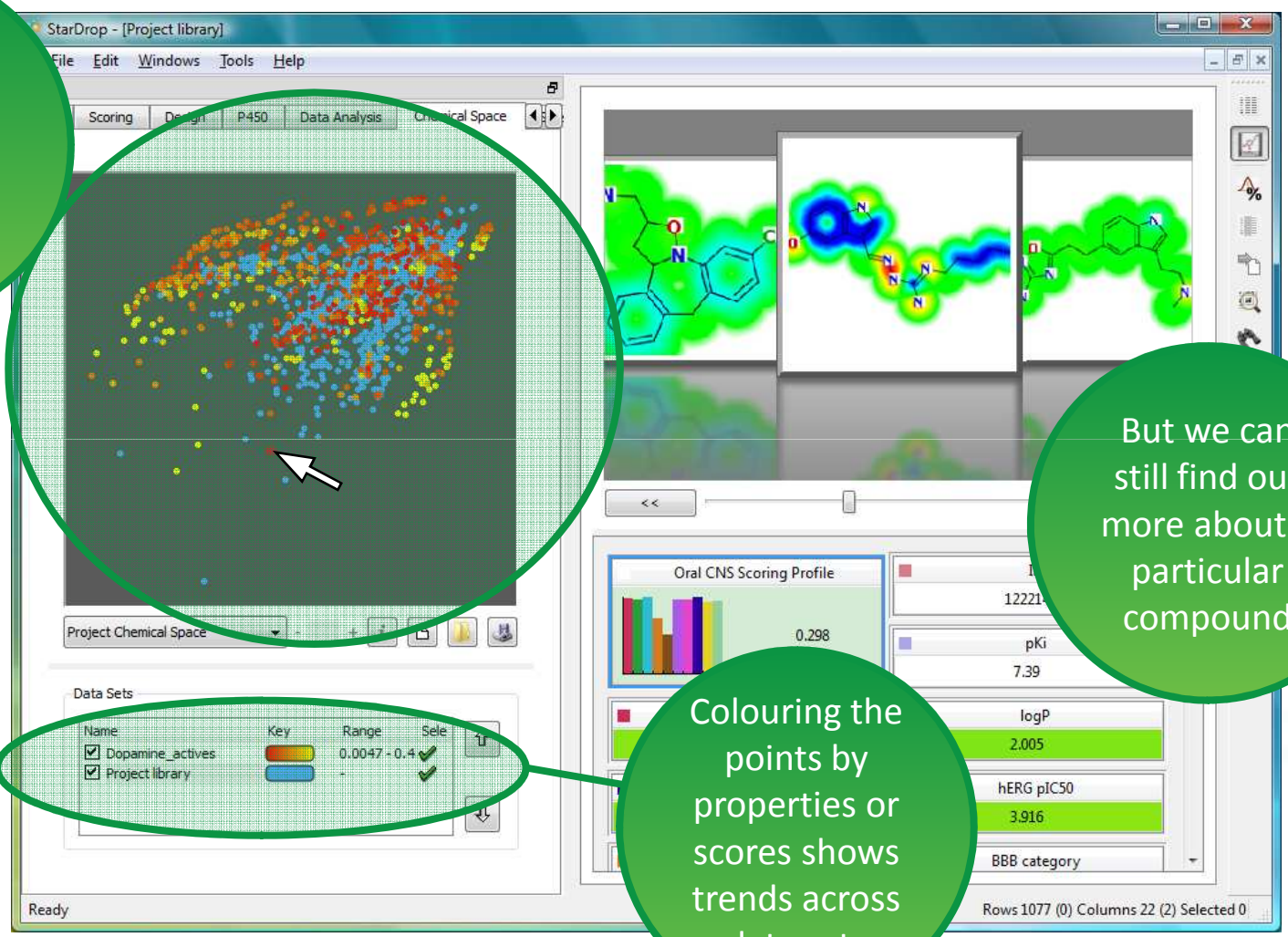
optibrium



Chemical Space
Visualise chemical diversity

"How can I get a good high level view of my data?"
"How do I get to the detail?"

Two compounds will be close together if they are structurally similar

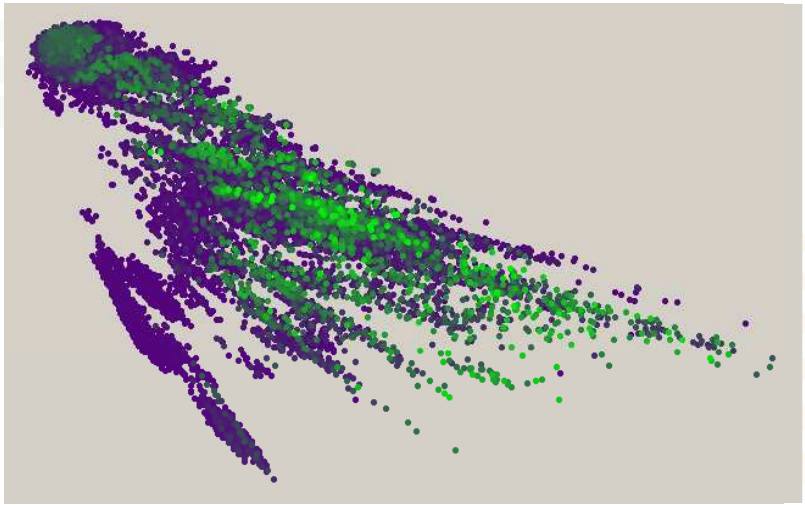


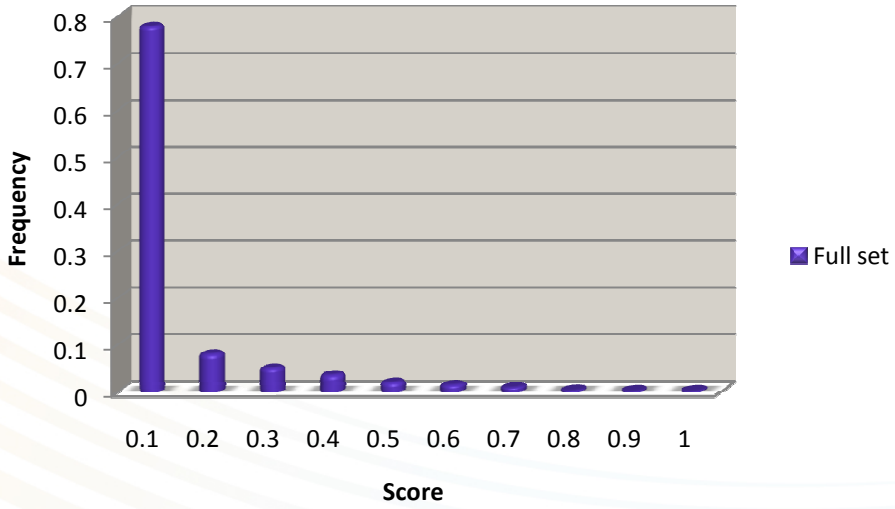
But we can still find out more about a particular compound

Colouring the points by properties or scores shows trends across data sets

Selection:
Balancing Quality and Diversity

"How can I get a good high level view of my data?"

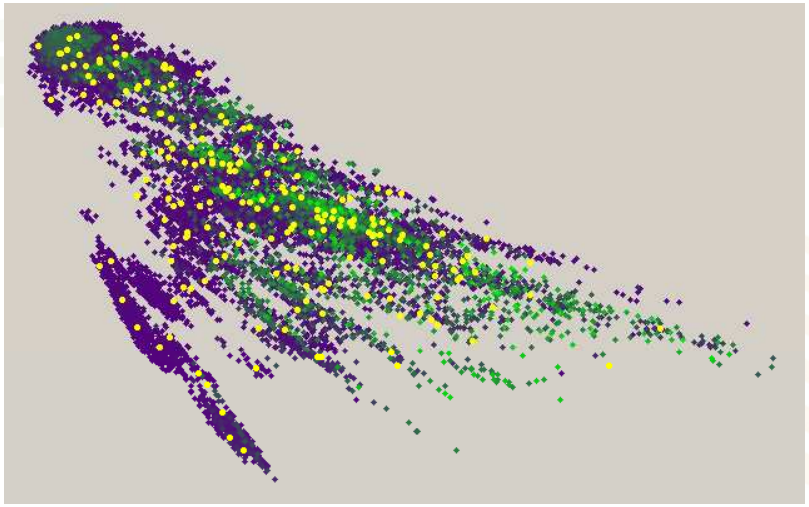


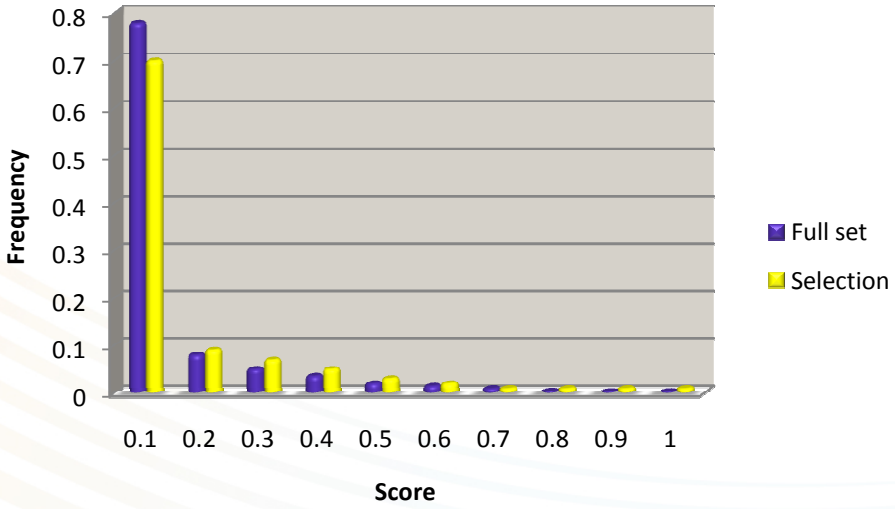


Objective: Select 200 compounds from scored library of 13,000 compounds

Selection:
Balancing Quality and Diversity

"How can I get a good high level view of my data?"

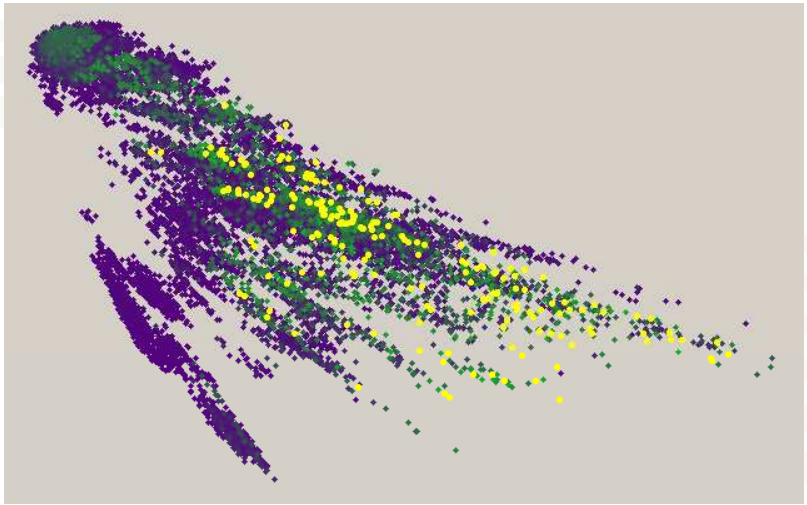


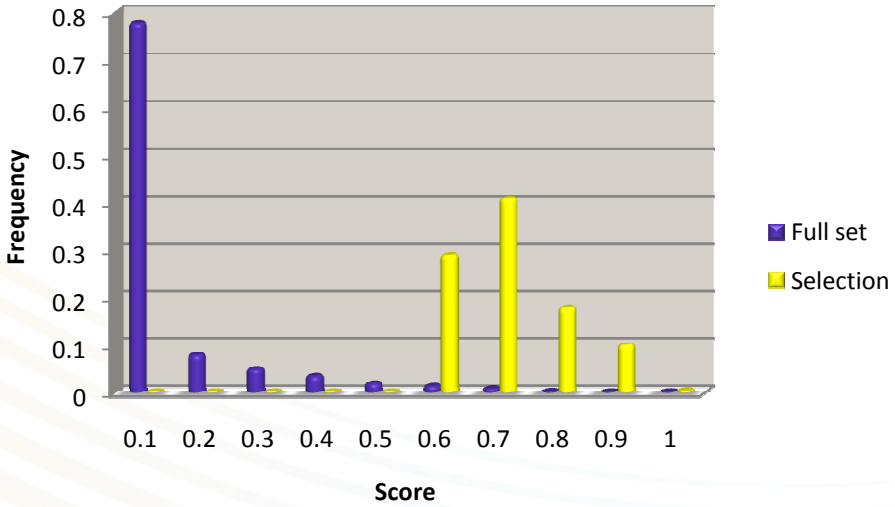


Diverse Sample

Selection:
Balancing Quality and Diversity

"How can I get a good high level view of my data?"

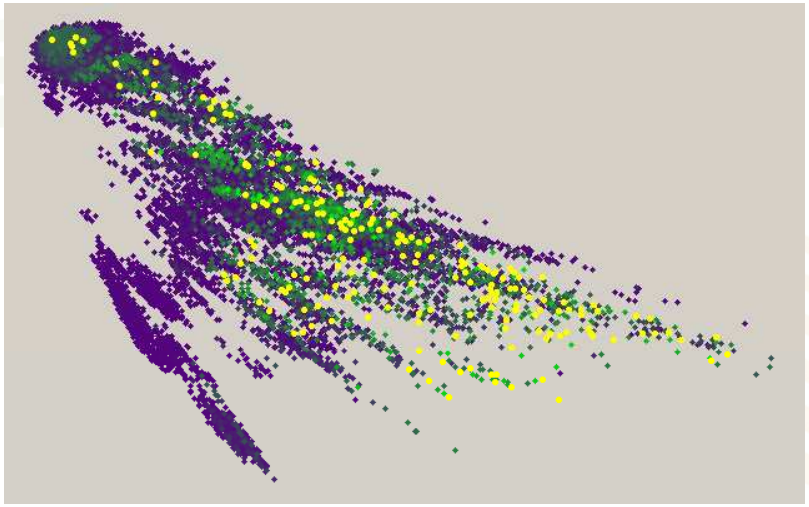


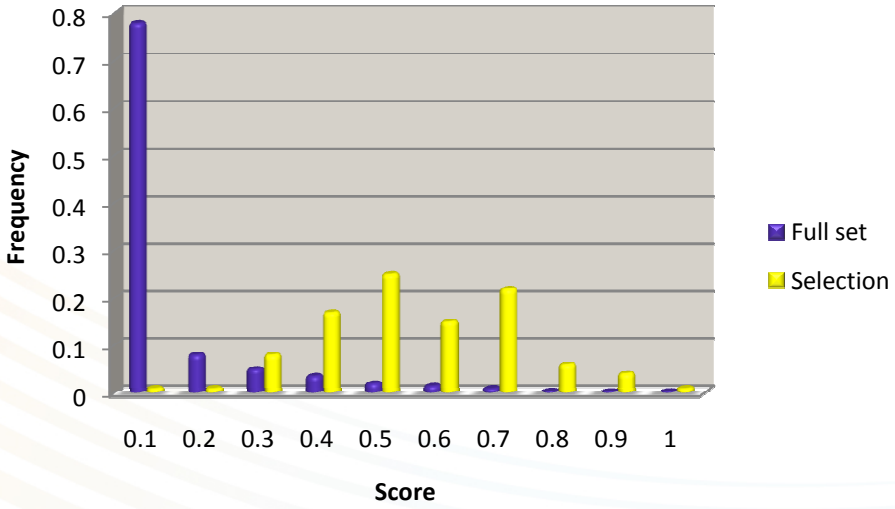


Top 200 ranked compounds

Selection:
Balancing Quality and Diversity

"How can I get a good high level view of my data?"

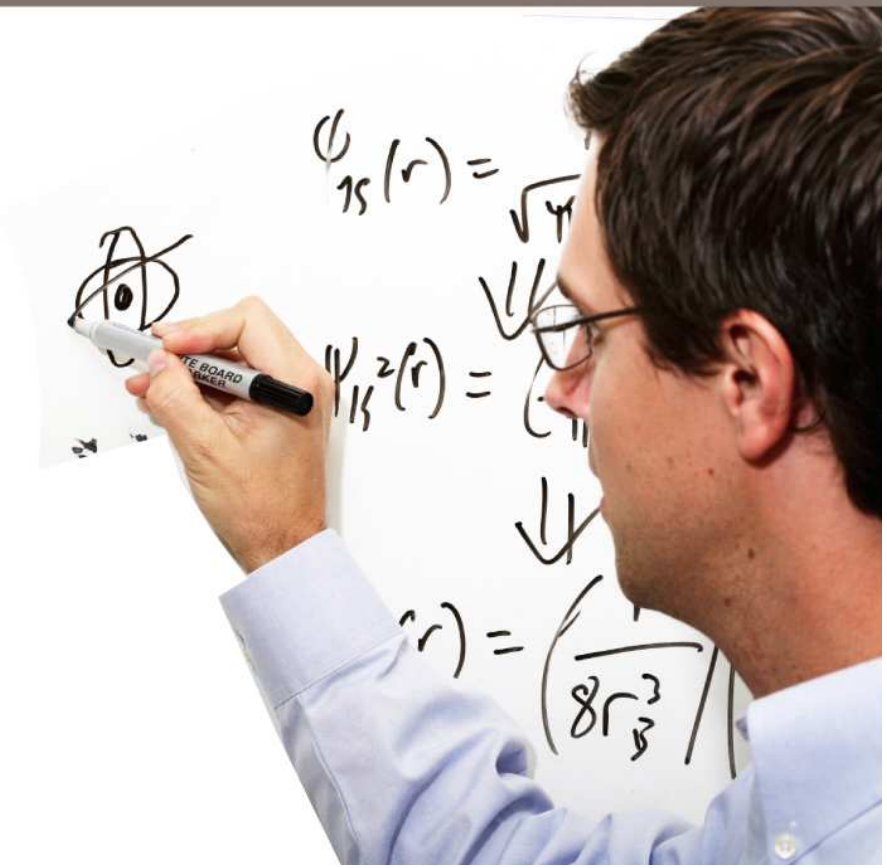




Balance Diversity:Rank = 80:20

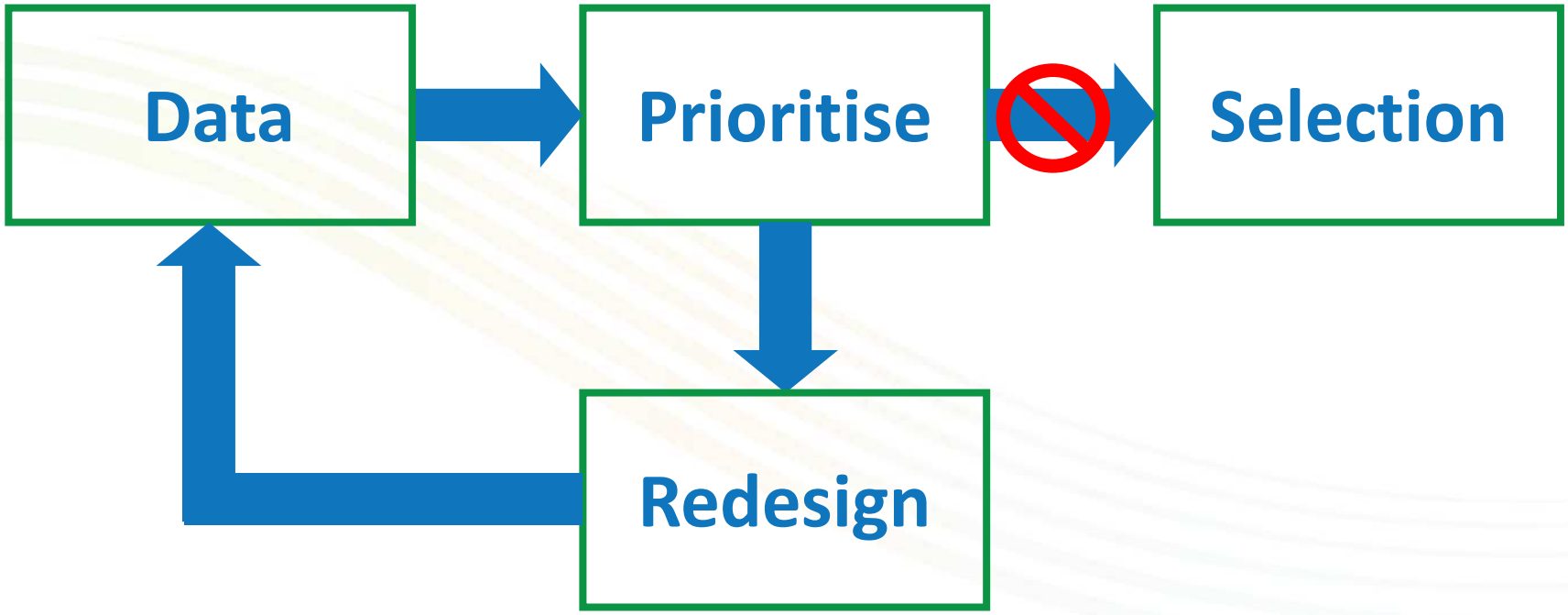
Design/Redesign

optibrium



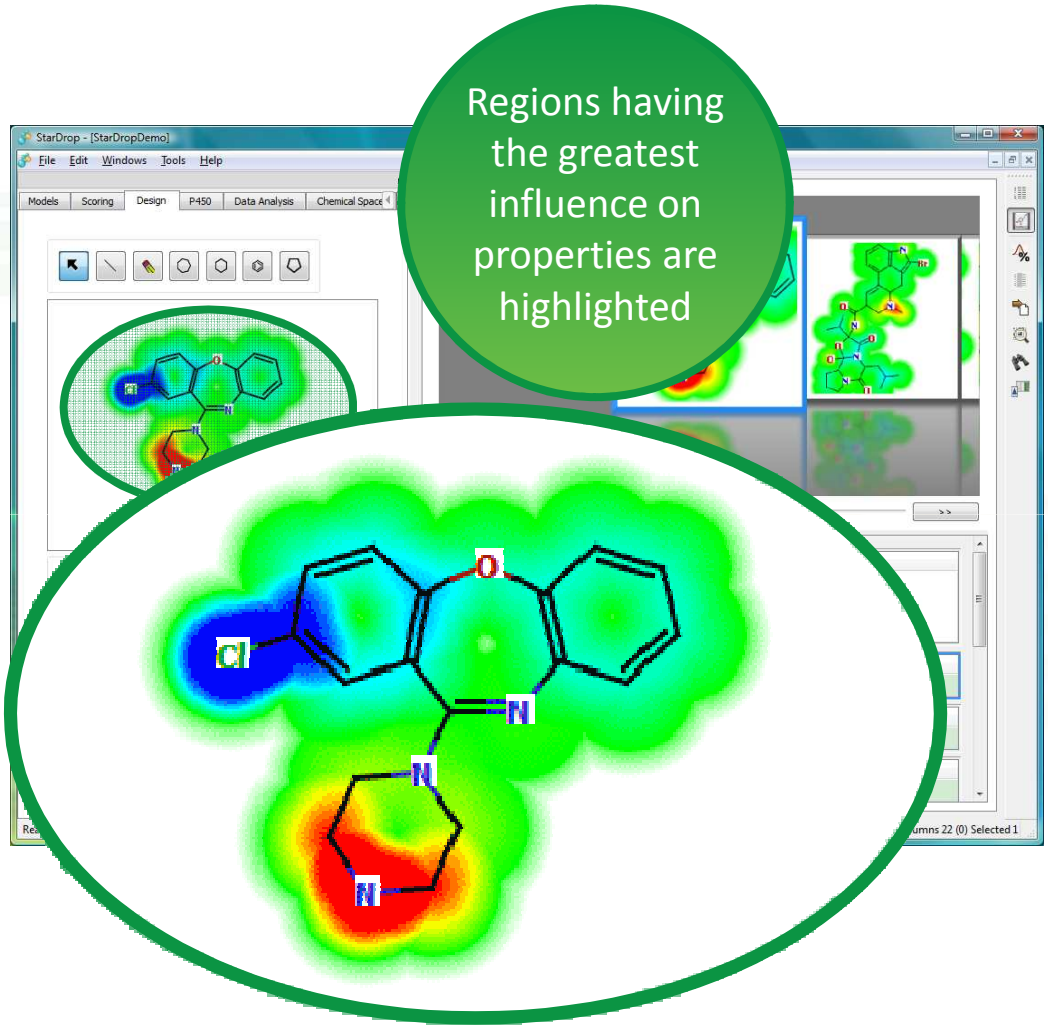
Workflow

Applying data to guide decisions



Interactive Redesign:
The 'Glowing Molecule'

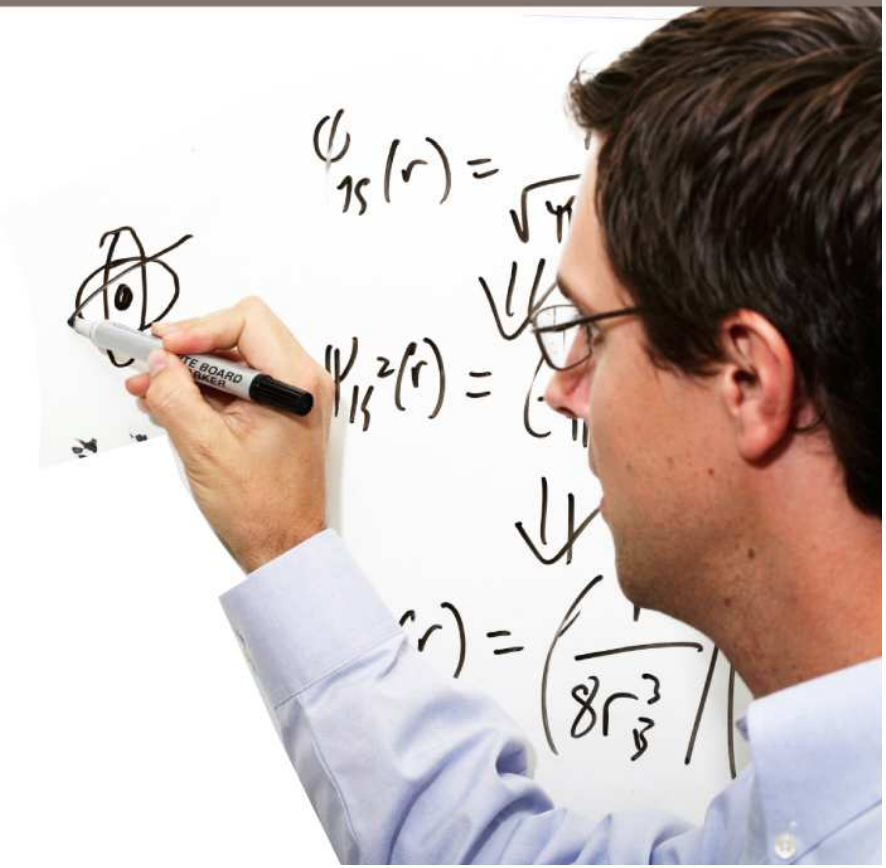
- "Why does this structure have that property value?"
- "What should I do to my molecule to change a property?"



- Interactive redesign to explore new ideas
- Visual feedback on structural influences on predicted properties
  - Interpret SAR to guide redesign of molecules
  - Individual properties or scores

Conclusions

optibrium



Conclusions

- The data we have available is only as valuable as the decisions we are able to make with it
- The decisions we make are dependant on the quality of our data, but we can take the uncertainty into account
- Given that we have multiple parameters with uncertainty wrapped into our decisions we need both high level and detailed views together allows us to understand the decisions our data is leading us towards

Acknowledgements

- Matt Segall
- Chris Leeding
- Andy Usher
- Olga Obrezanova
- Brett Saunders

optibrium StarDrop

optibrium StarVue