optibrium

Finding and Applying Rules for Successful Drug Discovery

GTC Drug Design & Medicinal Chemistry, May 9th 2013

Matthew Segall, CEO, Optibrium Ltd.

© 2013 Optibrium Ltd.
Optibrium™, StarDrop™, Auto-Modeller™ and Glowing Molecule™ are trademarks of Optibrium Ltd.

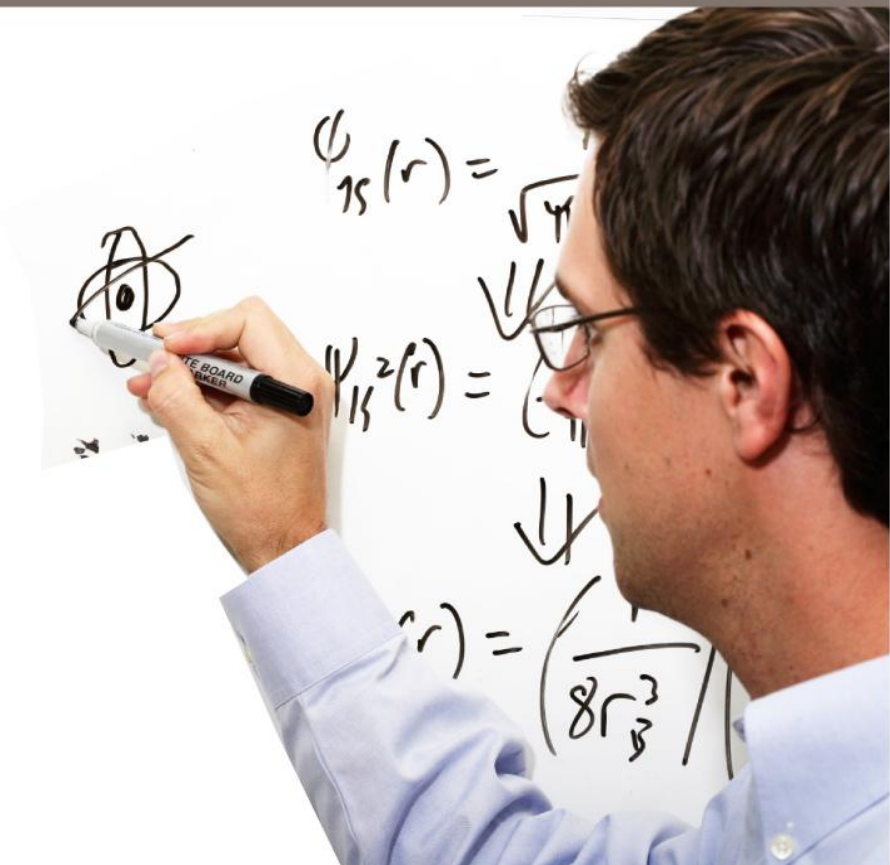
Overview

- Multi-parameter optimization (MPO) in drug discovery
- Case study: Balancing properties in lead optimization
- Finding multi-parameter rules for drug discovery
- Example application: ‘Drug-like’ properties
- Conclusions

© 2013 Optibrium Ltd.

Multi-parameter Optimization in Drug Discovery

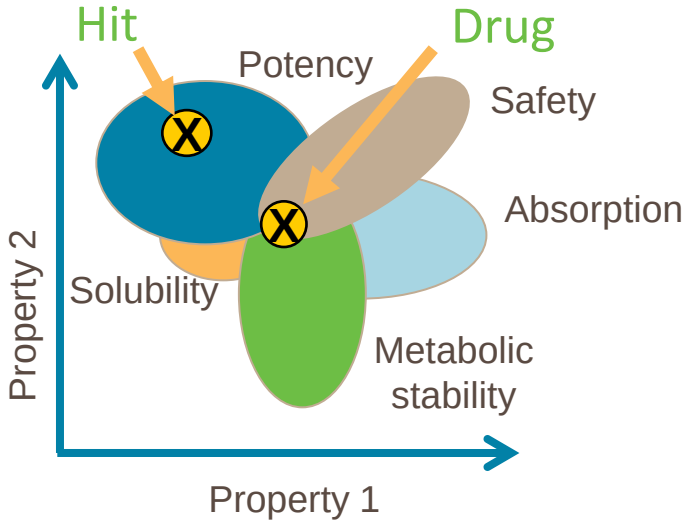
optibrium

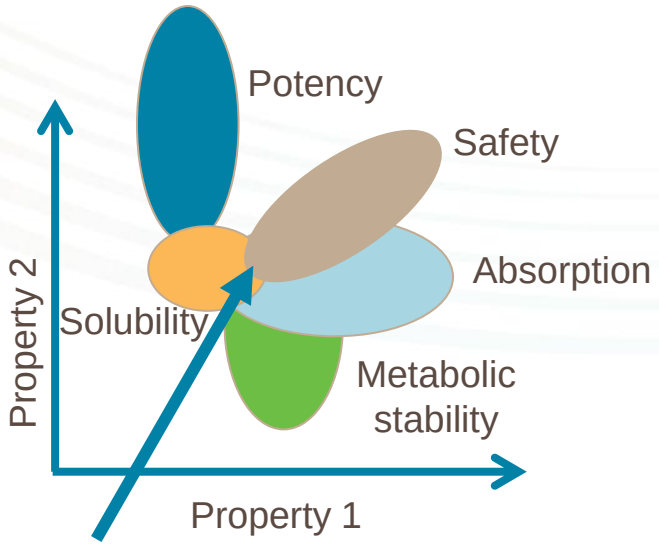


The Objectives of Drug Discovery
Multi-parameter optimisation

- 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**

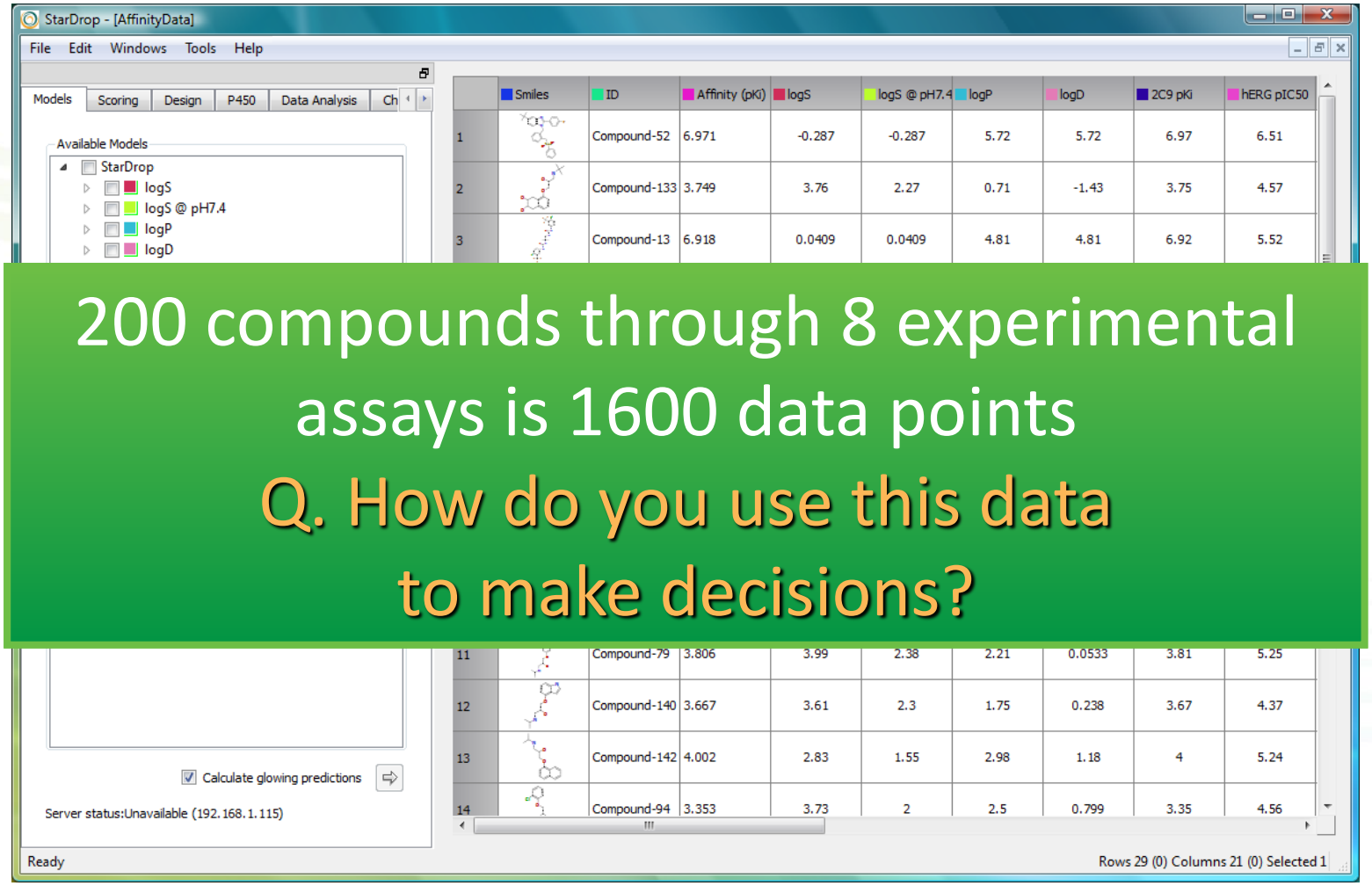




No good drug

© 2013 Optibrium Ltd.

The Challenge



	Smiles	ID	Affinity (pKi)	logS	logS @ pH7.4	logP	logD	2C9 pKi	hERG pIC50
1		Compound-52	6.971	-0.287	-0.287	5.72	5.72	6.97	6.51
2		Compound-133	3.749	3.76	2.27	0.71	-1.43	3.75	4.57
3		Compound-13	6.918	0.0409	0.0409	4.81	4.81	6.92	5.52
11		Compound-79	3.806	3.99	2.38	2.21	0.0533	3.81	5.25
12		Compound-140	3.667	3.61	2.3	1.75	0.238	3.67	4.37
13		Compound-142	4.002	2.83	1.55	2.98	1.18	4	5.24
14		Compound-94	3.353	3.73	2	2.5	0.799	3.35	4.56

200 compounds through 8 experimental assays is 1600 data points

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

© 2013 Optibrium Ltd.

Requirements for MPO in Drug Discovery

- Interpretable
  - Easy to understand compound priority and how to improve compounds' chances of success
- Flexibility
  - Define criteria depending on therapeutic objectives of project
- Weighting
  - Take into account relative importance of different endpoints to success of project
- Uncertainty
  - Take uncertainty into account, avoid missed opportunities

© 2013 Optibrium Ltd.

Approaches for MPO

- Many methods have been applied for MPO in drug discovery
  - Rules-of-thumb
  - **Filtering**
  - Calculated metrics
  - Pareto optimisation
  - Desirability functions
  - **Probabilistic scoring**
- For a detailed review, please see:
  - M.D. Segall Curr. Pharm. Des. **18**(9) pp. 1292-1310 (2012)
  - Download from http://bit.ly/1140lS1

© 2013 Optibrium Ltd.

Approaches for MPO
Filtering

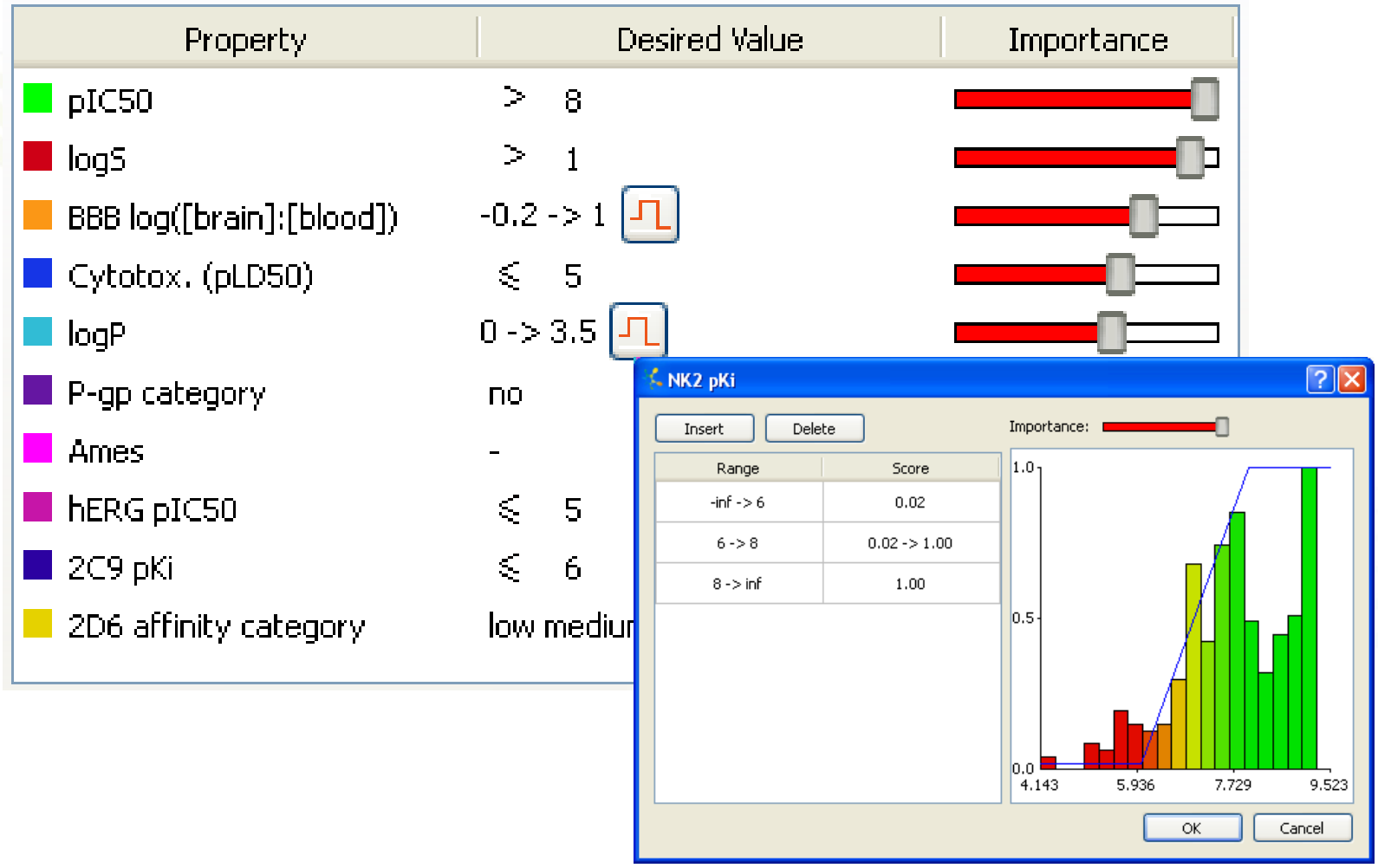
Potency

Absorpti

Metabolic Stability

© 2013 Optibrium Ltd.

Approaches for MPO
Probabilistic Scoring* – Scoring Profile



© 2013 Optibrium Ltd.

* Segall *et al.* (2009) Chem. & Biodiv. **6** p. 2144

Probabilistic Scoring*

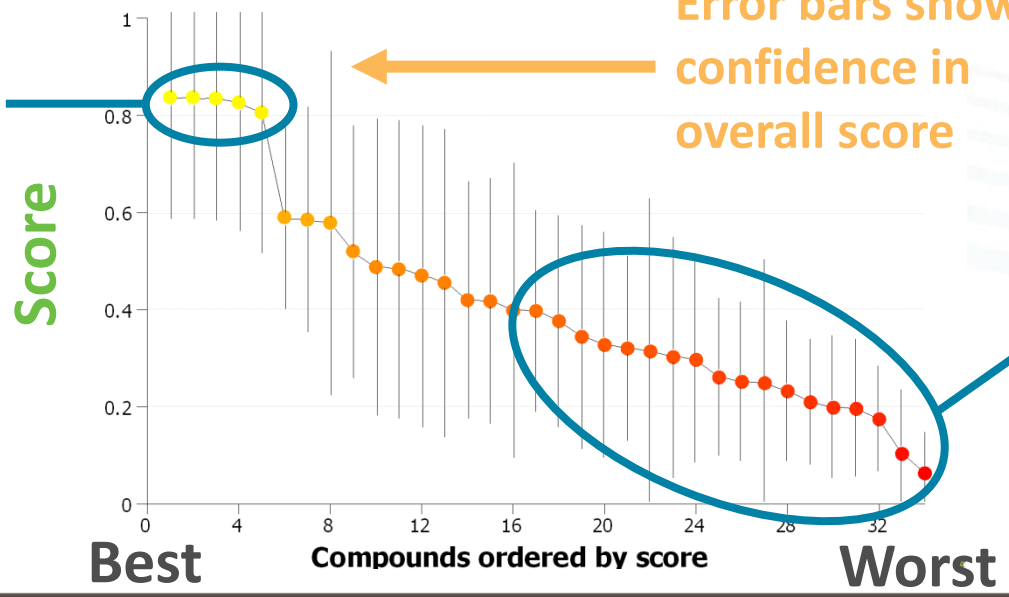
- **Property data**
  - Experimental or predicted
- **Criteria for success**
  - Relative importance
- **Uncertainties in data**
  - Experimental or statistical

- Score (Likelihood of Success)
- Confidence in score

Data do not separate these as error bars overlap

Error bars show confidence in overall score

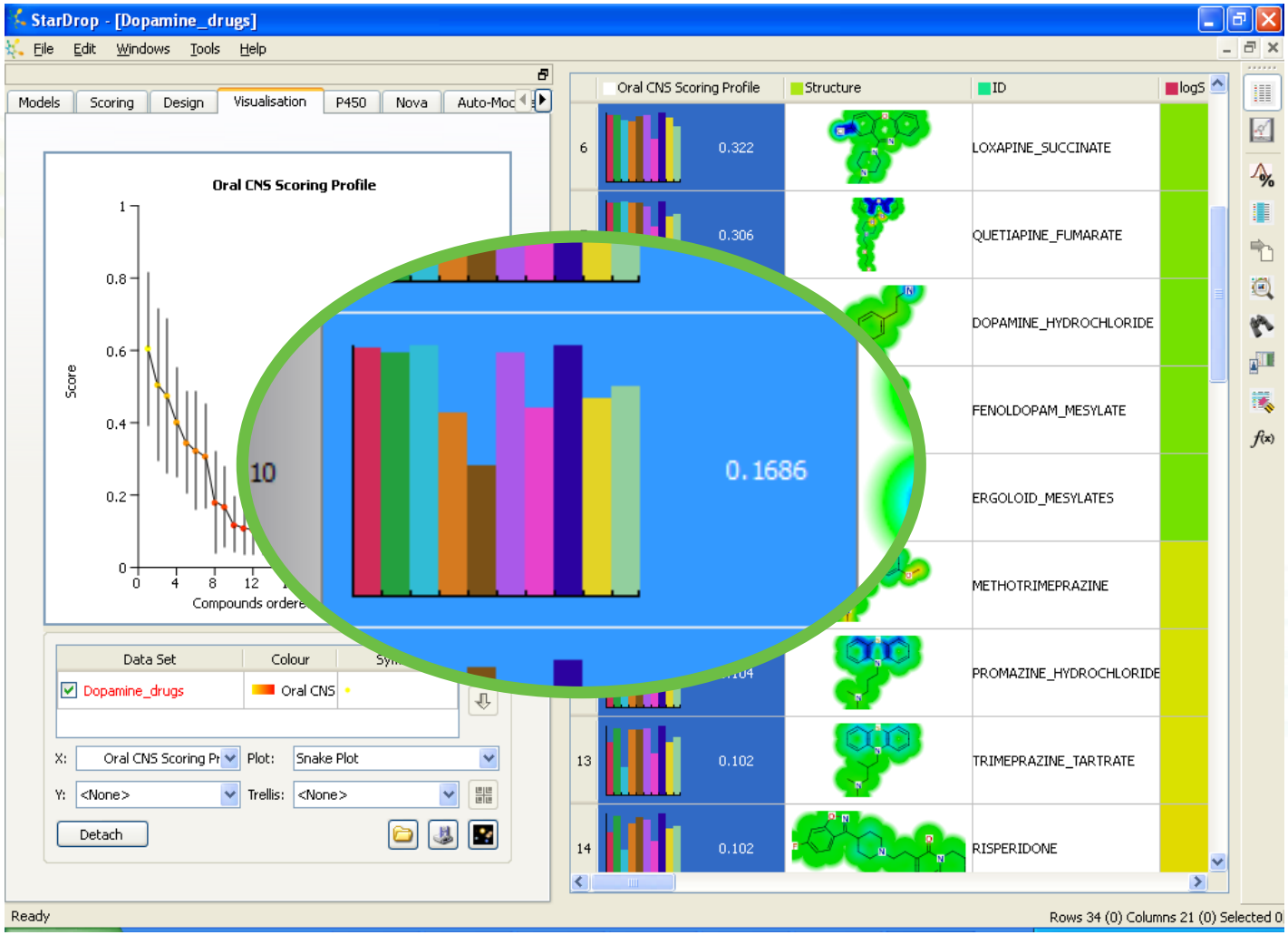
Bottom 50% may be rejected with confidence



© 2013 Optibrium Ltd.

* Segall *et al.* (2009) Chem. & Biodiv. **6** p. 2144

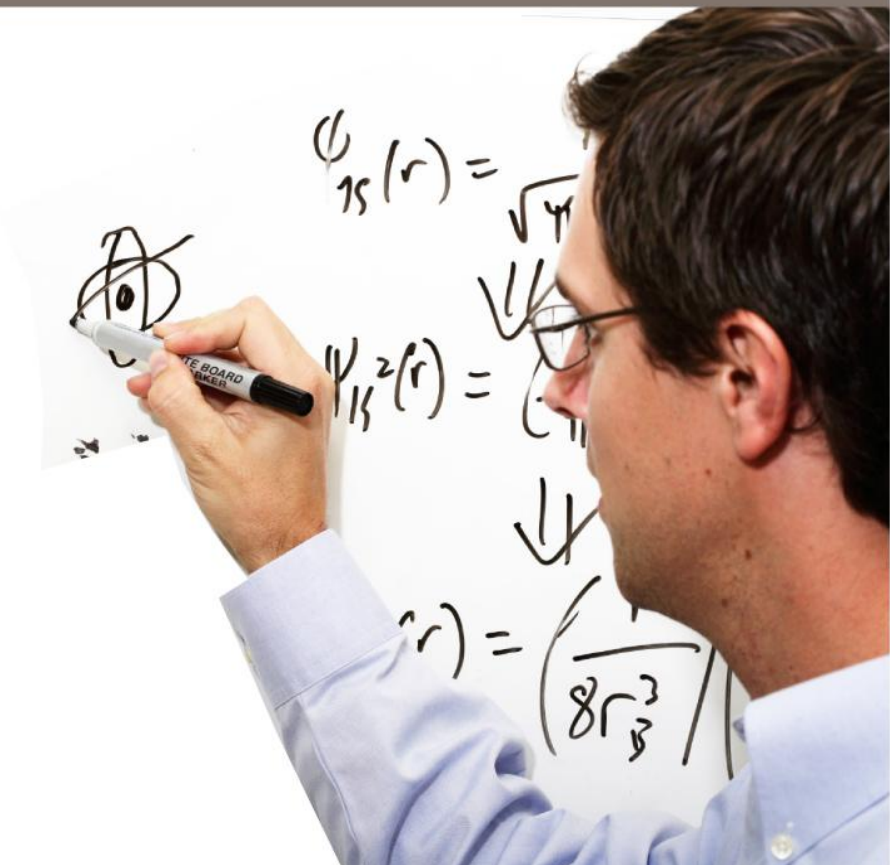
Provide Feedback on Influence of Properties
Guide redesign to improve chance of success



© 2013 Optibrium Ltd.

Case Study
Balancing Properties in Lead Optimization

optibrium



Case Study*
Goal: Orally dosed compound against CV target

- *In vitro* data generated on ~150 compounds
  - Potency
  - Selectivity
  - Solubility
  - Microsomal stability (rat and human: RLM and HLM)
- Original process focused on potency and selectivity, filtering compounds that did not meet requirements
- Results
  - Low but prolonged activity after IP dosing
  - No correlation between *in vitro* and *in vivo* potency
  - Problems with solubility and metabolic stability

© 2013 Optibrium Ltd.
*Segall *et al.,* Expert Opin. Drug. Metab. Toxicol., **2** pp. 325-37 (2006)

Comparison of Strategies

Potency and selectivity
No uncertainty - filter

	Name	pIC50	Selectivity (log)
1	XXX322	6.49	1.36
2	XXX313	6.8	1.14
3	XXX137	6.72	1.24
4	XXX540	6.7	1.17
5	XXX572	6.68	1.05
6	XXX541	6.66	0.94
7	XXX160	6.64	1.14
8	XXX326	6.64	1.33
9	XXX502	6.18	1.13
10	XXX292	6.28	1.22
11	XXX318	6.2	1.21
12	XXX537	6.4	0.98
13	XXX280	6.23	1.13
14	XXX282	6.16	0.95
15	XXX104	6.3	1.21
16	XXX295	6.25	0.99
17	XXX582	6.01	1.07
18	XXX561	6.92	0.84
19	XXX560	6.41	0.72
20	XXX133	6.38	0.58
21	XXX573	6.34	0.26
22	XXX293	6.3	0.79
23	XXX023	6.3	0.54
24	XXX294	6.28	0.89
25	XXX649	6.28	0.57
...			
50	XXX015	5.95	0.42
51	XXX136	5.94	0.46
52	XXX027	5.92	0.59
53	XXX284	5.9	0.77
54	XXX323	5.9	0.95
55	XXX017	5.9	0.36
56	XXX025	5.89	0.71
57	XXX319	5.89	0.56
58	XXX289	5.89	1.14
59	XXX297	5.88	0.5
60	XXX574	5.88	0.18

Potency and selectivity
Consider uncertainty

	Name	pIC50	Selectivity (log)
1	XXX326	6.64	1.33
2	XXX137	6.72	1.24
3	XXX322	6.49	1.36
4	XXX313	6.8	1.14
5	XXX540	6.7	1.17
6	XXX160	6.64	1.14
7	XXX572	6.68	1.05
8	XXX104	6.3	1.21
9	XXX292	6.28	1.22
10	XXX541	6.66	0.94
11	XXX561	6.92	0.84
12	XXX318	6.2	1.21
13	XXX537	6.4	0.98
14	XXX280	6.23	1.13
15	XXX502	6.18	1.13
16	XXX295	6.25	0.99
17	XXX294	6.28	0.89
18	XXX278	6.24	0.9
19	XXX282	6.16	0.95
20	XXX293	6.3	0.79
21	XXX560	6.41	0.72
22	XXX582	6.01	1.07
23	XXX289	5.89	1.14
24	XXX879	6.13	0.76
25	XXX133	6.38	0.58
...			
50	XXX316	5.63	1.02
51	XXX319	5.89	0.56
52	XXX655	6.11	0.34
53	XXX518	5.78	0.67
54	XXX110	6.12	0.32
55	XXX136	5.94	0.46
56	XXX062	5.98	0.42
57	XXX297	5.88	0.5
58	XXX186	6.27	0.21
59	XXX015	5.95	0.42
60	XXX315	5.8	0.51

All properties
Consider uncertainty

	Name	pIC50	Selectivity (log)	Expt. Solubility	Expt. HLM	Expt. RLM
1	XXX572	6.68	1.05	136	36.5	85.6
2	XXX518	5.78	0.67	148	4.03	38
3	XXX582	6.01	1.07	132	84.1	29.9
4	XXX295	6.25	0.99	146	63	77
5	XXX321	6	0.87	193	55.8	71.9
6	XXX502	6.18	1.13	127	95.6	64.6
7	XXX292	6.28	1.22	192	89	88
8	XXX274	5.81	0.89	124	91.9	49.2
9	XXX025	5.89	0.71	136	54.2	77.5
10	XXX280	6.23	1.13	165	83.6	76
11	XXX316	5.63	1.02	190	57.8	78.1
12	XXX278	6.24	0.9	144	88.9	70.8
13	XXX294	6.28	0.89	196	87	92
14	XXX282	6.16	0.95	185	78.4	84.6
15	XXX293	6.3	0.79	111	97.8	81.7
16	XXX319	5.89	0.56	178	60.3	87.9
17	XXX284	5.9	0.77	185	70	71.3
18	XXX111	5.73	0.6	116	59.9	93.9
19	XXX289	5.89	1.14	103	95.3	66.7
20	XXX277	5.85	0.98	194	91.4	83.8
21	XXX313	6.8	1.14	8.95	71.4	53.5
22	XXX517	5.82	0.85	137	90	95.3
23	XXX160	6.64	1.14	23.7	80.9	29.9
24	XXX468	5.69	0.97	118	82.1	87.7
25	XXX537	6.4	0.98	6.59	75.8	47

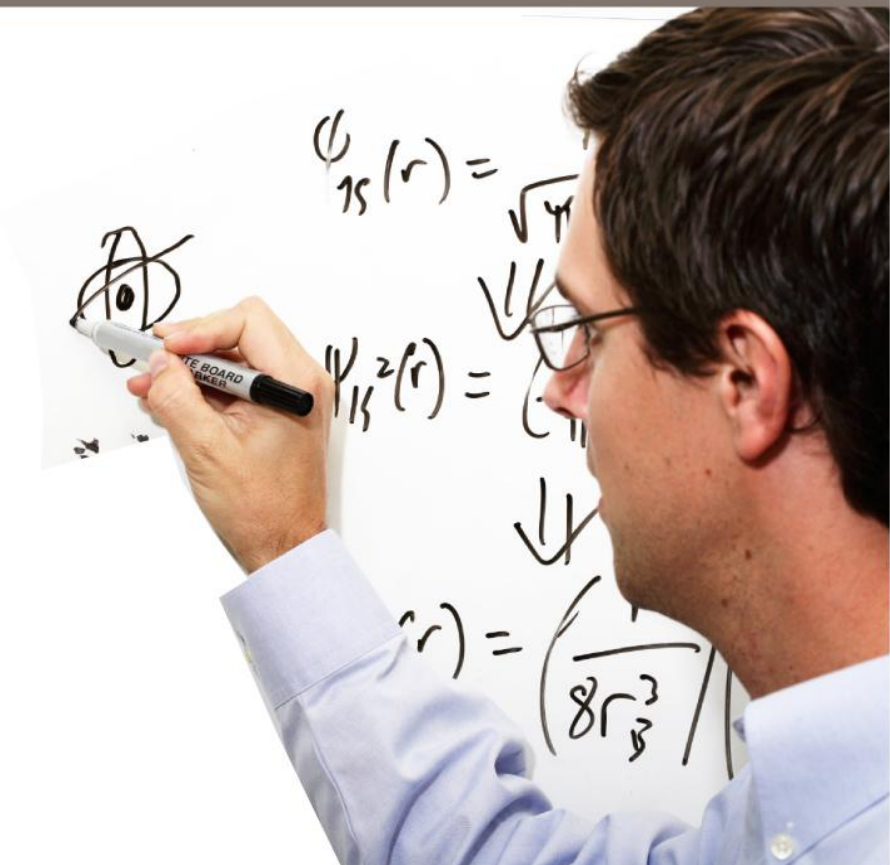
- New series identified with oral bioavailability and efficacy
- New direction for project

© 2013 Optibrium Ltd.

Finding Multi-Parameter Rules for Drug Discovery

Patent pending

optibrium



The Next Challenge

- Property criteria and importance are usually defined by expert opinion
  - What if we don't know what criteria to use for our project's objective?
- Can we **automatically** find scoring profiles with which to identify compounds with improved chance of success based on historic data
  - Any drug discovery objective, e.g. clinical, PK, toxicity...
  - Once developed, profile can be applied prospectively to find new compounds
- Identify the most important data with which to distinguish between successful and unsuccessful compounds
  - Any data can be used as input, calculated or experimental
- Results must be interpretable and modifiable
  - Avoid black boxes
  - Synergy between computer and experts

© 2013 Optibrium Ltd.

What is a Rule?

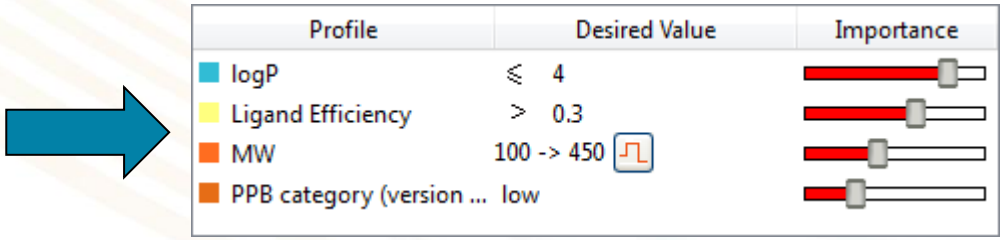
- A **Rule** is a set of property criteria that in **combination** identify 'good' compounds, e.g.

logP < 4

Ligand efficiency > 0.3

100 < MW < 450

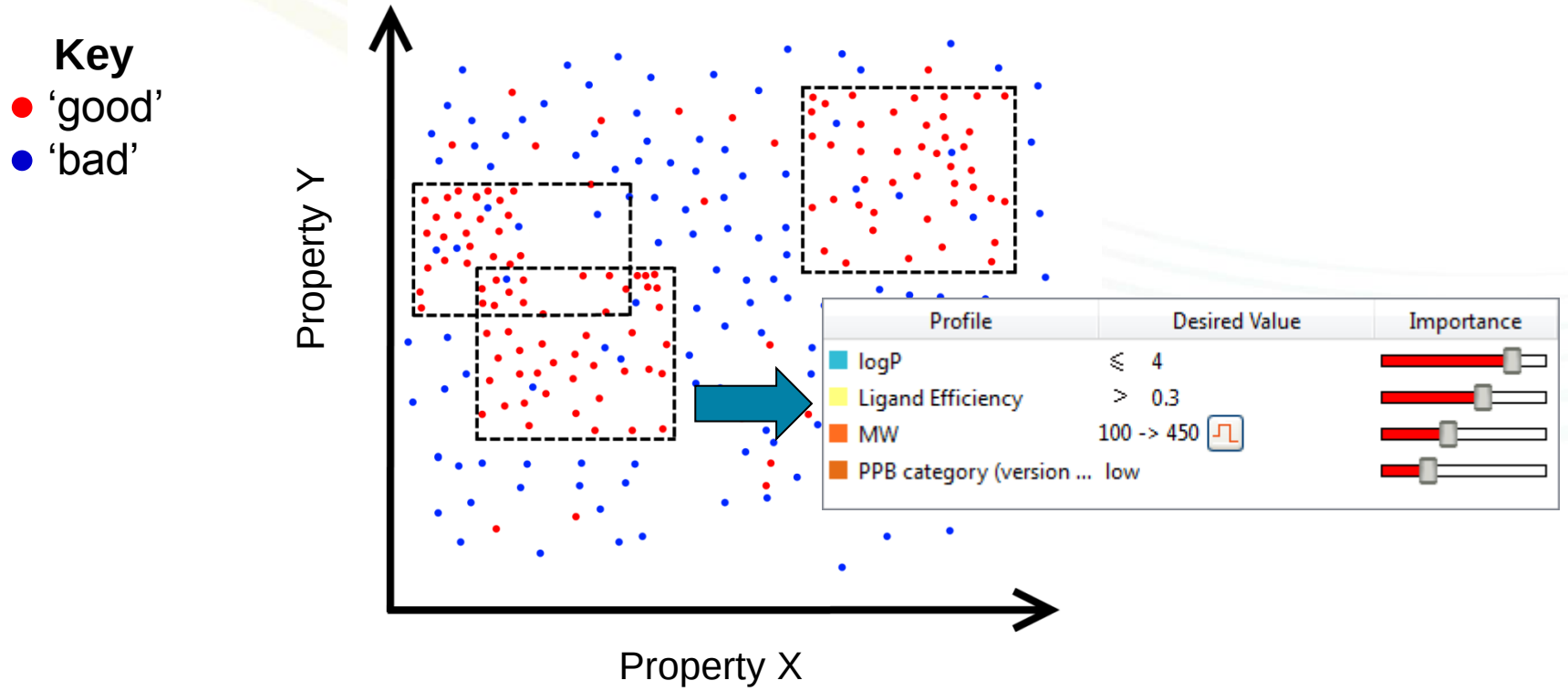
PPB category = low



© 2013 Optibrium Ltd.

What is a Rule?

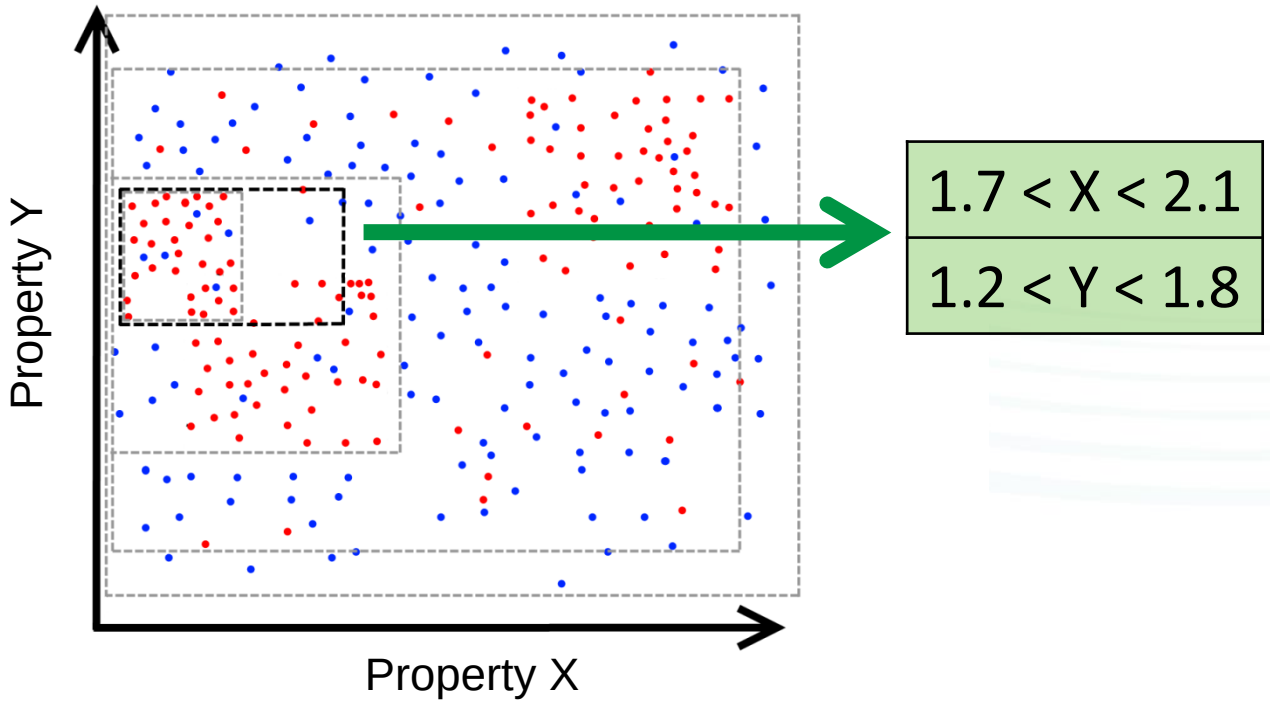
- A **Rule** is a box in multi-dimensional property space containing significantly more ‘good’ than ‘bad’ compounds
  - Equivalent to a scoring profile



© 2013 Optibrium Ltd.

Rule Induction with PRIM

- The Patient Rule Induction Method (PRIM) by Friedman and Fisher* is an effective way to find rules in multi-dimensional data

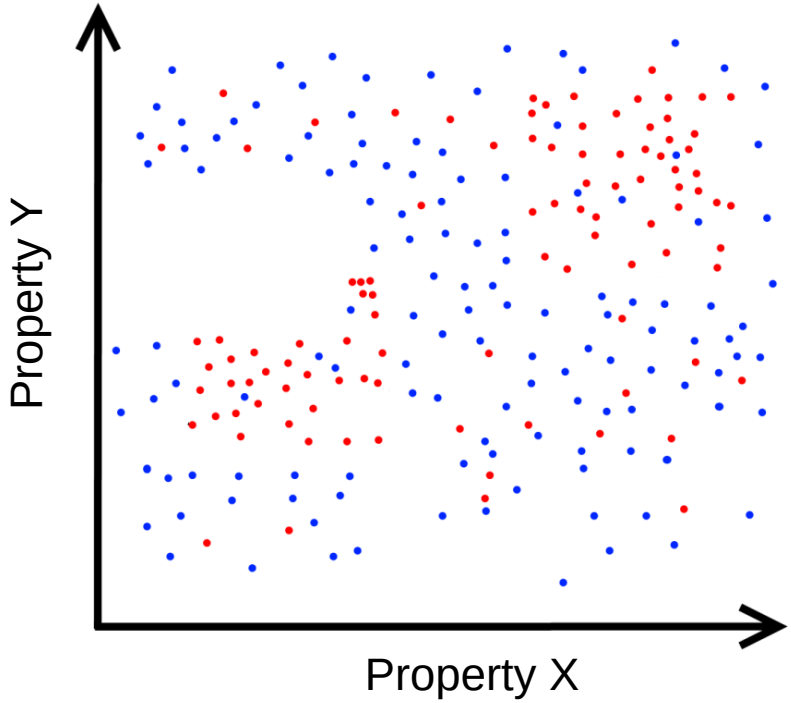


© 2013 Optibrium Ltd.

* Friedman & Fisher Stat. and Comp. **9**(2), p. 123 (1999)

Rule Induction with PRIM

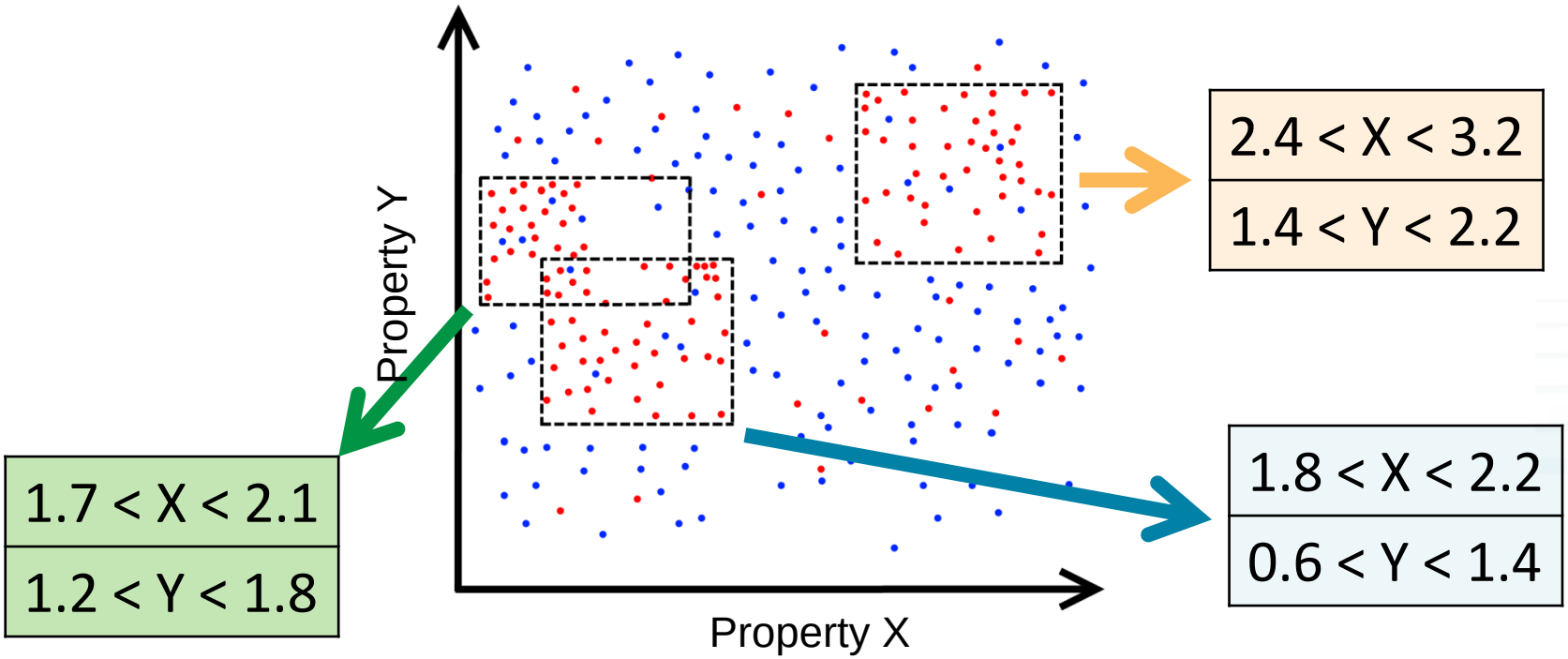
- After finding one box, we remove the box's compounds from the dataset and start over



© 2013 Optibrium Ltd.

Rule Induction with PRIM

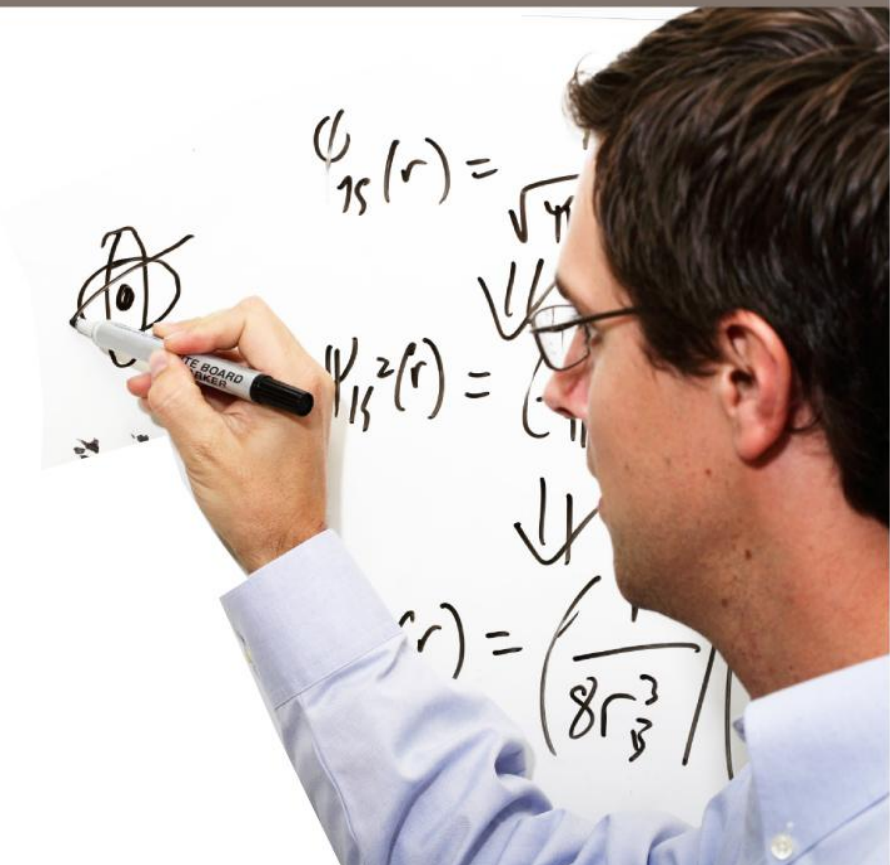
- The result is a series of boxes, each corresponding to an individual rule or scoring profile



© 2013 Optibrium Ltd.

Example Application
‘Drug-like’ Properties

optibrium

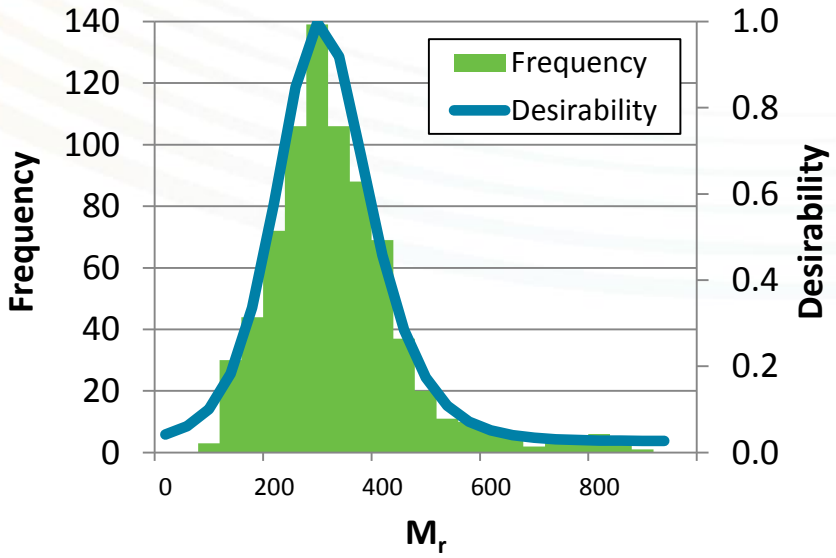


Drug-Like Properties
QED*

- Quantitative Estimate of Drug-Likeness (QED)*
- Combine values for 8 properties

MW	logP	HBD	HBA
PSA	ROTB	AROM	ALERT

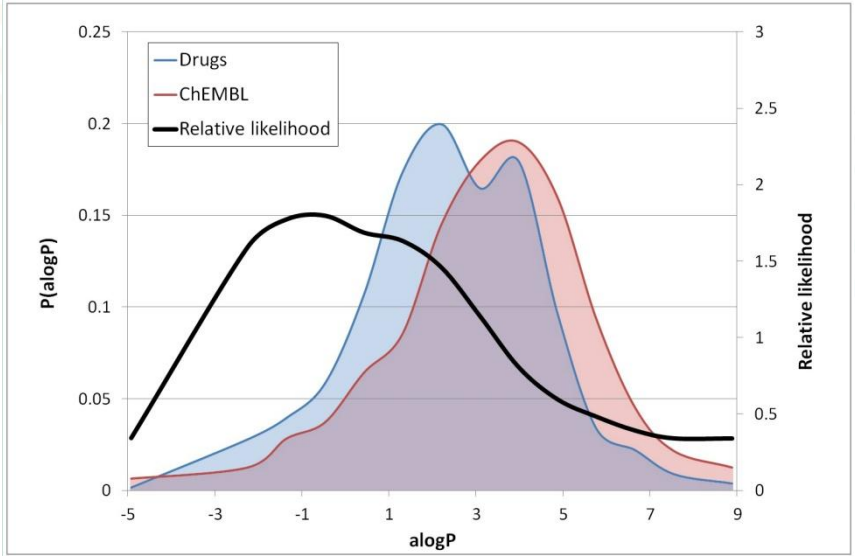
- For each individual property desirability function fitted to distribution for 771 **oral** drugs
- QED calculated as geometric mean of **individual** desirabilities



© 2013 Optibrium Ltd.
*Bickerton *et al.* Nature Chem. 4, pp. 90-98 (2012)

Example: Drug-Like Properties
RDL*

- Relative Drug Likelihood (RDL)*
- Compare characteristics of 771 **oral** drugs with 1000 randomly selected compounds from ChEMBL database
  - What property values increase likelihood of compound being an oral drug?
- Used same 8 properties as QED
- RDL calculated as geometric mean of **individual** likelihoods



© 2013 Optibrium Ltd.
*Yusof and Segall, Drug Discov. Today DOI: 10.1016/j.drudis.2013.02.008

Example: Drug-Like Properties
Rule Induction

- Rule induction applied to data set of 771 **oral** drugs and 1000 randomly selected compounds from ChEMBL
  - Random split 70:30 training:validation sets
- Used same 8 properties as QED and RDL as inputs
- 2 Rules:

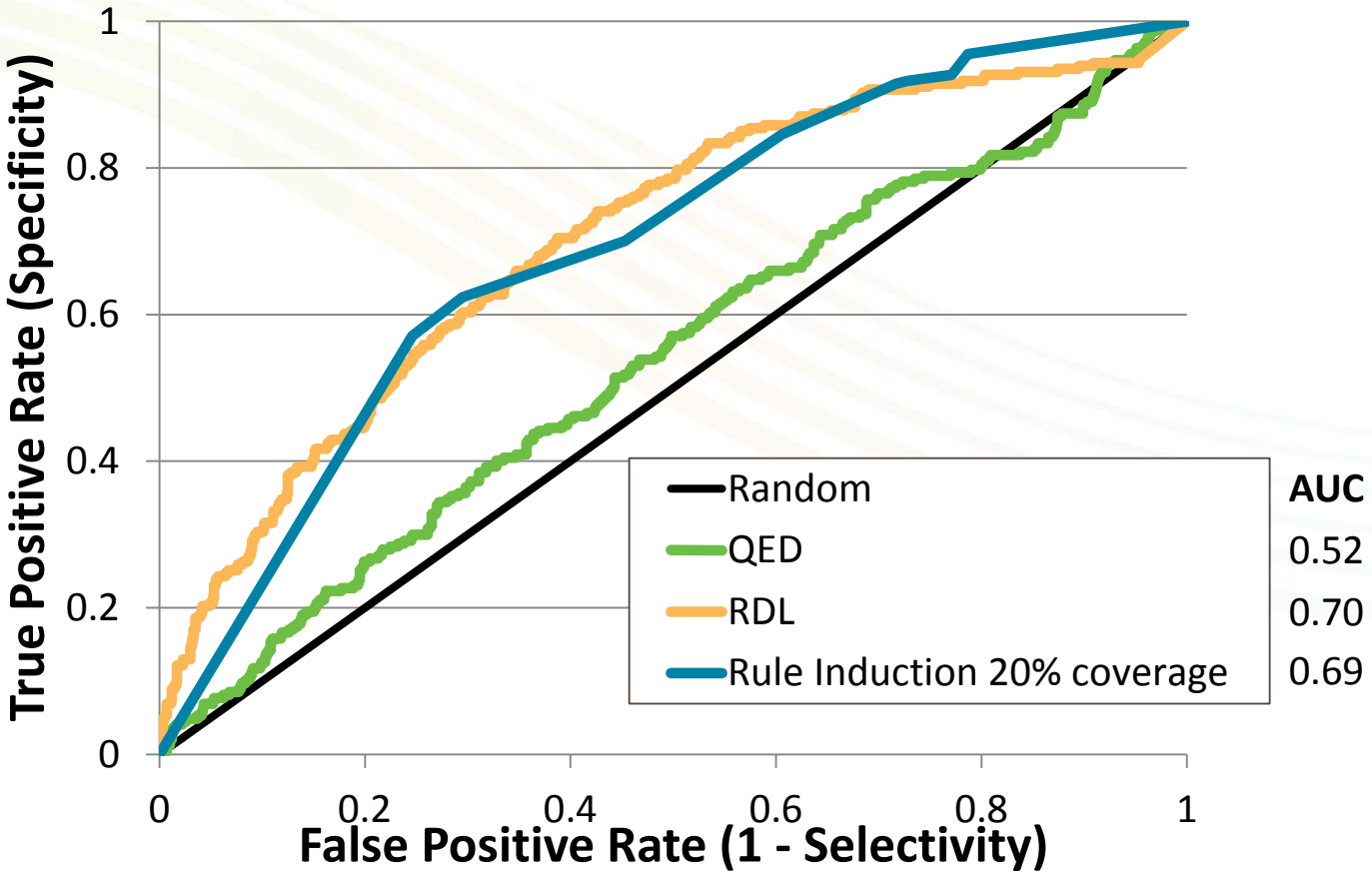
Profile	Desired Value	Importance
Rule 1		
MW	⩽ 444.855	
AROM	⩽ 1.01	
ALERTS	⩽ 1.01	

Profile	Desired Value	Importance
Rule 2		
ROTB	⩽ 4.04	
ALOGP	⩽ 2.727	

© 2013 Optibrium Ltd.

Example: Drug-Like Properties
Results

- Applied to **independent** test set of 247 oral drugs and 1000 compounds randomly selected from ChEMBL



© 2013 Optibrium Ltd.

Example: Oral CNS
Rule Induction

- Data set of 1191 drugs labelled as orally administered and CNS active or not
  - By approved route of administration and therapeutic indication (noisy)
- Divided into training (667), validation (286) and test (238) sets
- Calculated ADME properties from StarDrop™ used as input:

logP	Solubility (logS)	Human Intestinal Absorption category (HIA)
Blood-brain barrier penetration (BBB log)	Plasma protein binding category	P-gp substrate category

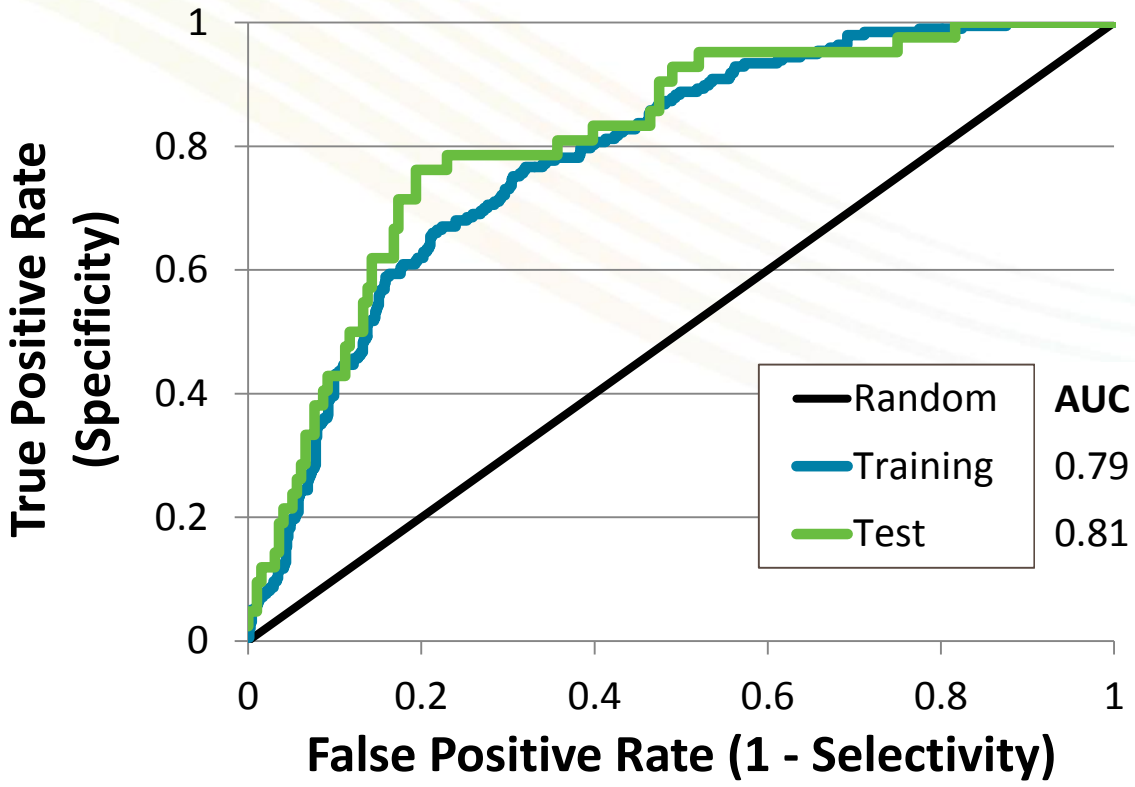
- Minimum coverage 20%

© 2013 Optibrium Ltd.

Example: Oral CNS
Results

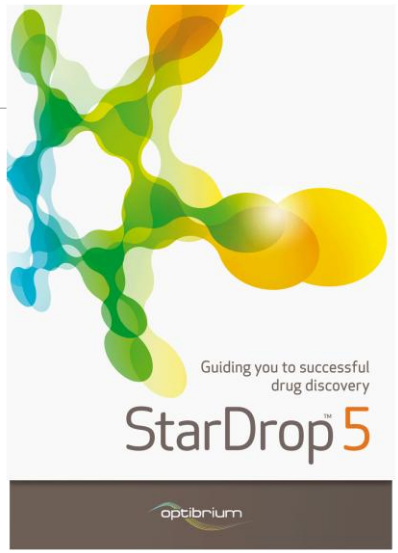
Profile	Desired Value	Importance
HIA category	+	
logS	> 1.00601	
BBB log([brain]:[blood])	> -0.424961	
logP	> 1.12432	
P-gp category	no	

Set	Mean Improvement (%)	Coverage (%)
Train	132	23
Val	123	20



© 2013 Optibrium Ltd.

Conclusion



- MPO is a powerful approach to select and design compounds with a high chance of success
  - Focus quickly on high quality compounds
- Rule Induction helps to guide the development of scoring profiles to select compounds for a drug discovery objective
  - Apply to any objective
  - Use experimental or calculated data
  - Not black box – synergy between computer and expert
- Identify most important data to guide selection of successful compounds
  - Optimise screening strategy and prioritise experimental resources
- For more information:
  - matt.segall@optibrium.com
  - www.optibrium.com

© 2013 Optibrium Ltd.

Acknowledgements

- Tatsu Hashimoto – MIT

- Opribrium team, including:
  - Iskander Yusof
  - Ed Champness
  - Chris Leeding
  - James Chisholm
  - Hector Garcia Martinez

© 2013 Optibrium Ltd.