HelmholtzZentrum münchen
German Research Center for Environmental Health

Early ADME/Tox predictions: toy or tool?

Igor V. Tetko

Helmholtz Zentrum München - German Research Center for Environmental Health (GmbH)
Institute of Bioinformatics & Systems Biology

Obernai, France, 21 June 2010

HELMHOLTZ ASSOCIATION

Helmholtz Zentrum München statistics

- Previous name (before 2008): GSF (Forschungszentrum für Umwelt und Gesundheit GmbH)
- Part of Helmholtz Network (2.35 Milliards Euro, 26500 people, 15 centers)
- Leading center for Environmental Health in Germany
- 25 institutes (1797 people, ca 700 scientists & 300 PhD students)*
- 70 contracts with EU
- Strong IPR and management support

- Institute for Bioinformatics & Systems Biology
  - 50 peoples, strong expertise in *in silico* data analysis, machine learning methods, software development, data dissemination (Web, Internet)

*January 2008


HelmholtzZentrum münchen
German Research Center for Environmental Health

Layout of presentation

Productivity of R&D companies

Importance of ADMETox parameters

Overview of eADMETox properties/data

Applicability Domain challenges

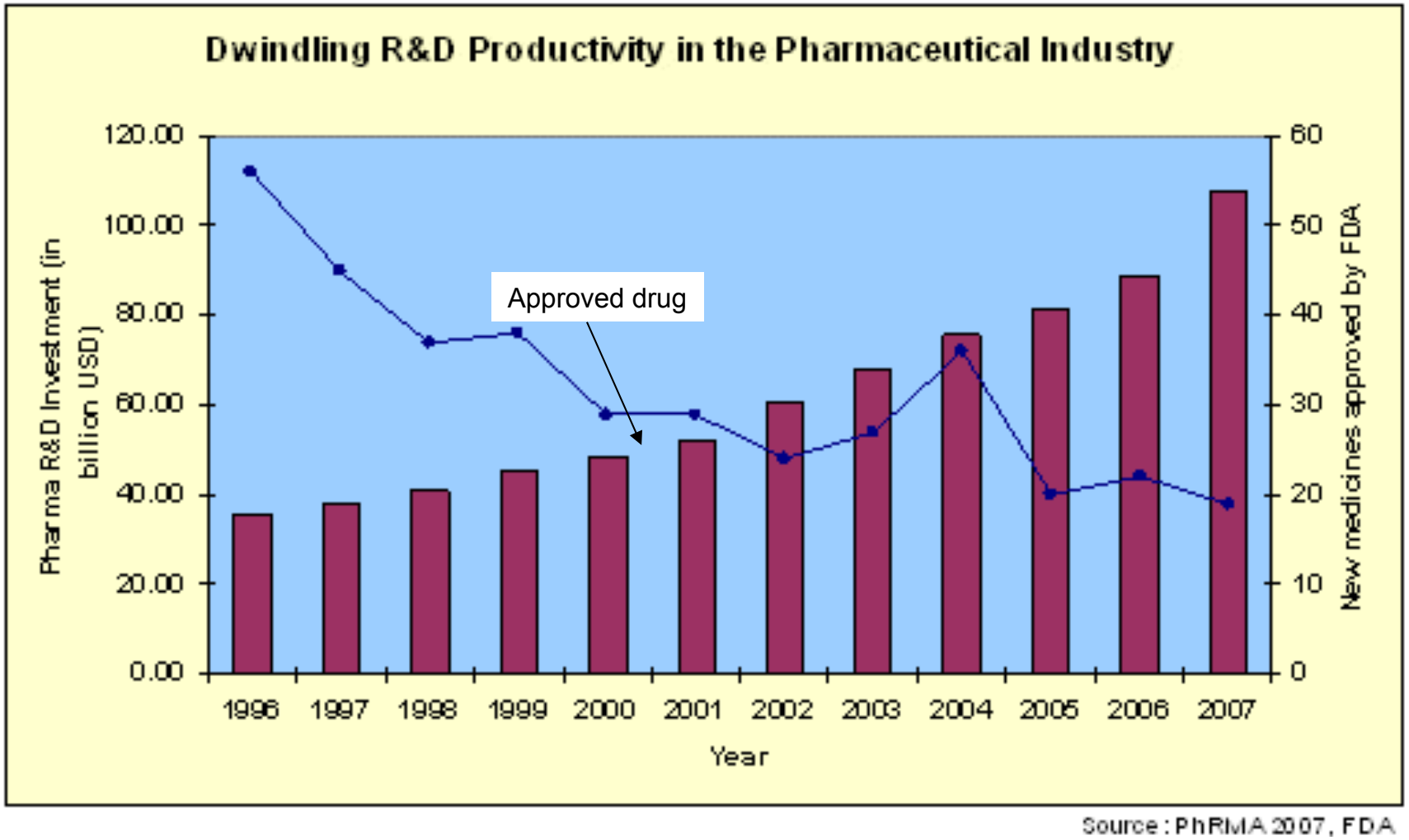
- LogP benchmarking study
- AD for qualitative models – AMES test

Data integration

OCHEM – On-line CHEmical database & Modeling environment

Conclusions

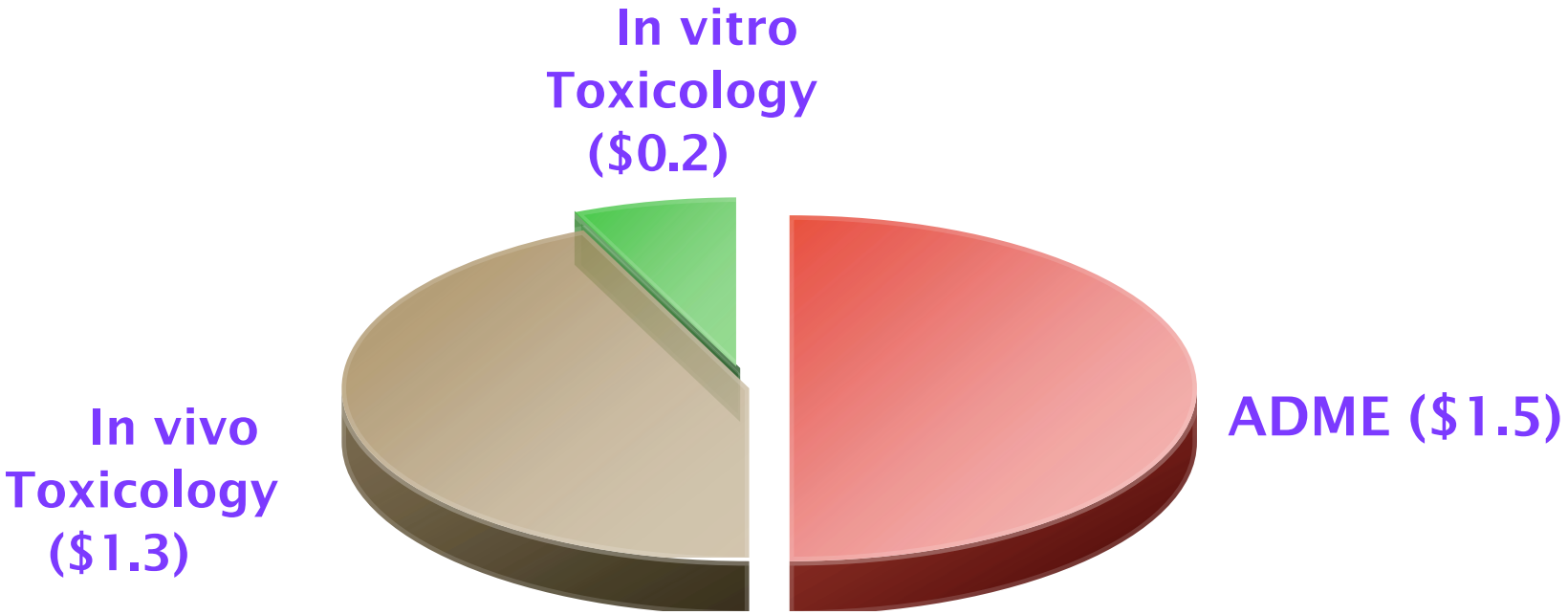
Pharma R&D Cost and Productivity: Fewer drugs, more expenditure



2008 – 24(3*); 2009 – 25(6*)

*Biological license applications

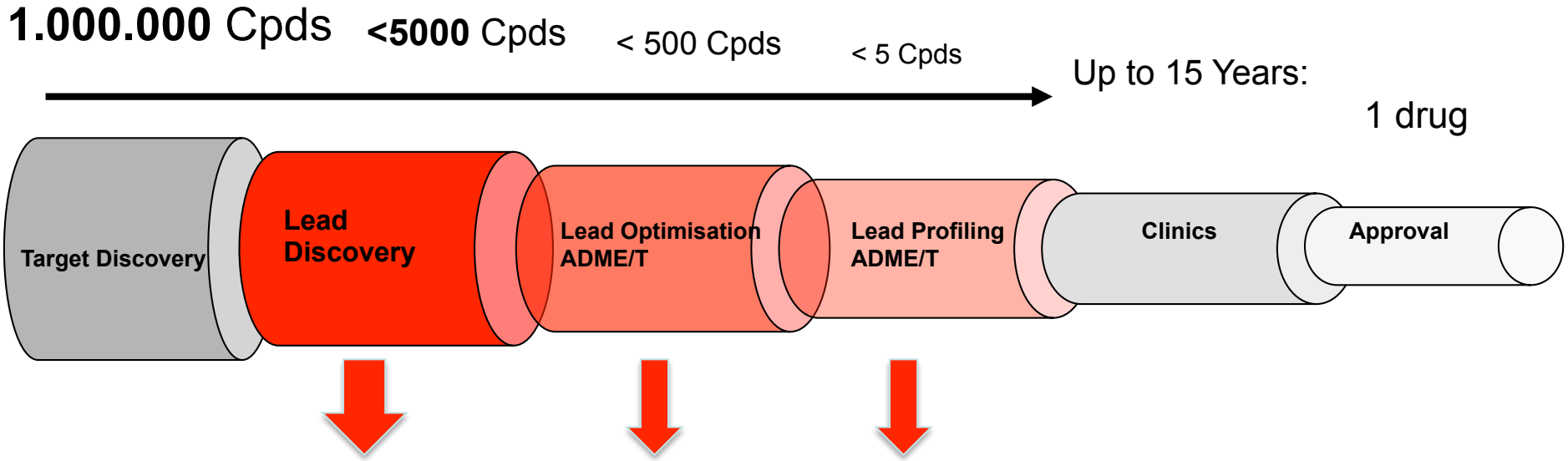
Potential ADME/T market (US $ billions)[1]



It will grow up to US$ 4.4 billion up to 2012[2]

1) Razvi, E.S. *Drug and Market Development* (2003).
2) http://www.researchandmarkets.com/reports/c84850

Pharma R&D: Cost and Productivity issues
Compound numbers



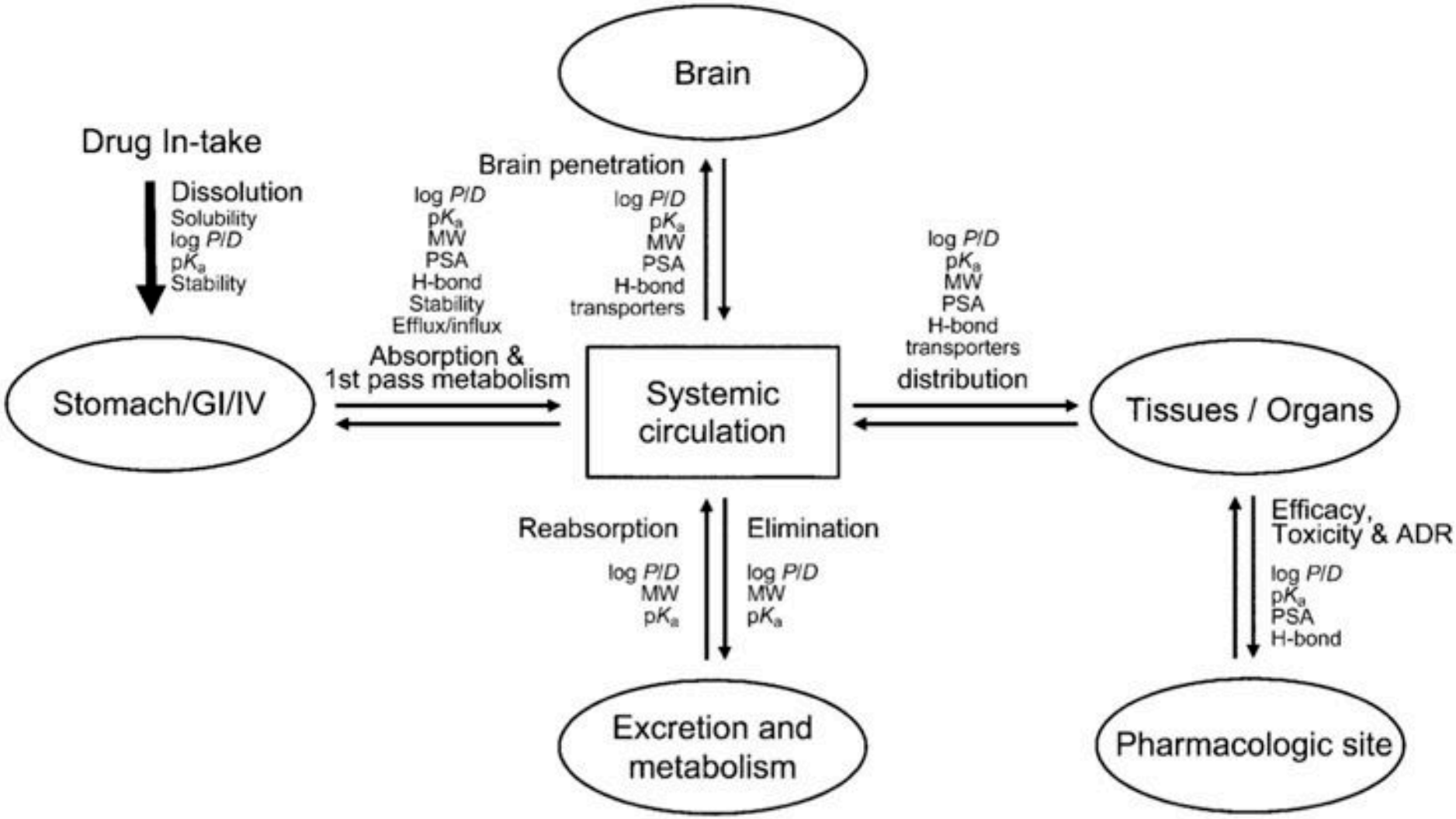
In vitro and in vivo ADME/T property determination:
Millions of screens for solubility, stability, absorption, metabolism, transport, reactive products, drug interactions, etc etc

Preclinics Costs: > $300m PER COMPOUND to reach approval

ADME/T

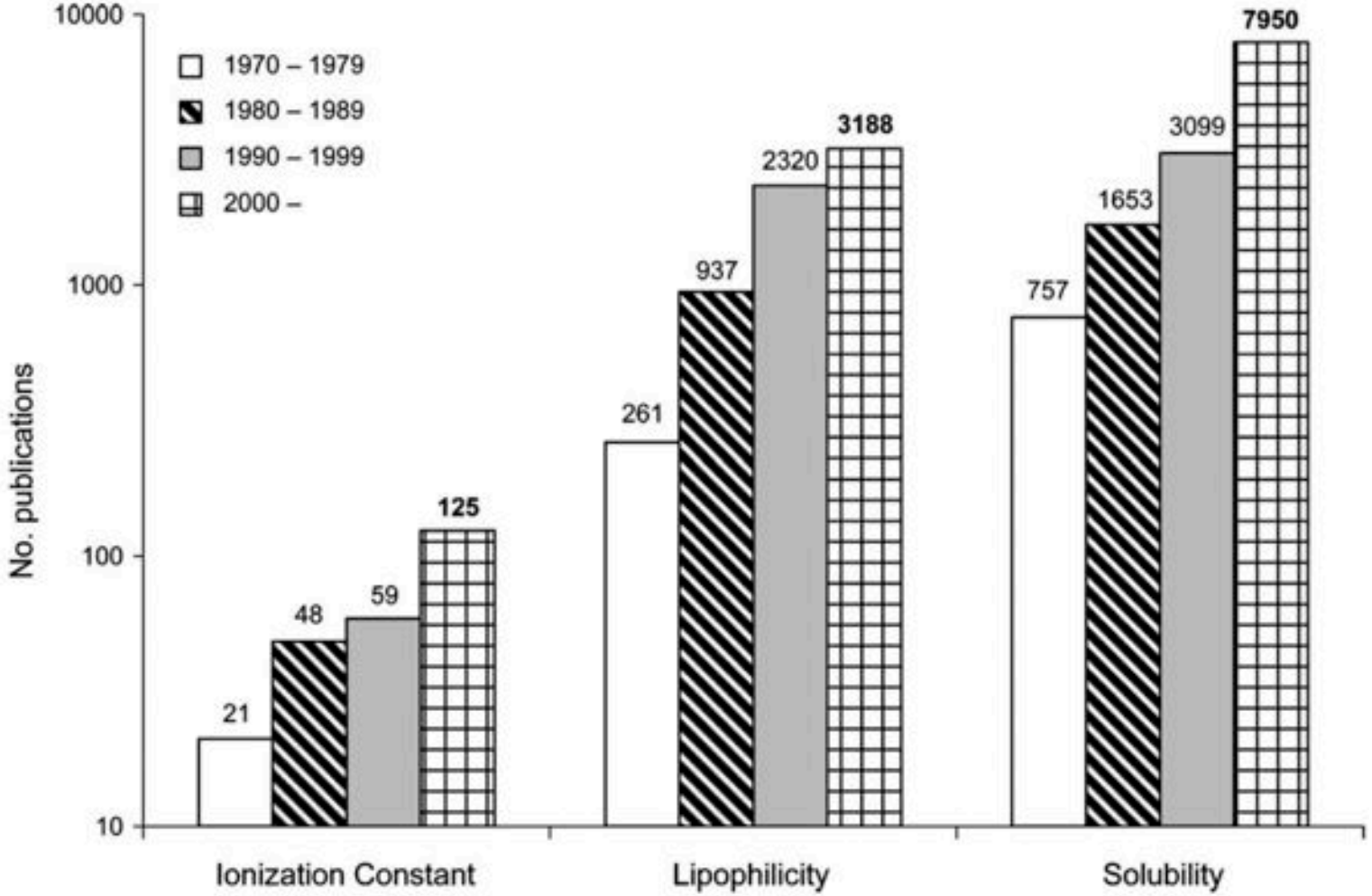
Absorption enters organism (by oral administration)	Size, lipophilicity, solubility, ionization, permeability, active transport
Distribution distributed between blood/ plasma/tissues (e.g. brain)	Affinity to different tissues, permeability, active transport
Metabolism bio-converting	Affinity to different enzymes
Elimination mechanisms and pathways for excretion of drugs	Active transport, size, lipophilicity, ionization, permeability (also for metabolites)
Toxicity undesired interactions of drug or its metabolites	Presence of toxicological pharmacophores, liophilicity

Interplay of physico-chemical properties with *in vivo* pharmacological activities/data



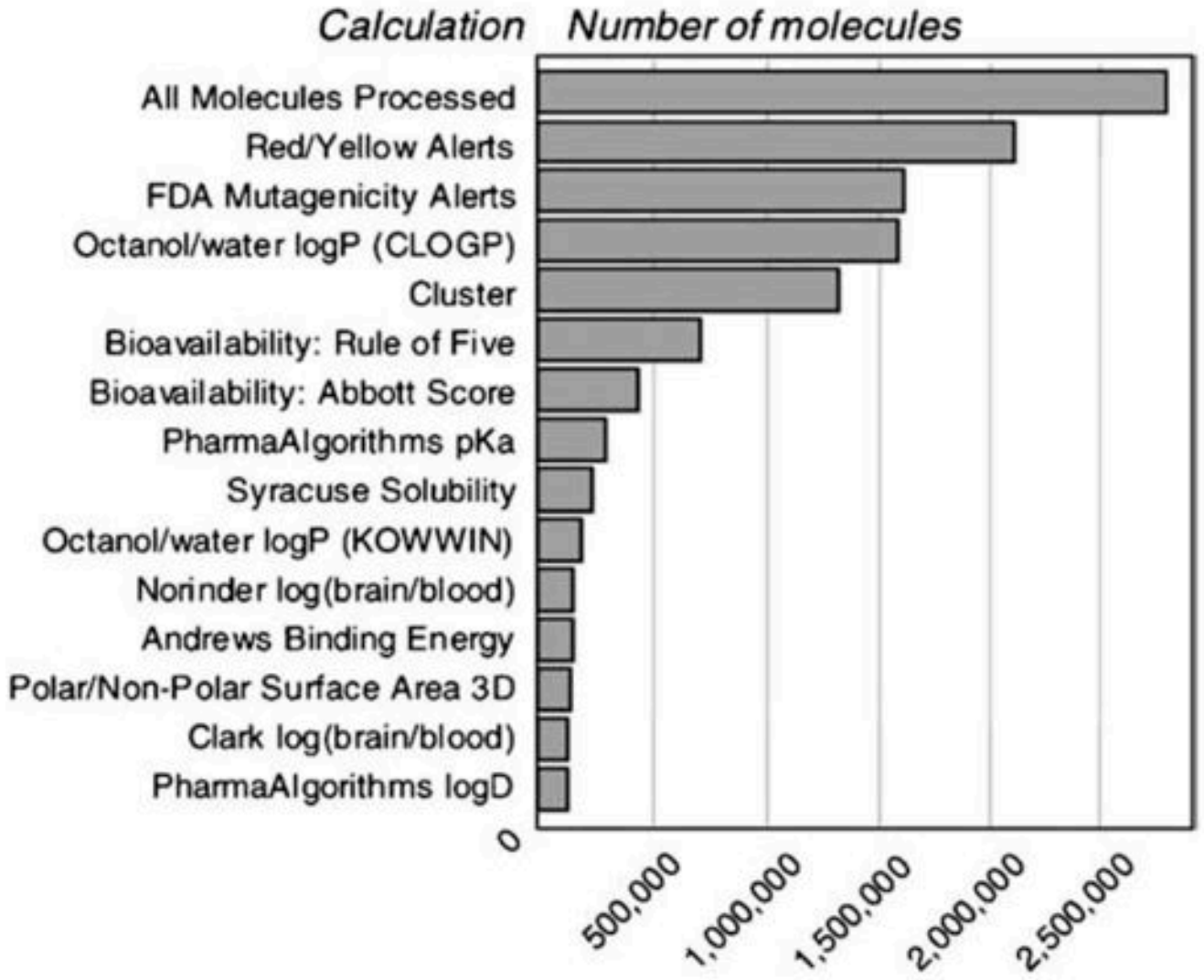
Wang & Shkolnik, Chem. & Biodiversity, 2009, 6, 1887.

Interest in Phys-Chem properties



Wang & Shkolnik, Chem. & Biodiversity, 2009, 6, 1887.

Number of molecules processed at the Abbot site through the various algorithms available on the property calculation web page



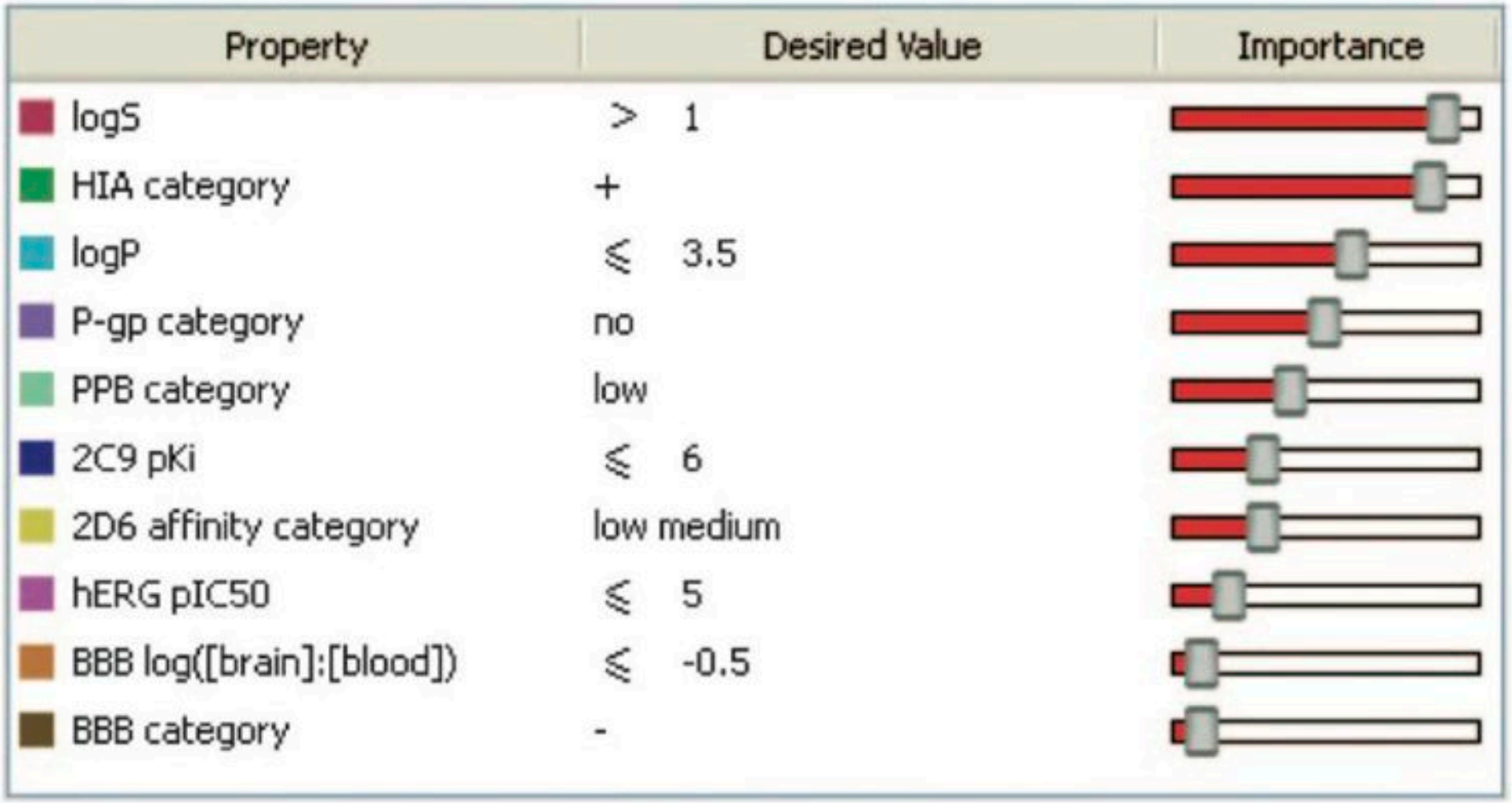
Y. C. Martin, QSAR Comb. Sci., 2006, 25, 1192.

Properties Used to Define Drug-Likeness

Property	Drugs	CNS-Drugs	Leads	Fragments
MW	<500	<450	<400	<300
logP	<5	0-4	<4	<3
HA	<10		<8	<3
HD	<5		<4	<3
$logD_{7.4}$	1-3	1-4		
PSA	<140	<80	<120	<90

van de Waterbeemd, Chem. Biodiver. 2009, 1760.

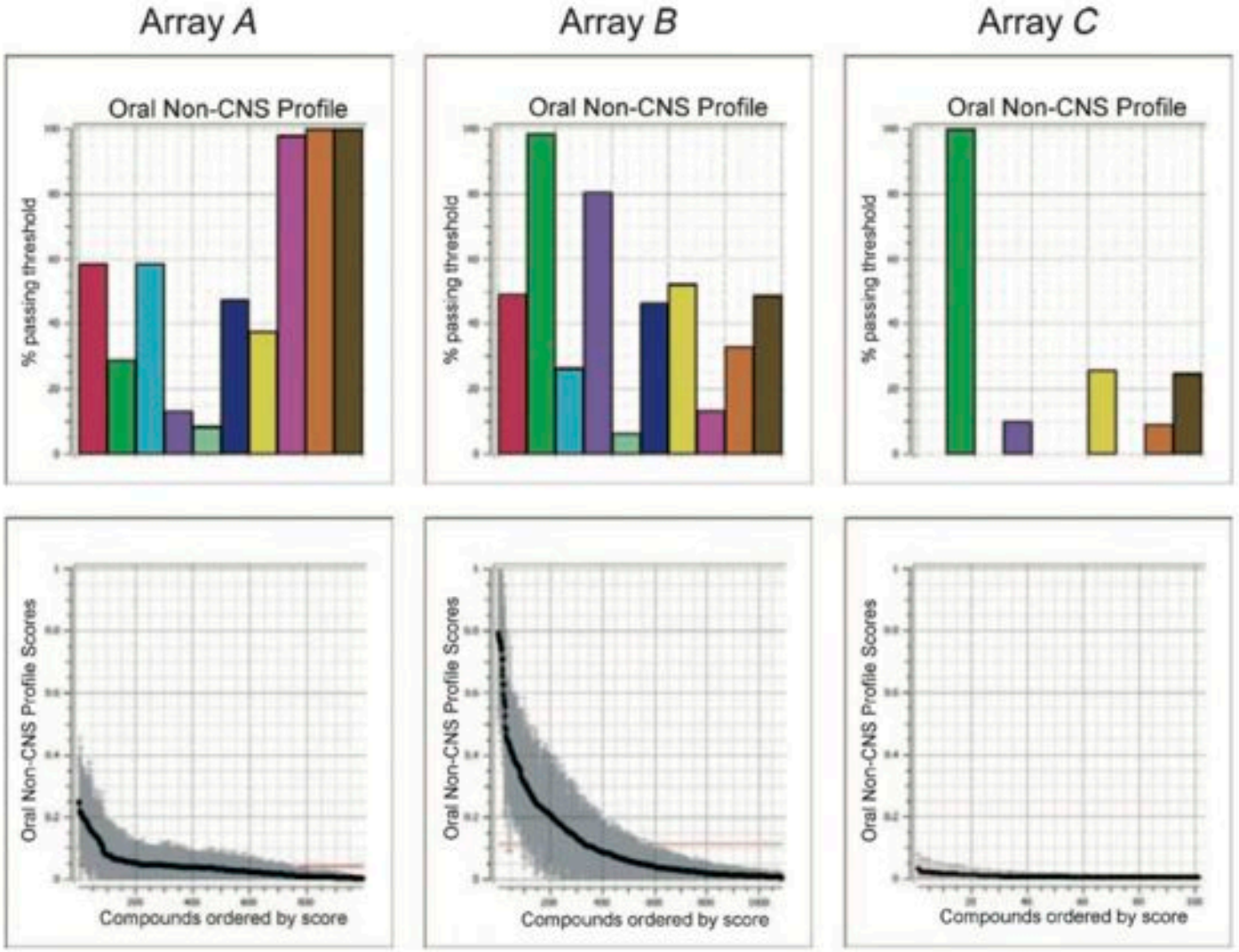
Profiling of chemical compounds (Optibrium Ltd)



Property	Desired Value	Importance
logS	> 1	
HIA category	+	
logP	≤ 3.5	
P-gp category	no	
PPB category	low	
2C9 pKi	≤ 6	
2D6 affinity category	low medium	
hERG pIC50	≤ 5	
BBB log([brain]:[blood])	≤ -0.5	
BBB category	-	

Segal et al, Chem. Biodiver., 2009, 6, 2144.

Profiling of chemical compounds (Optibrium Ltd)



Segal et al, Chem. Biodiver., 2009, 6, 2144.

ADMET parameters

Absorption
Caco-2 (colorectal carcinoma cells)
MDCK (Madin-Darby Canin Kidney)
PAMPA (parallel artificial membrane permeability assay)
Human Intestinal Absorption
%FA (% of absorbed drug mass)
Active transporters (P-glycoprotein, multi drug resistance protein – efflux, peptide, amino-acid transporters – absorption)

Metabolism
CYP450
Aldehyde/Alcohol dehydrogenases
Hydrolases, oxidases, esterases
Microsomes, hepatocytes
Genetically modified cell lines to express single CYP450

Distribution
BBB (Blood-Brain Barrier)
PPB (Plasma protein Binding)
HSA (Human Serine Albumin) binding
Tissue partitioning
Volume of distribution=dose/C_0 (C_0 is initial concentration of a drug in plasma)

Elimination
Route of elimination (renal, liver (bile->faeces, metabolism)

Toxicity
AMES test
DDI (inhibition/activation of CYP450)
hERG (human ether-a-go-go-related gene) potassium channel inhibition
Toxicity alerts

Physico-chemical properties

In vitro:

logP

logD

Solubility in water

Solubility in DMSO

pKa

Solubility in simulated intestinal fluid

Absorption:

In vitro:

CaCo-2

MDCK

PAMPA

In vivo:

%FA

Distribution:

In vitro:

PPB

HSA

BBB

In vivo:

BBB – animal models

Tissue partitioning

Metabolism:

In vitro:

CYP450

Genetically engineered cell lines to study individual CYP

Microsomes

Hepatocytes

In vivo:

MS analysis

Toxicity

In vitro:

AMES mutagenicity

hERG toxicity

In vivo:

Animal models (LD50)

ADMETox properties

Physico-chemical

Lipophilicity (logP/logD)

- ~ 20k (>100k)

Aqueous solubility

- ~ 10k (~100k)

pKa

- ~ 10k (~100k ?)

Solubility in DMSO

- ~1k (>100k)

Biological properties

%FA (Fraction Absorbed) ~1k

Blood-Brain barrier ~1k

CYP450 affinities ~10k

Transporters (PgP) ~10k

Ion channels (hERG) ~10k

Microsomes ~100k

Hepatocytes ~10k

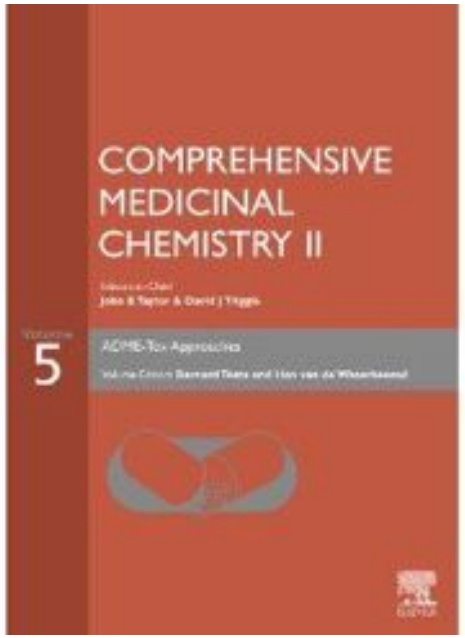
VD ~300

Available sources: WOMBAT, Symyx, CHeMBL, PHYSPROP, ChemSpider, OCHEM

Additional readings:

In Silico ADME Prediction: Data, Models, Facts and Myths, Lombardo, F.; Gifford, E.; Shalaeva, M.Y. Mini Reviews in Medicinal Chemistry, 2003, 3, 861-875

Comprehensive Medicinal Chemistry II: In silico tools in ADMET; Testa, B., van de Waterbeemd, H., Eds.; Elsevier: 2006; Vol. 5.



Chemistry & Biodiversity, vol. 6, 2009.

What are the goals of modeling?

Decrease number of experimental measurements by substitution of them with computational predictions.

This can be achieved when computational accuracy of models is similar (or better!) to that of experimental measurements.

Can we achieve it?

"One can not embrace the unembraceable."

Possible: 10^{60} - 10^{100} molecules theoretically exist

Achievable: 10^{20} - 10^{24} can be synthesized now by companies (weight of the Moon is ca 10^{23} kg)

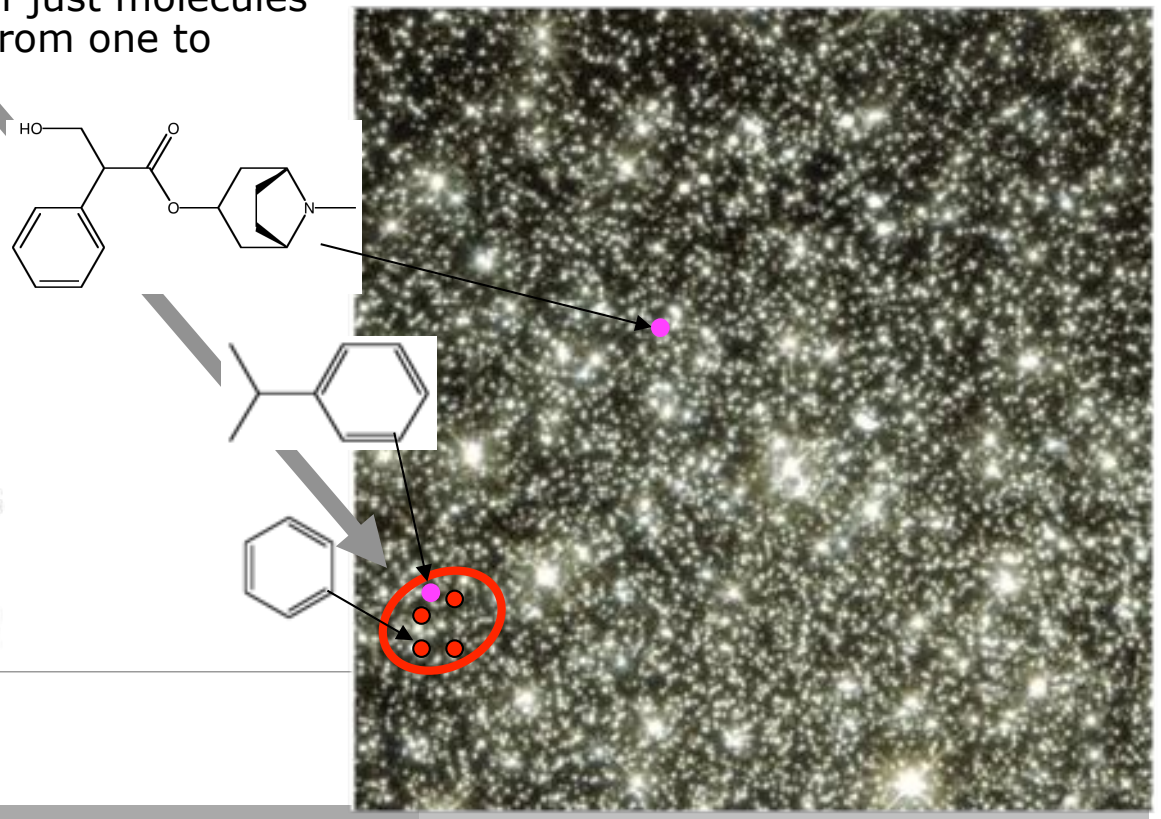
Available: $2*10^{7}$ molecules are on the market

Measured: 10^{2} - 10^{5} molecules with ADME/T data

Problem: To predict ADME/T properties of just molecules on the market we must extrapolate data from one to 1,000 - 100,000 molecules!

Kozma Prutkov

10^{80} atoms in the Universe



There is a need for methods which can estimate the accuracy of predictions!

Current dogma about prediction of physico-chemical properties

- Prediction of physico-chemical properties, in particular **log P**, is simple
- There is no need to measure them
- We have enough number of good computational methods

Is this true?

Statistics of logP benchmarking

30 (18) methods - major commercial providers and public software

Public dataset:
N=266 molecules

in house data:
N=95809 molecules from Prizer
N=889 molecules from Nycomed

Arithmetic Average Model (AAM):
mean log*P* was used as a prediction (one value for all molecules)

Rank III: models with errors (*RMSE*) **≥ AAM**, i.e. non-predictive
Rank I: models with *RMSE* identical or close to the best method
Rank II: remaining models

Mannhold et al, *J. Pharm. Sci.*, **2009**, *98*, 861-893.

Performance of algorithms for the public dataset

rank I – high accuracy predictions, RMSE ~ best model

rank II – good predictions, RMSE<AAM

rank III – low accuracy predictions, RMSE ≥ AAM

AAM = *base ("no model") model, R^2=0, it used just one logP value as predicted value for all 95809 or 882 molecules, respectively.*

Method	Star set (*N* = 223)					Non-Star set (*N* = 43)				
			% within error range					% within error range		
	RMSE	rank	<0.5	0.5-1	>1	*RMSE*	rank	<0.5	0.5-1	>1
AB/LogP	0.41	I	84	12	4	1.00	I	42	23	35
S+logP	0.45	I	76	22	3	0.87	I	40	35	26
ACD/logP	0.50	I	75	17	7	1.00	I	44	33	23
Consensus log *P*	0.50	I	74	18	8	0.80	I	47	28	26
CLOGP	0.52	II	74	20	6	0.91	I	47	28	26
VLOGP OPS	0.52	II	64	21	7	1.07	I	33	28	26
ALOGPS	0.53	II	71	23	6	0.82	I	42	30	28
MiLogP	0.57	II	69	22	9	0.86	I	49	30	21
XLOGP	0.62	II	60	30	10	0.89	I	47	23	30
KowWIN	0.64	II	68	21	11	1.05	I	40	30	30
CSlogP	0.65	II	66	22	12	0.93	I	58	19	23
ALOGP (Dragon)	0.69	II	60	25	16	0.92	I	28	40	33
MolLogP	0.69	II	61	25	14	0.93	I	40	35	26
ALOGP98	0.70	II	61	26	13	1.00	I	30	37	33
OsirisP	0.71	II	59	26	16	0.94	I	42	26	33
VLOGP	0.72	II	65	22	14	1.13	I	40	28	33
TLOGP	0.74	II	67	16	13	1.12	I	30	37	30
ABSOLV	0.75	II	53	30	17	1.02	I	49	28	23
QikProp	0.77	II	53	30	17	1.24	II	40	26	35
QuantlogP	0.80	II	47	30	22	1.17	II	35	26	40
SLIPPER-2002	0.80	II	62	22	15	1.16	II	35	23	42
COSMOFrag	0.84	II	48	26	19	1.23	II	26	40	33
XLOGP2	0.87	II	57	22	20	1.16	II	35	23	42
QLOGP	0.96	II	48	26	25	1.42	II	21	26	53
VEGA	1.04	II	47	27	26	1.24	II	28	30	42
CLIP	1.05	II	41	25	30	1.54	III	33	9	49
LSER	1.07	II	44	26	30	1.26	II	35	16	49
MLOGP (Sim+)	1.26	II	38	30	33	1.56	III	26	28	47
NC+NHET	1.35	III	29	26	45	1.71	III	19	16	65
SPARC	1.36	III	45	22	32	1.70	III	28	21	49
MLOGP(Dragon)	1.52	III	39	26	35	2.45	III	23	30	47
LSER UFZ	1.60	III	36	23	41	2.79	III	19	12	67
AAM	**1.62**	**III**	**22**	**24**	**53**	**2.10**	**III**	**19**	**28**	**53**
VLOGP-NOPS	1.76	III	1	1	7	1.39	III	7	0	7
HINT	1.80	III	34	22	44	2.72	III	30	5	65
GBLOGP	1.98	III	32	26	42	1.75	III	19	16	65

Our methodology is top-ranked in a recent benchmarking

- Benchmarking was done by Pfizer and Nycomed – no data were available to participants
- Our **ALOGPS** algorithm was top-ranked (according to the lowest RMSE errors)
- Several methods performed worse than making no prediction,

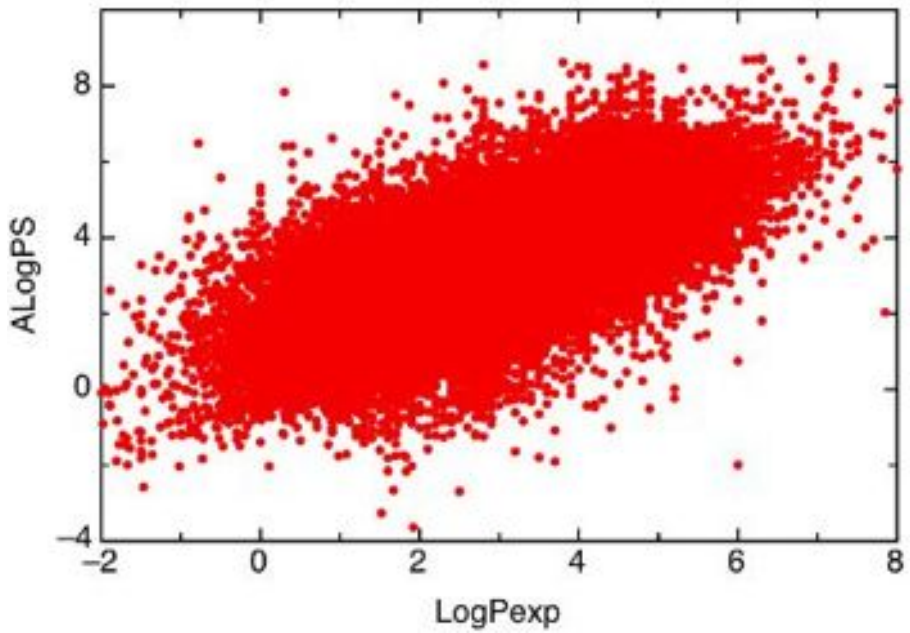
AAM *base ("no model") model, R^2=0, it used just one logP value as predicted value for all 95809 or 882 molecules, respectively.*

Method	Pfizer set (*N* = 95 809)							Nycomed set (*N* = 882)				
	RMSE	Failed[1]	rank	% in error range <0.5	0.5-1	>1	*RMSE*, zwitterions excluded[2]	*RMSE*	rank	% in error range <0.5	0.5-1	>1
Consensus log *P*	0.95		I	48	29	24	0.94	0.58	I	61	32	7
ALOGPS	1.02		I	41	30	29	1.01	0.68	I	51	34	15
S+logP	1.02		I	44	29	27	1.00	0.69	I	58	27	15
NC+NHET	1.04		II	38	30	32	1.04	0.88	III	42	32	26
MLOGP(S+)	1.05		II	40	29	31	1.05	1.17	III	32	26	41
XLOGP3	1.07		II	43	28	29	1.06	0.65	I	55	34	12
MiLogP	1.10	27	II	41	28	30	1.09	0.67	I	60	26	14
AB/LogP	1.12	24	II	39	29	33	1.11	0.88	III	45	28	27
ALOGP	1.12		II	39	29	32	1.12	0.72	II	52	33	15
ALOGP98	1.12		II	40	28	32	1.10	0.73	II	52	31	17
OsirisP	1.13	6	II	39	28	33	1.12	0.85	II	43	33	24
AAM	**1.16**		**III**	**33**	**29**	**38**	**1.16**	**0.94**	**III**	**42**	**31**	**27**
CLOGP	1.23		III	37	28	35	1.21	1.01	III	46	28	22
ACD/logP	1.28		III	35	27	38	1.28	0.87	III	46	34	21
CSlogP	1.29	20	III	37	27	36	1.28	1.06	III	38	29	33
COSMOFrag	1.30	1088[3]	III	32	27	40	1.30	1.06	III	29	31	40
QikProp	1.32	103	III	31	26	43	1.32	1.17	III	27	24	49
KowWIN	1.32	16	III	33	26	41	1.31	1.20	III	29	27	44
QLogP	1.33	24	III	34	27	39	1.32	0.80	II	50	33	17
XLOGP2	1.80		III	15	17	68	1.80	0.94	III	39	31	29
MLOGP(Dragon)	2.03		III	34	24	42	2.03	0.90	III	45	30	25

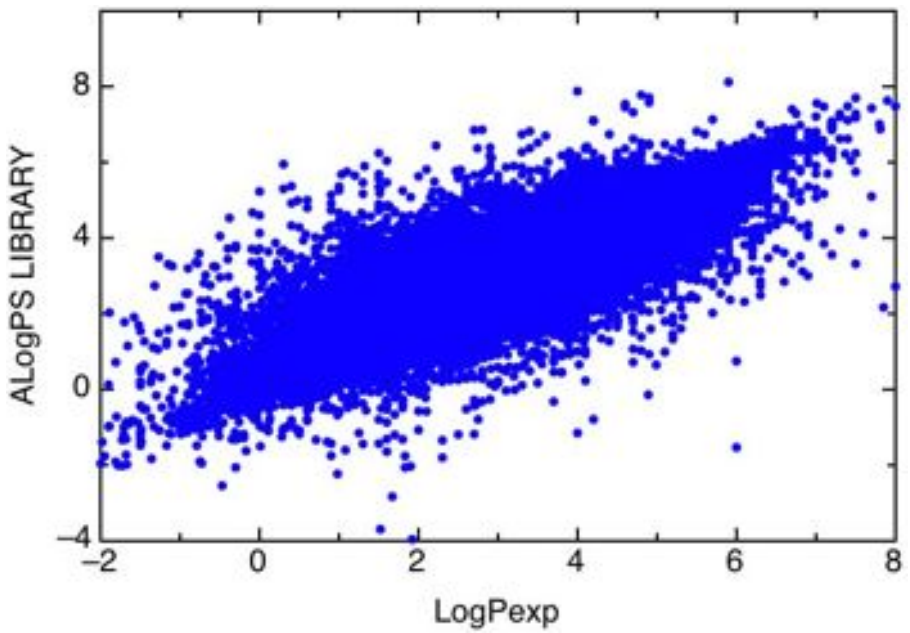
Mannhold, R. et al, J. Pharm. Sci., 2009, 98(3), 861-893.

ALOGPS decreases errors about twice using local corrections for *N*=95809 *in house* Pfizer molecules

ALOGPS Blind prediction



ALOGPS LIBRARY

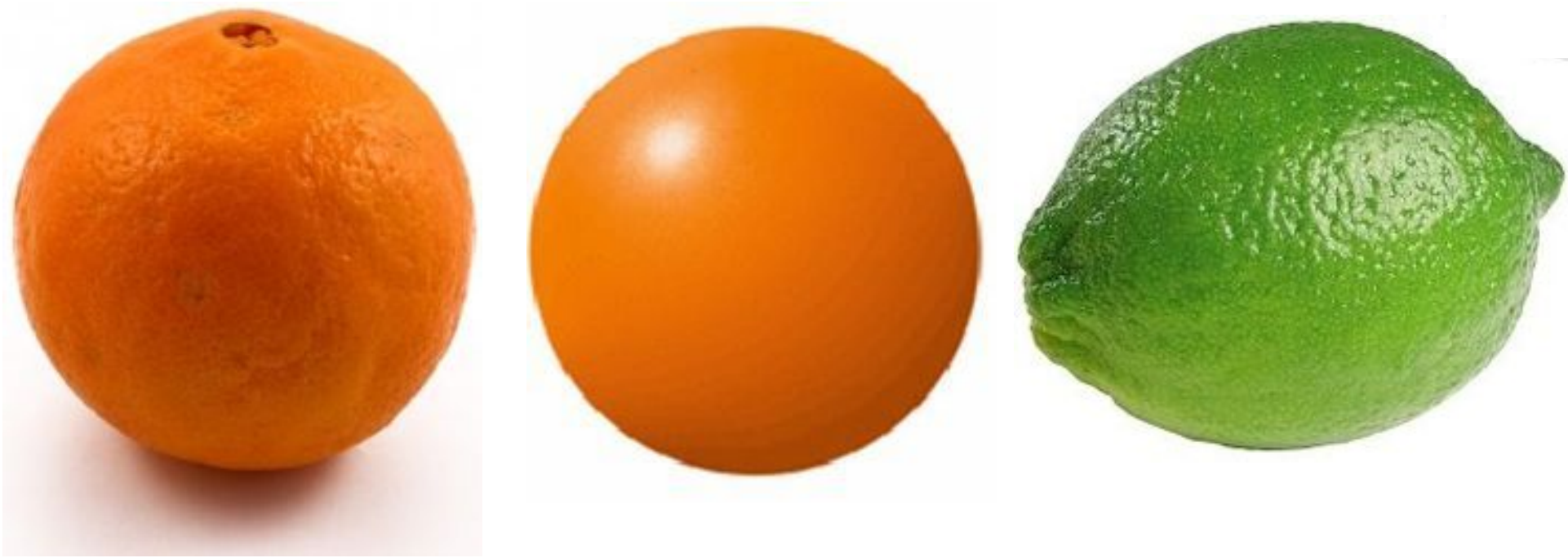


RMSE=1.02 → RMSE=0.59

ca 30 minutes of calculations on a notebook!

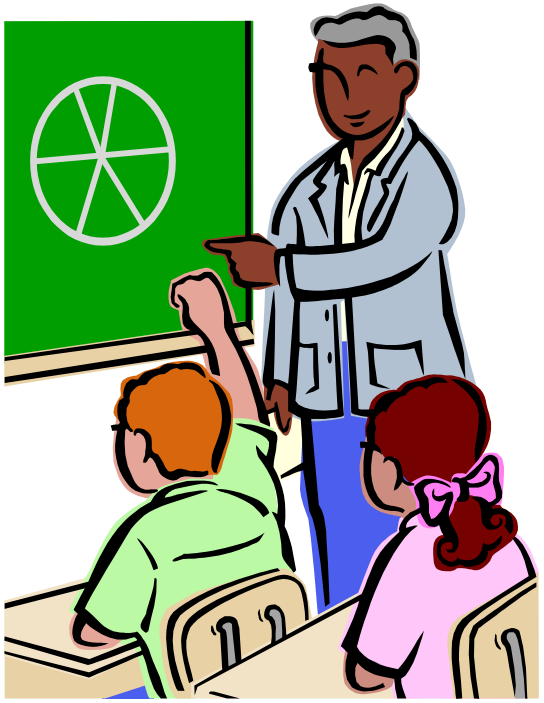
Tetko et al, *QSAR Comb. Sci.*, **2009**, *28*, 845-9.

The descriptor space challenge



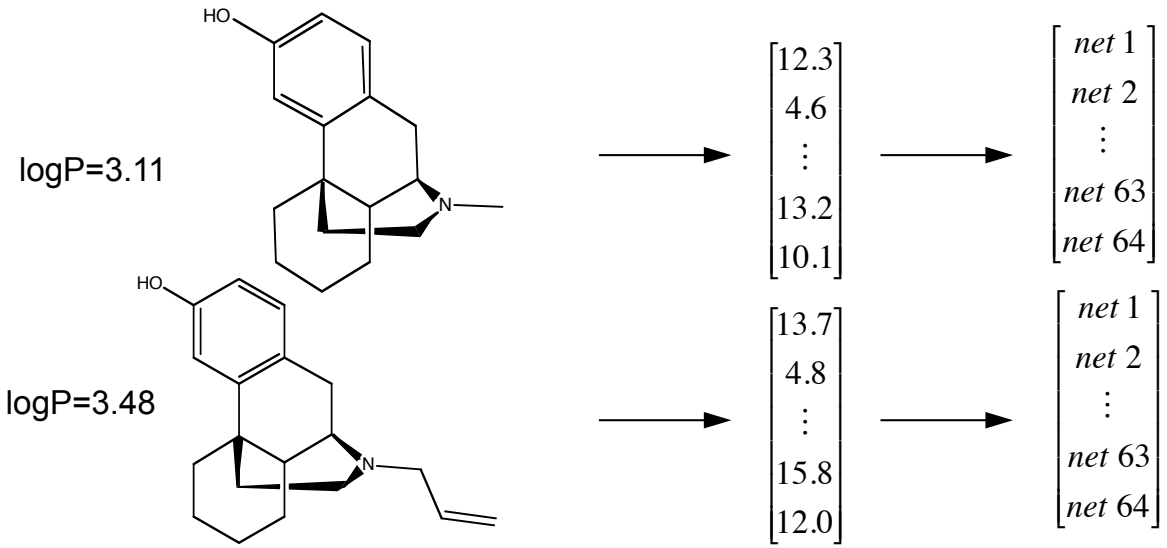
We need to know the target property and select correct descriptors!

Property-based space similarity illustration



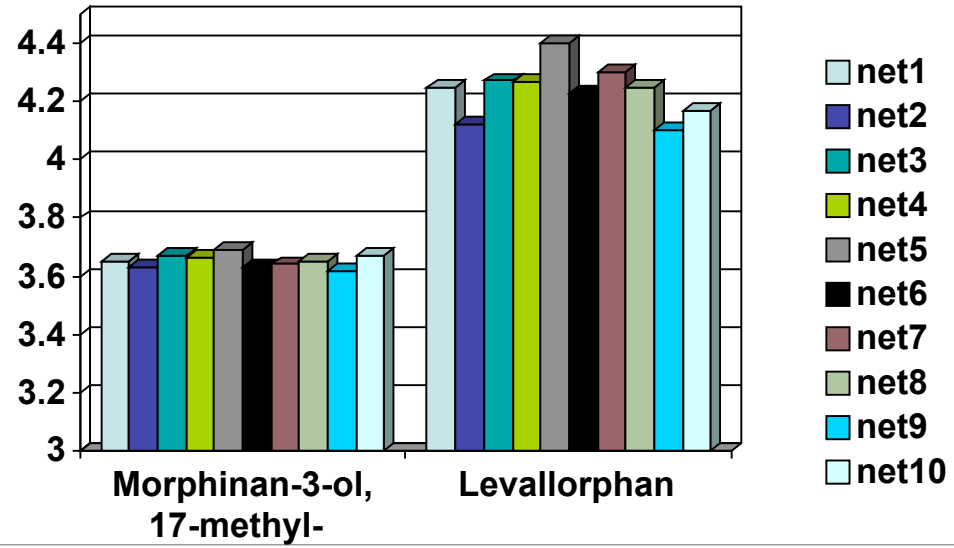
*Do they agree in their votes (**STD**)?*
*Do they have the same pattern of votes (**CORREL**)?*

Associative Neural Network Property-Based DMs



Morphinan-3-ol, 17-methyl-

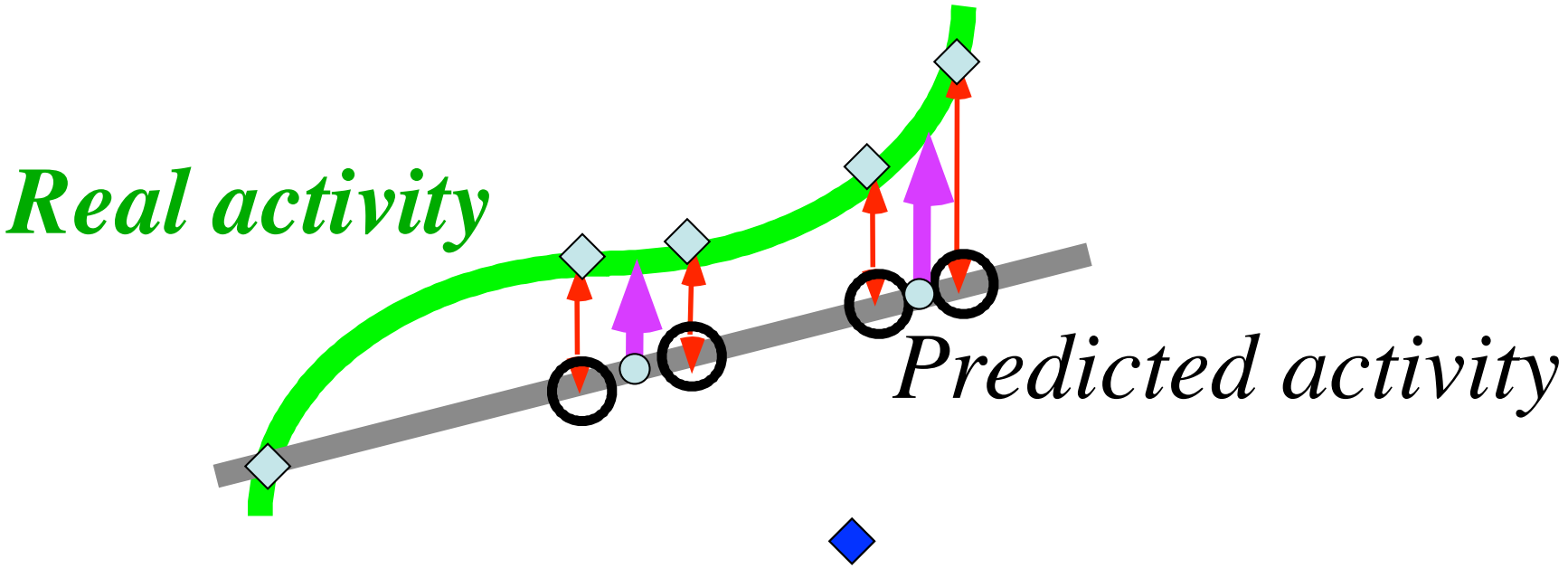
Levallorphan



CORREL - correlation between vectors of predictions

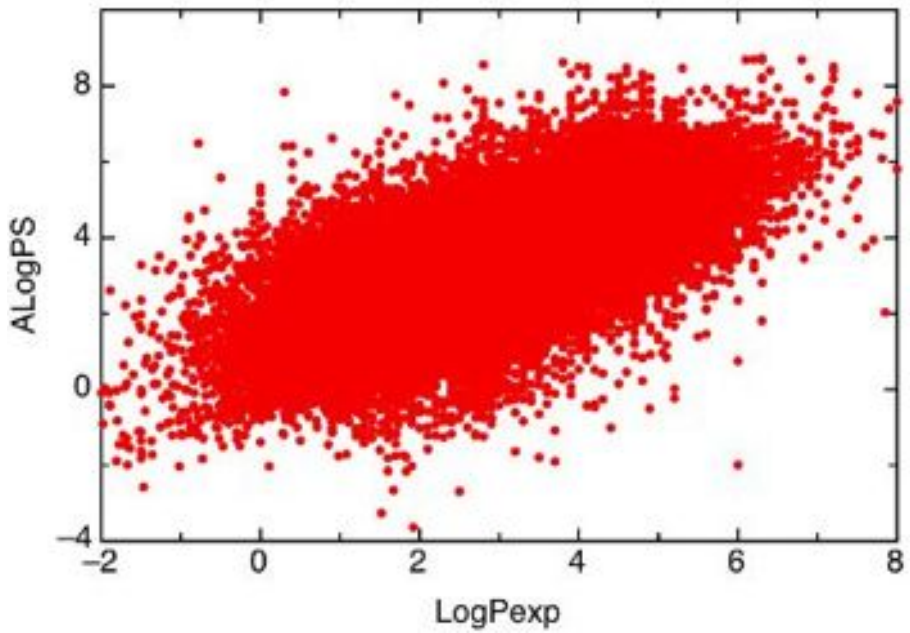
STD - standard deviation of ensemble predictions

Illustration of local correction using nearest neighbors

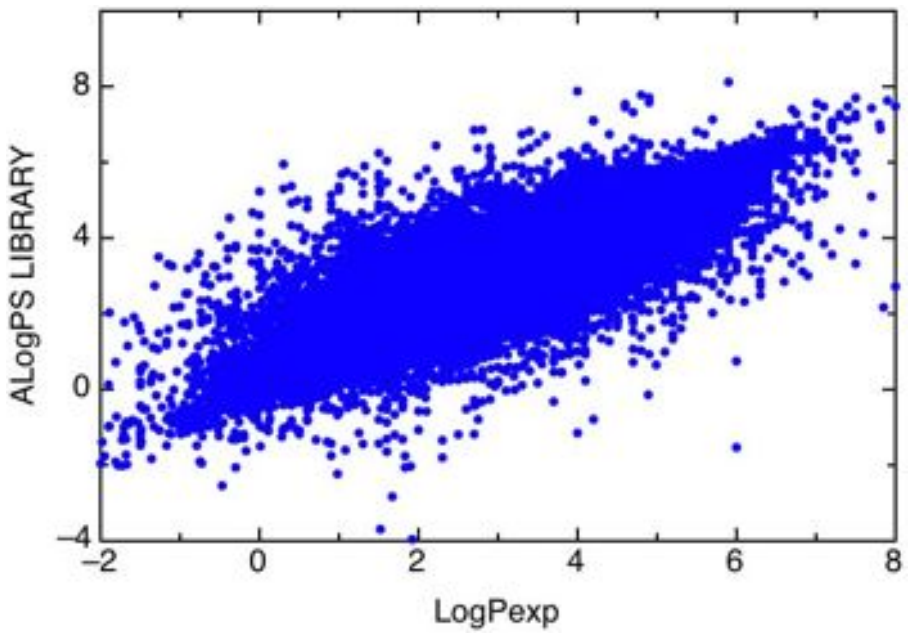


ALOGPS decreases errors about twice using local corrections for *N*=95809 *in house* Pfizer molecules

ALOGPS Blind prediction



ALOGPS LIBRARY

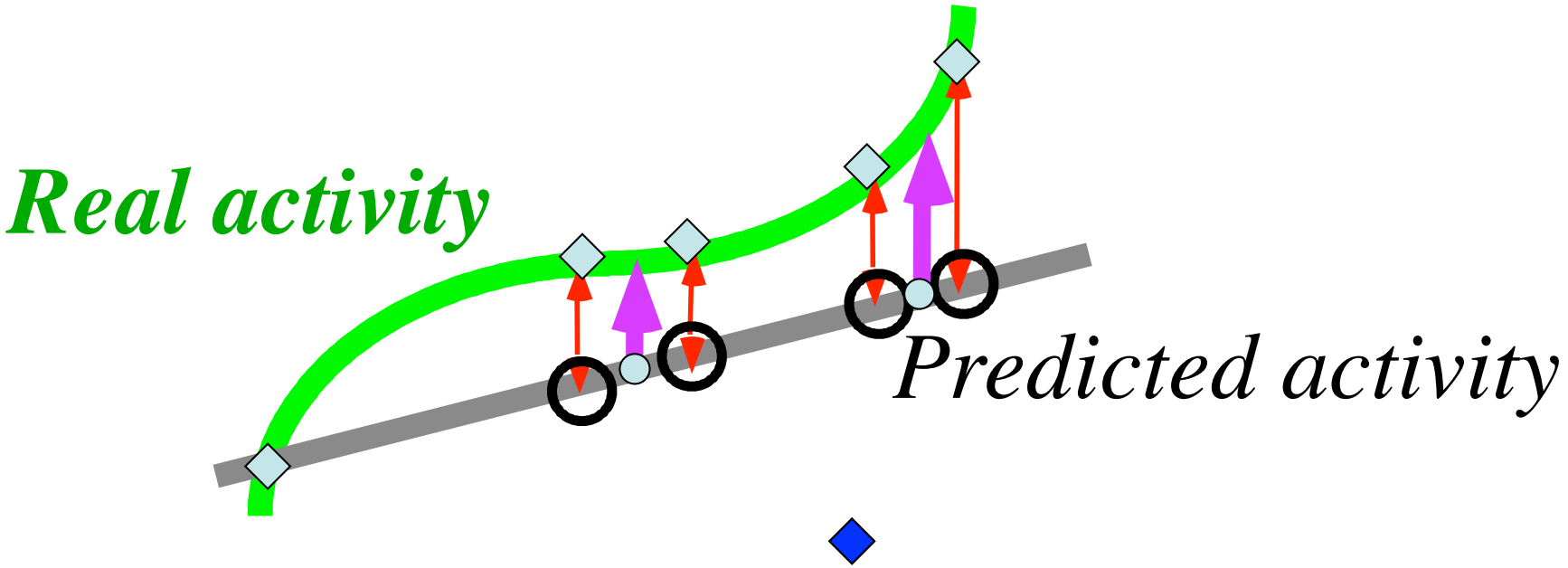


RMSE=1.02 → RMSE=0.59

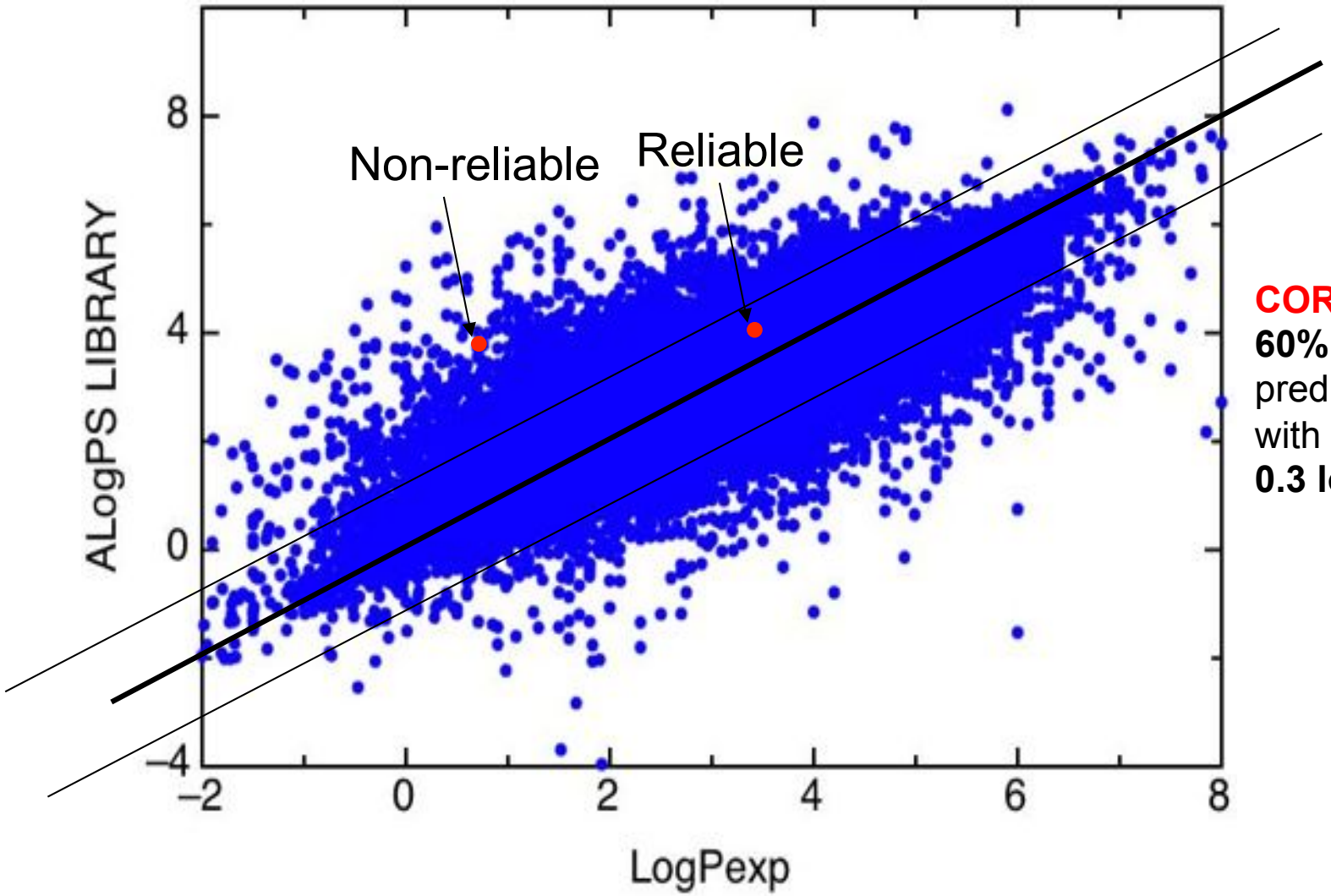
ca 30 minutes of calculations on a notebook!

Tetko et al, *QSAR Comb. Sci.*, **2009**, *28*, 845-9.

Illustration of local correction using nearest neighbors



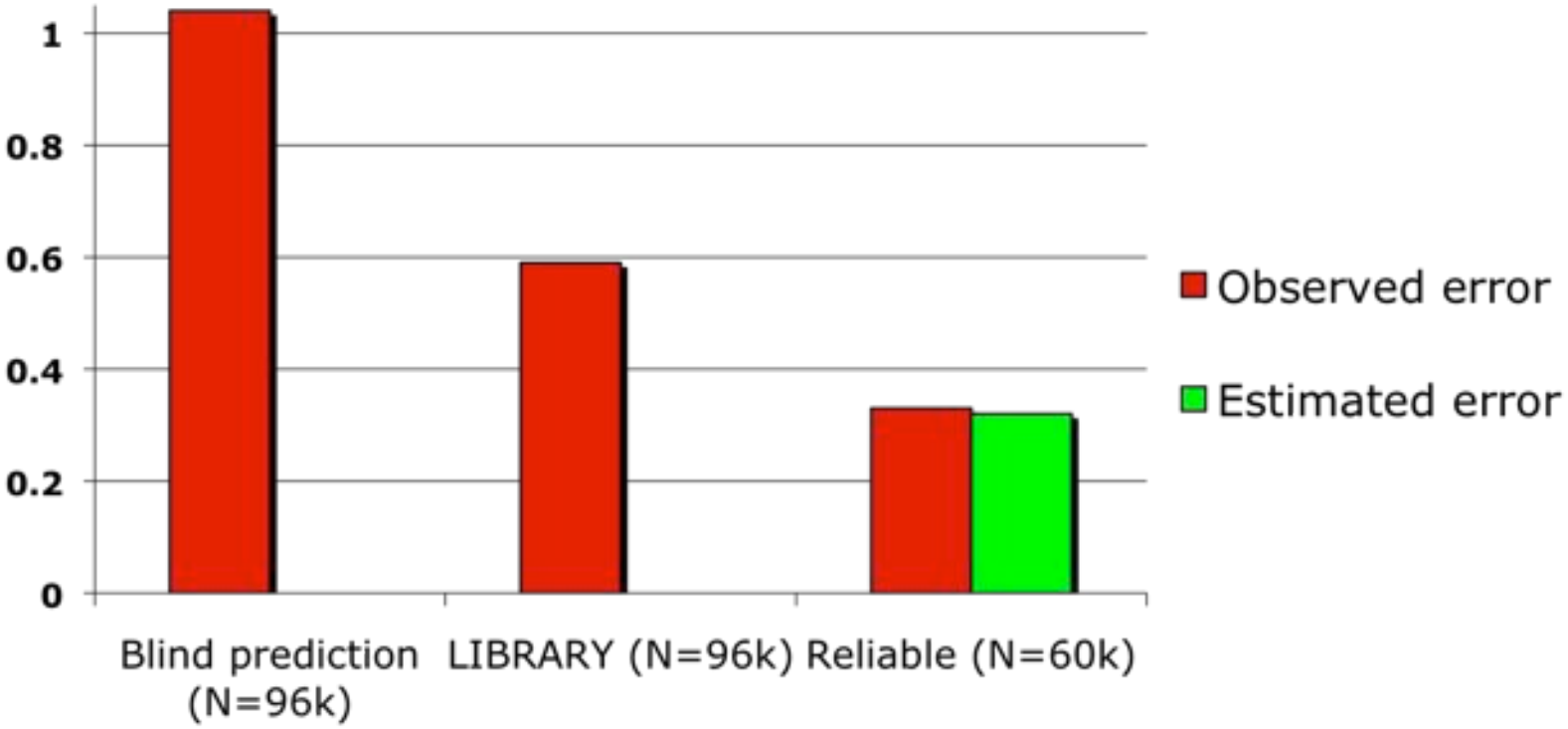
ALOGPS distinguishes reliable vs. non-reliable predictions in property-based space (CORREL)



CORREL identifies **60% of molecules** predicted with average accuracy of **0.3 log units**

Tetko et al, *Chemistry & Biodiversity*, **2009**, **6(11), 197-202.**

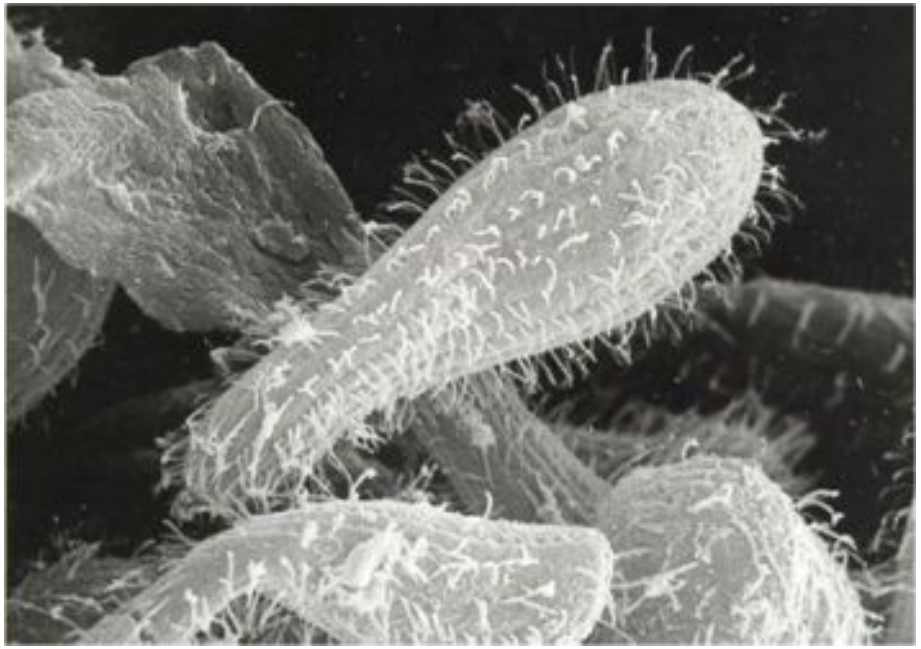
The use of ALOGPS advanced features dramatically increase prediction accuracy of the predictions



The experimental measurements accuracy was achieved for >60,000 Pfizer compounds.

Tetko et al, *Chem. & Biodivers.*, 2009, 6(11), 197-202.

Estimation of toxicity against *T. pyriformis*



T. pyriformis

Prof. T.W. Schultz

The overall goal is to predict and to assess the reliability of predictions toxicity against T. pyriformis for chemicals directly from their structure.

Dataset: 1093 molecules

Analyzed QSARs (Quantitative Structure Activity Relationship) and distances to models (DM)

country	modeling techniques	descriptors	abbreviation	distances to models (in space)	
				descriptors	property-based
	ensemble of 192 kNN models	MolconnZ	kNN-MZ	EUCLID	STD
	ensemble of 542 kNN models	Dragon	kNN-DR	EUCLID	STD
	SVM	MolconnZ	SVM-MZ		
	SVM	Dragon	SVM-DR		
	SVM	Fragments	SVM-FR		
	kNN	Fragments	kNN-FR	EUCLID, TANIMOTO	
	MLR	Fragments	MLR-FR		
	MLR	Molec. properties (CODESSA-Pro)	MLR-COD		
	OLS	Dragon	OLS-DR	LEVERAGE	
	PLS	Dragon	PLS-DR	LEVERAGE	PLSEU
	ensemble of 100 neural networks	E-state indices	ASNN-ESTATE		CORREL, STD
All	consensus model	-	CONS		STD

Tetko et al, *J Chem Inf Model*, **2008,** 48(9):1733-46.

Overview of analyzed distances to models (DMs)

EUCLID

$$EU_m = \frac{\sum_{j=1}^{k} d_j}{k}$$

$$EUCLID = \overline{EU}_m$$

k is number of nearest neighbors, m index of model

TANIMOTO

$$Tanimoto(a,b) = \frac{\sum x_{a,i}x_{b,i}}{\sum x_{a,i}x_{a,i} + \sum x_{b,i}x_{b,i} - \sum x_{a,i}x_{b,i}}$$

$x_{a,i}$ and $x_{b,i}$ are fragment counts

LEVERAGE

$$LEVERAGE = \mathbf{x}^T(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{x}$$

PLSEU (DModX)

Error in approximation (restoration) of the vector of input variables from the latent variables and PLS weights.

STD

$$STD = \frac{1}{N-1}\sum(y_i - \bar{y})^2$$

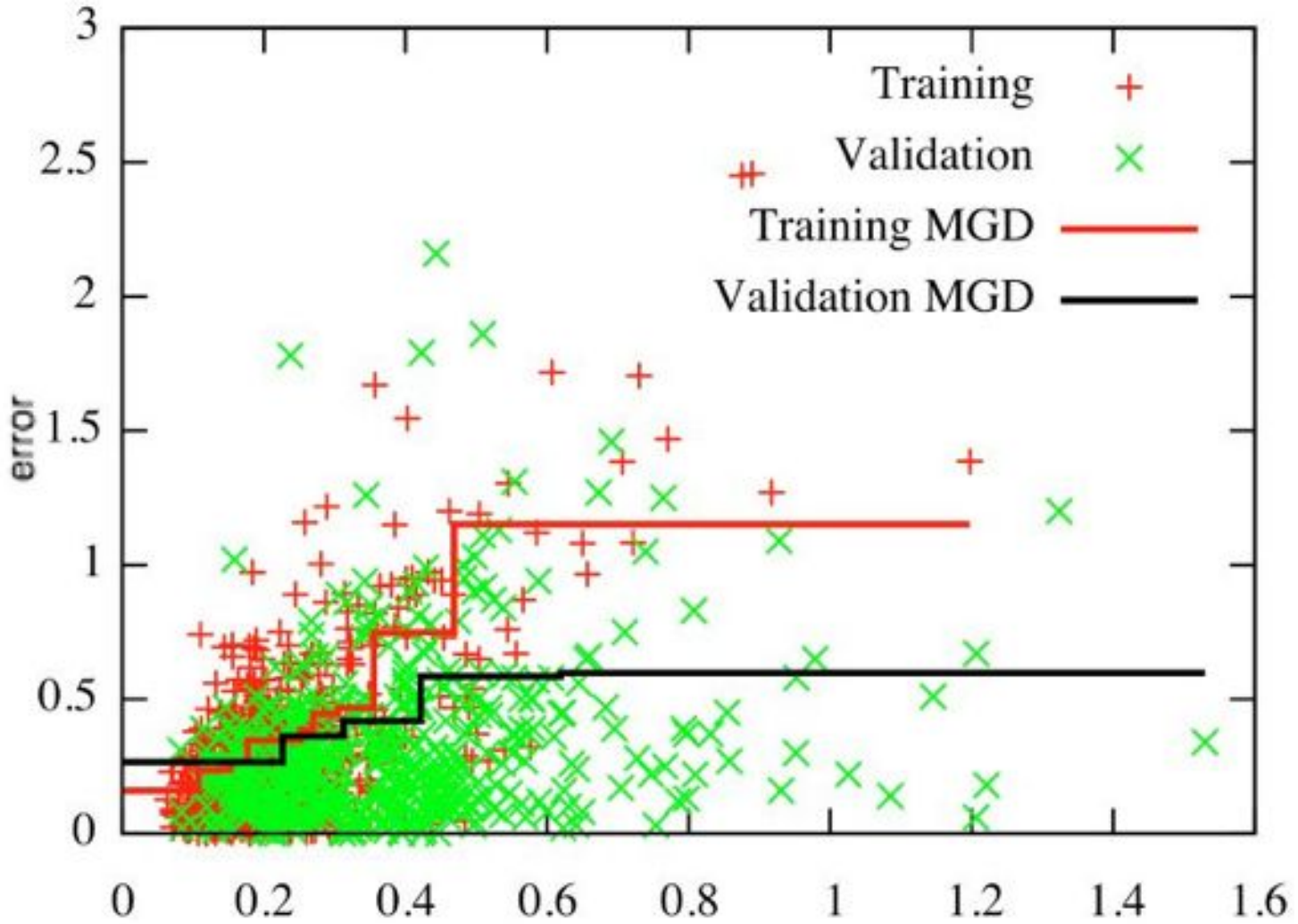
y_i is value calculated with model i and $\bar{y}$ is average value

CORREL

$$CORREL(a) = \max_j CORREL(a,j) = R^2(\mathbf{Y}^a_{calc}, \mathbf{Y}^j_{calc})$$

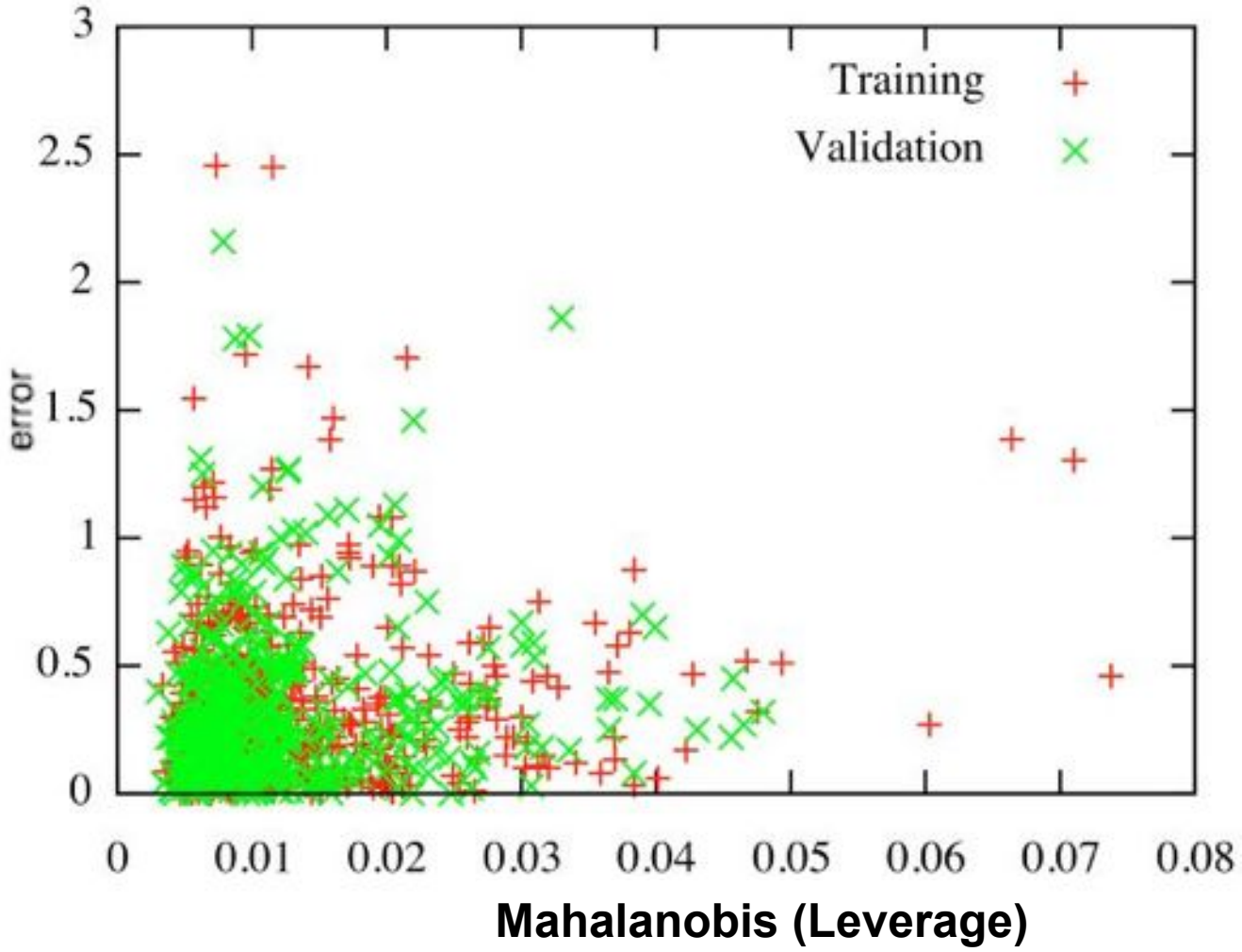
$\mathbf{Y}^a = (y_1,\ldots,y_N)$ is vector of predictions of molecule i

Property-based, ASNN model: DM does work!



Tetko et al, *J. Chem. Inf. Model*, **2008,** *48*, 1733-46.

Descriptor space, ASNN model: DM does not work



Tetko et al, *J. Chem. Inf. Model*, **2008,** *48*, 1733-46.

Ranking of Distance to Models (DM)

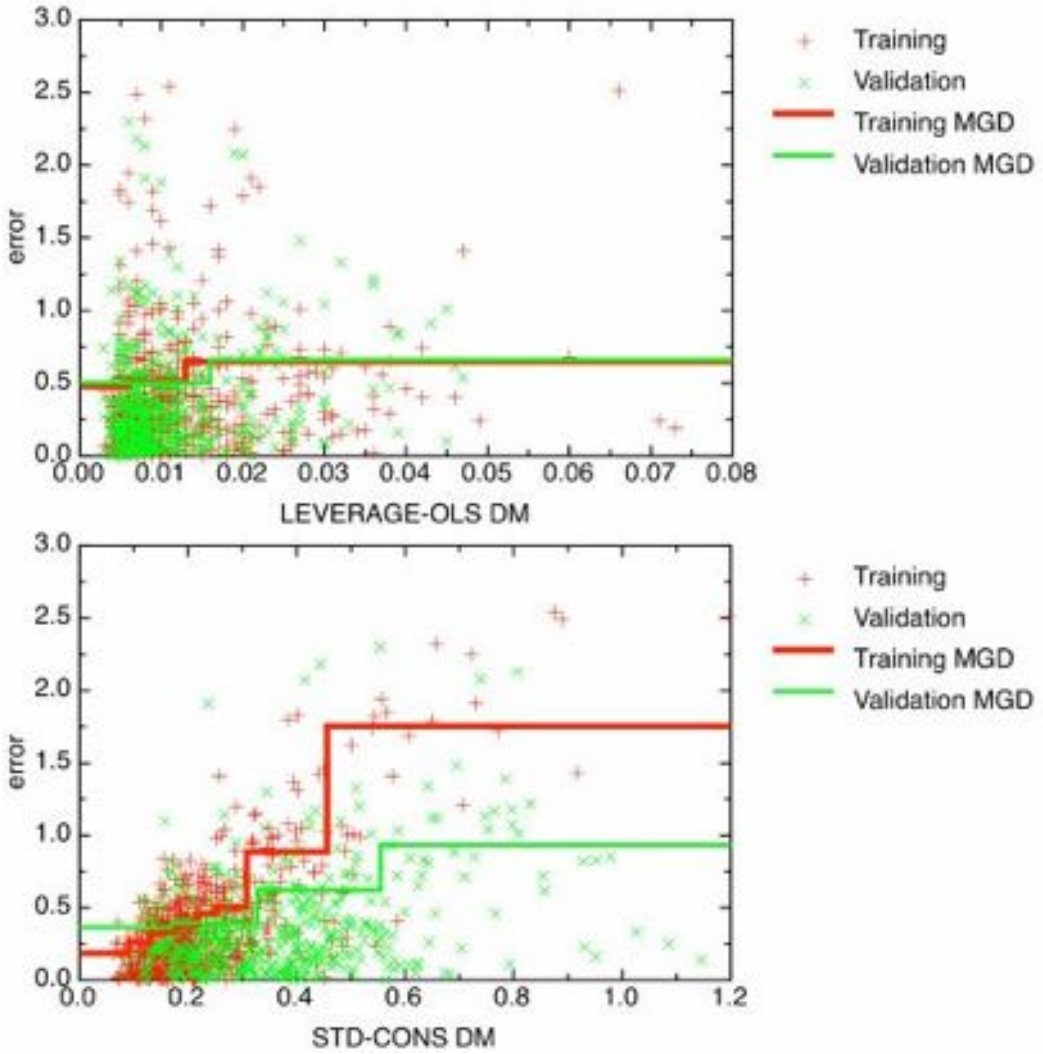
DM	average rank			highest rank[1]		
	LOO	5-CV	**Valid.***	LOO	5-CV	Valid.
STD-CONS	1	1.8	**1.1**	12	2	11
STD-ASNN	2	1.2	**2.5**		10	1
STD-kNN-DR	6.6	4.3	**4.1**			
STD-kNN-MZ	9.2	8.3	**5.3**			
EUCLID-kNN-DR	7.1	4.9	**5.4**			
LEVERAGE-PLS	8.4	5	**6.3**			
EUCLID-kNN-MZ	7.5	7.1	**6.4**			
TANIMOTO-kNN-FR	7	6.1	**6.8**			
TANIMOTO-MLR-FR	8.3	8.3	**9**			
CORREL-ASNN	10.7	10.8	**9.4**			
LEVERAGE-OLS-DR	12.3	12.6	**11.1**			
EUCLID-MLR-FR	7	9.3	**11.5**			
PLSEU-PLS	11.1	11.8	**11.5**			
EUCLID-kNN-FR	12.1	13.3	**12.1**			

*Ordered by performance of the DMs on the validation dataset

Tetko et al, *J. Chem. Inf. Model*, **2008,** *48*, 1733-46.

Analysis of DMs for a linear model

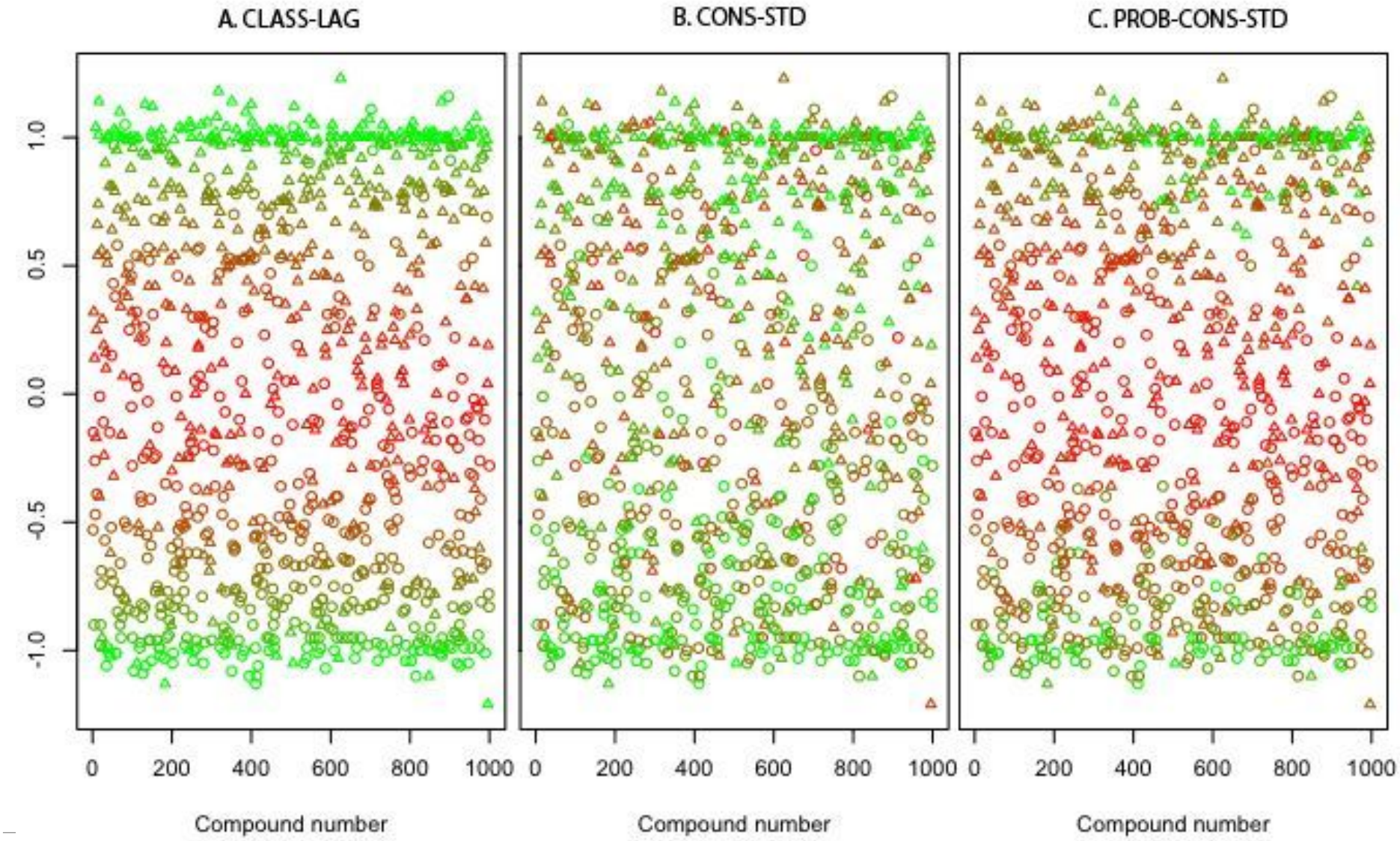
$Log(IGC_{50}^{-1})=$
-18(±0.7) +0.065(±0.002)**AMR**-0.50
(0.04)**O56**-0.30(0.03)**O58**
-0.29(0.02)**nHAcc**+0.046(0.005)
H-046+16(0.7)**Me**



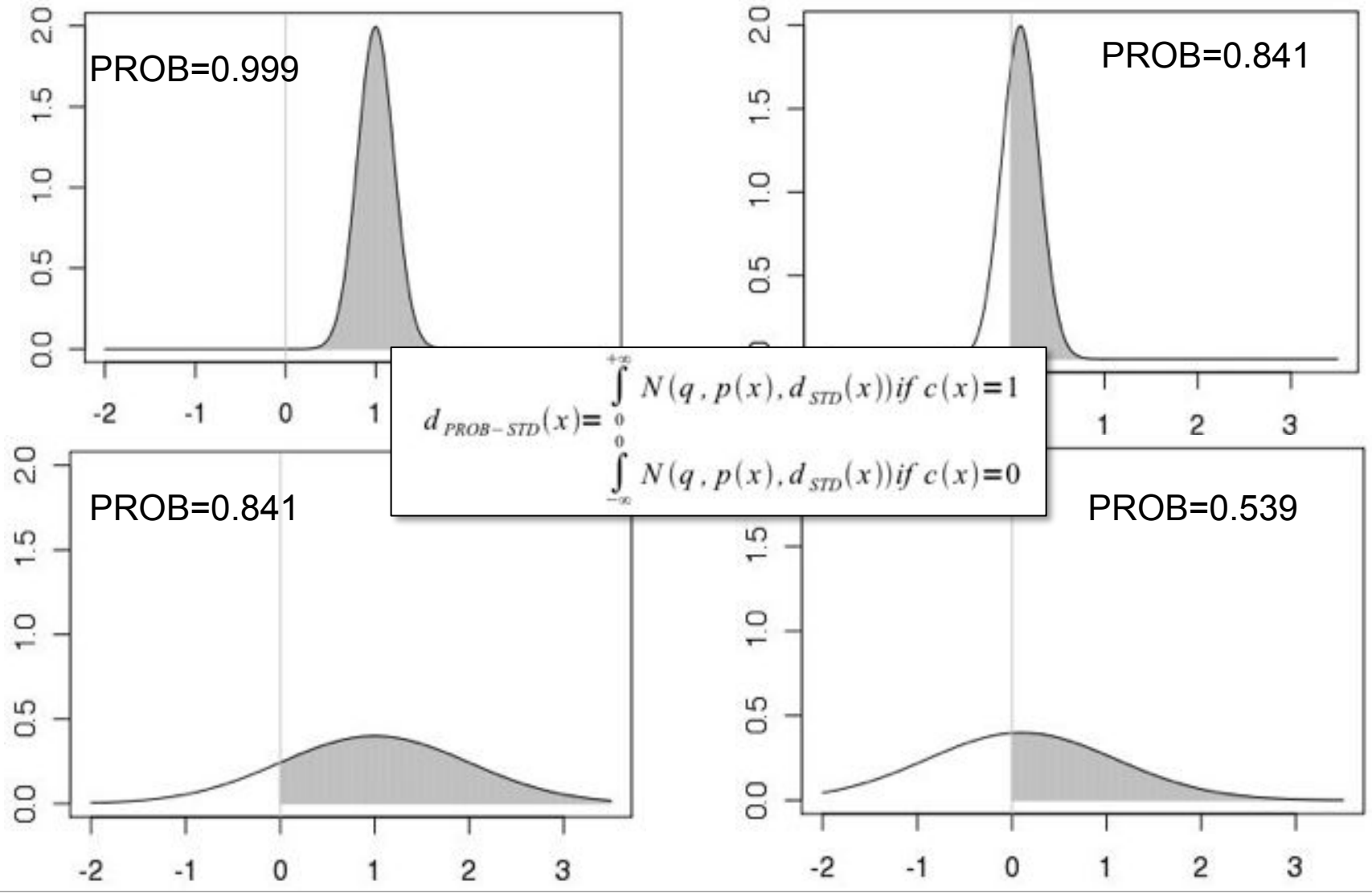
The results of DM performance are consistent across different models

Tetko et al, *J Chem Inf Model*, **2008,** 48(9):1733-46.

Classification task distance measures



Binary classification



Prediction of Ames Mutagenicity set

http://ml.cs.tu-berlin.de/toxbenchmark

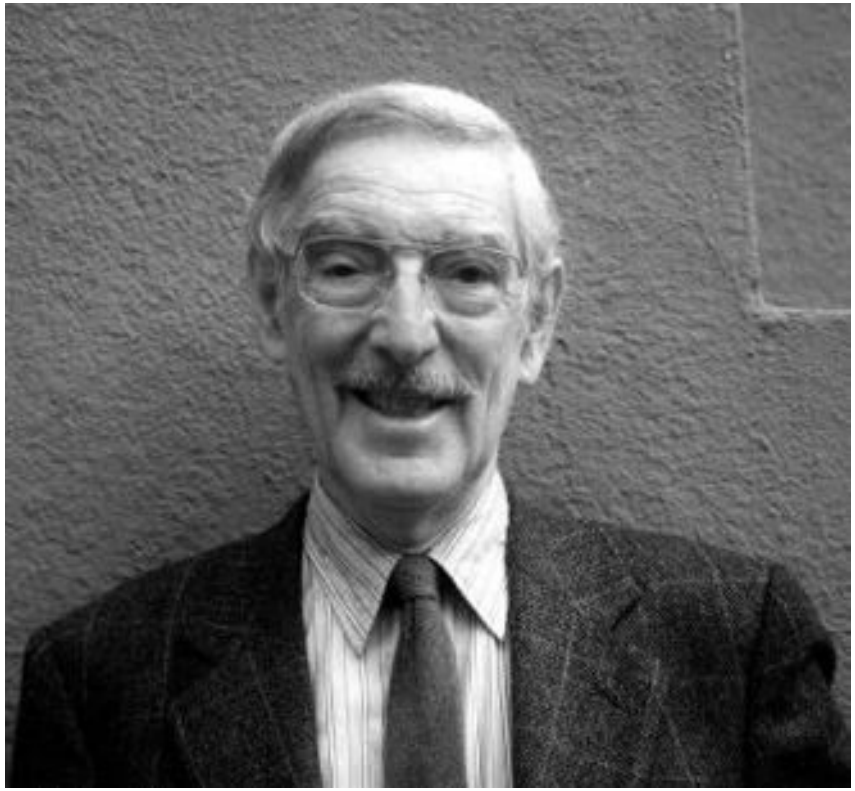
Toxicity against *Salmonella typhimurium*

Training dataset: 4361 molecules

"Blind" test dataset: 2181 molecules

54% with mutagenic effect

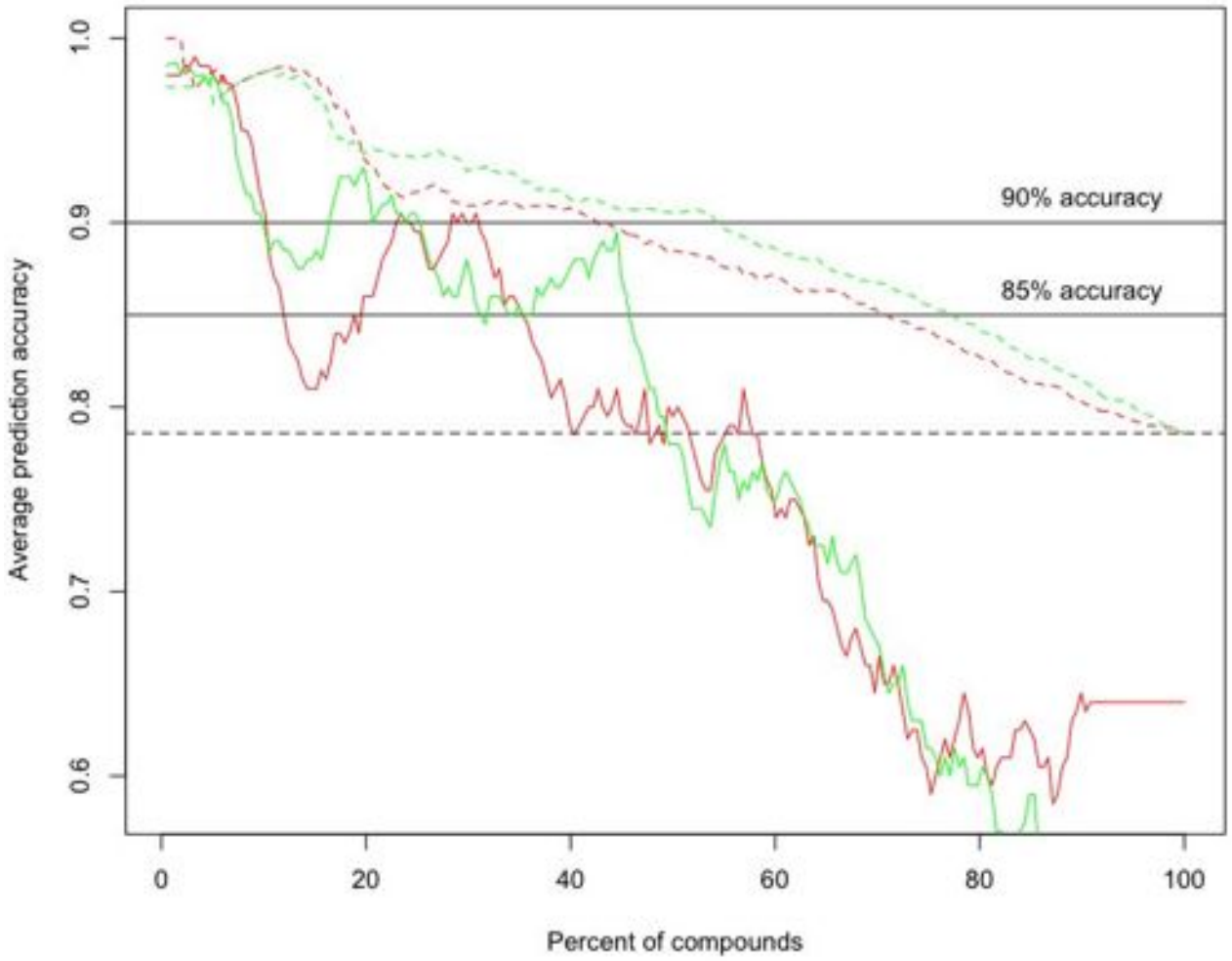
Large international collaboration effort of >10 labs from USA, Canada, EU, Russia, the Ukraine & China (see also poster P-22)



Prof. Bruce N. Ames
Inventor of the test (1975)

[1]Schwaighofer et al, *JCIM*, **2008**, 48, 785-96.

Accuracy of a AMES consensus model as function of two Distances to Models



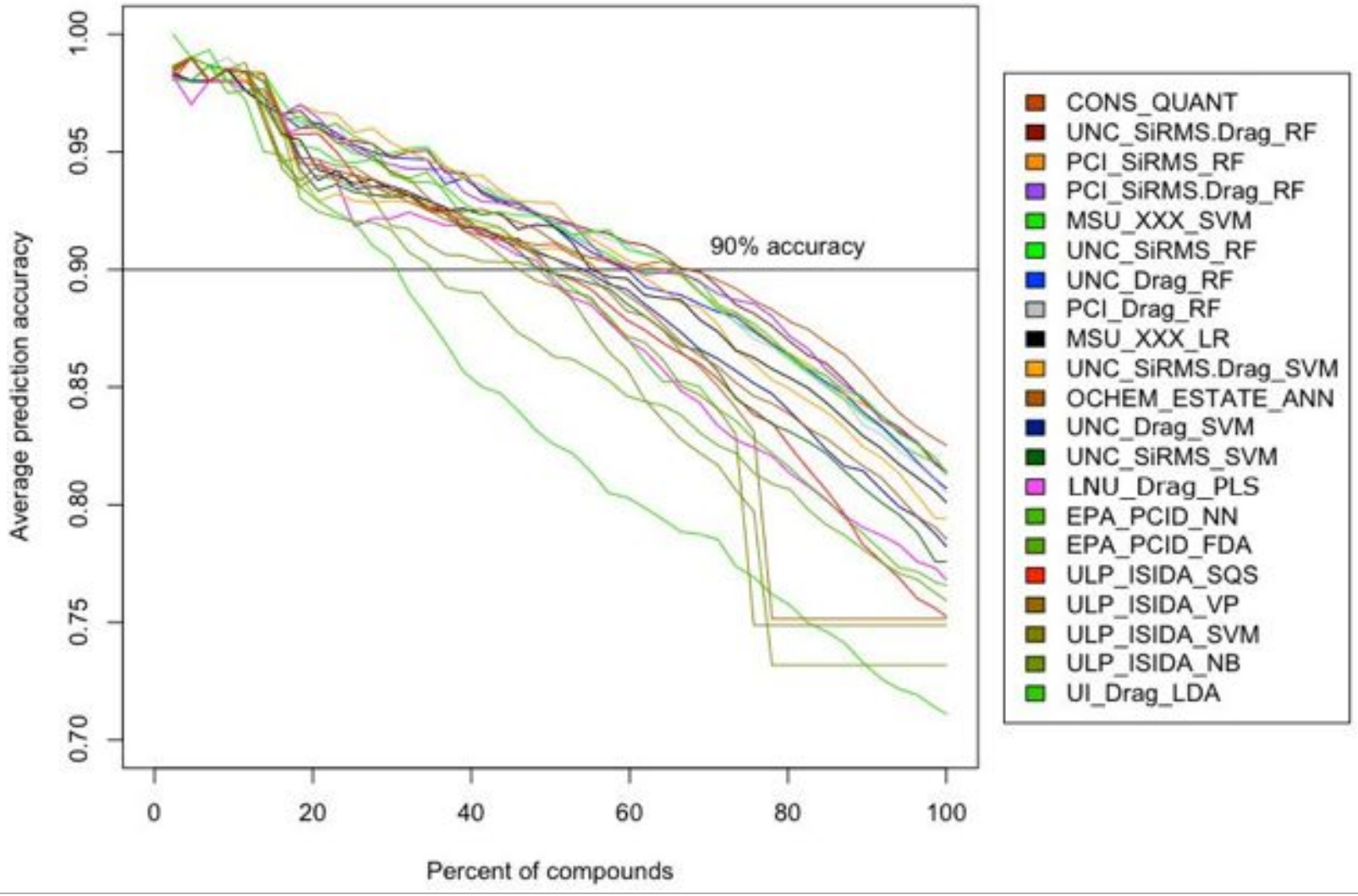
Accuracy=CORRECT/ALL

Averaged ranking of DMs according to the percentage of compounds with 90% accuracy for training and test sets.

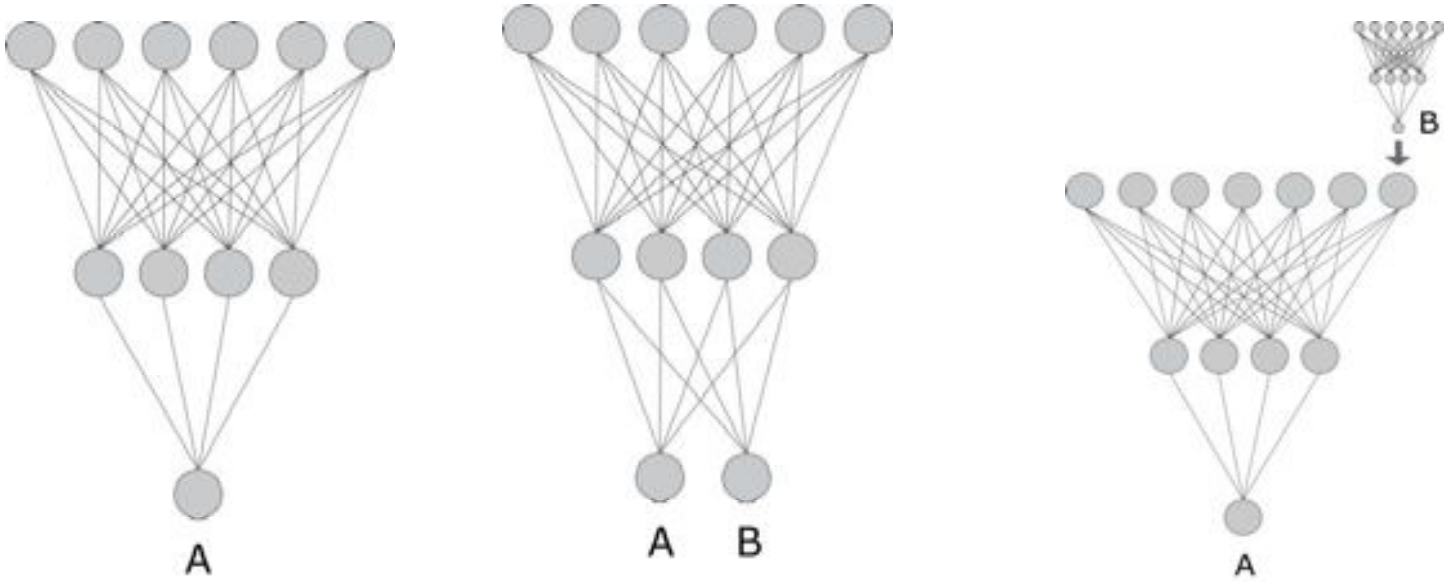
Distance to model	Average rank - training set	Average rank – test set
CONS-STD-QUAL-PROB	2.17	1.83
CONCORDANCE	1.62	2.1
CONS-STD-PROB	3.43	3.05
CONS-STD-QUAL	3.67	4.9
ASNN-STD-PROB	6.52	5.48
CONS-STD	4.83	5.6
CLASS-LAG	7.1	6.24
ASNN-STD	8.14	7.67
AD_MEAN1*	10.71	9.07
CORREL	9.26	10.26
AD_MEAN2*	9.71	10.86
LEVERAGE*	10.83	10.95

CONCORDANCE is the number of models that give the same prediction, as the current model does

Accuracy of different AMES model as function of a Distance to Models



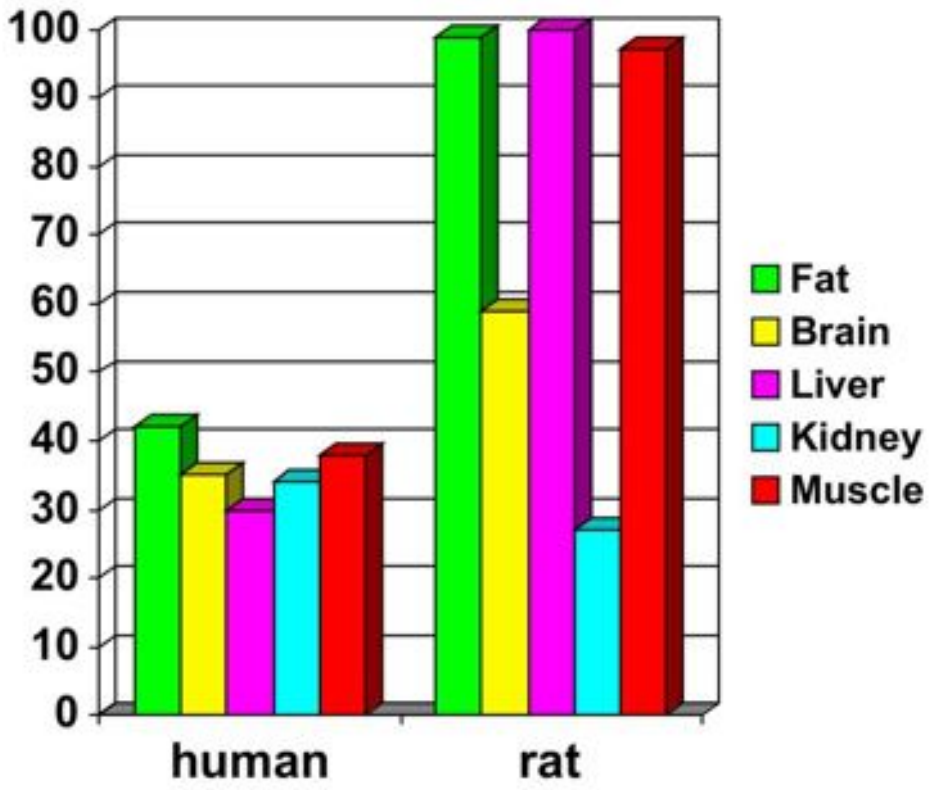
Multi-task learning



Multi-task learning: unequal number of data

Problem:

- prediction of tissue-air partition coefficients
- small datasets 30-100 molecules (human & rat data)



Varnek, et al, JCIM, 2009, 49(1):133-144.

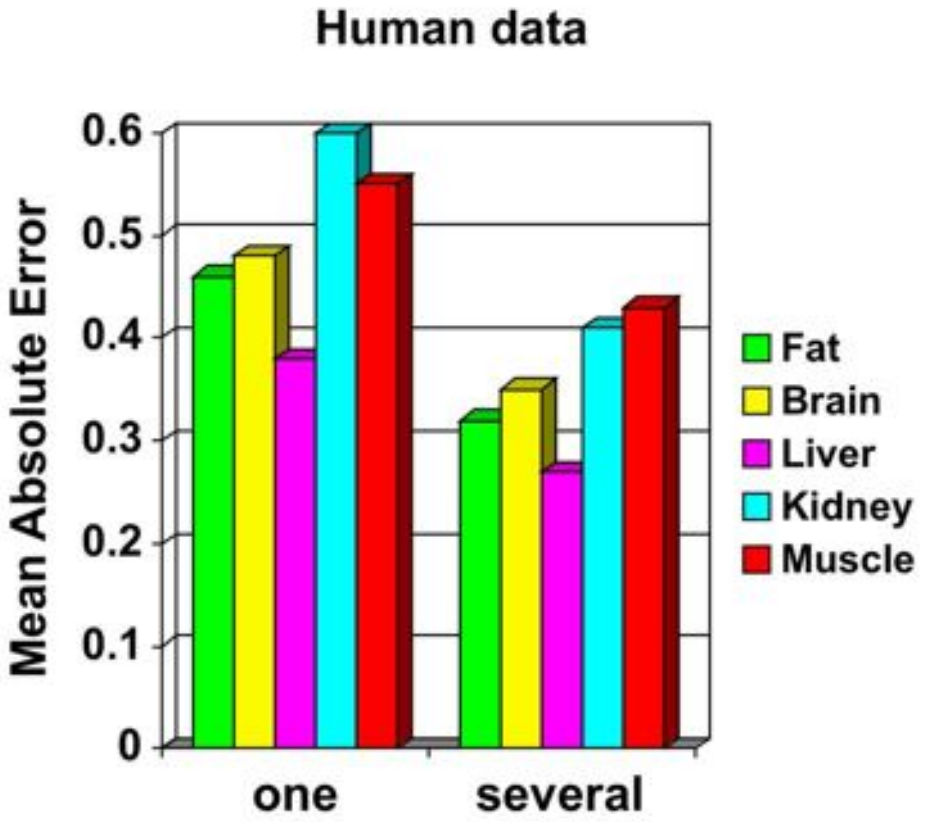
Multi-task learning can improve models for small sets

Problem:

• prediction of tissue-air partition coefficients
• small datasets 30-100 molecules (human & rat data)

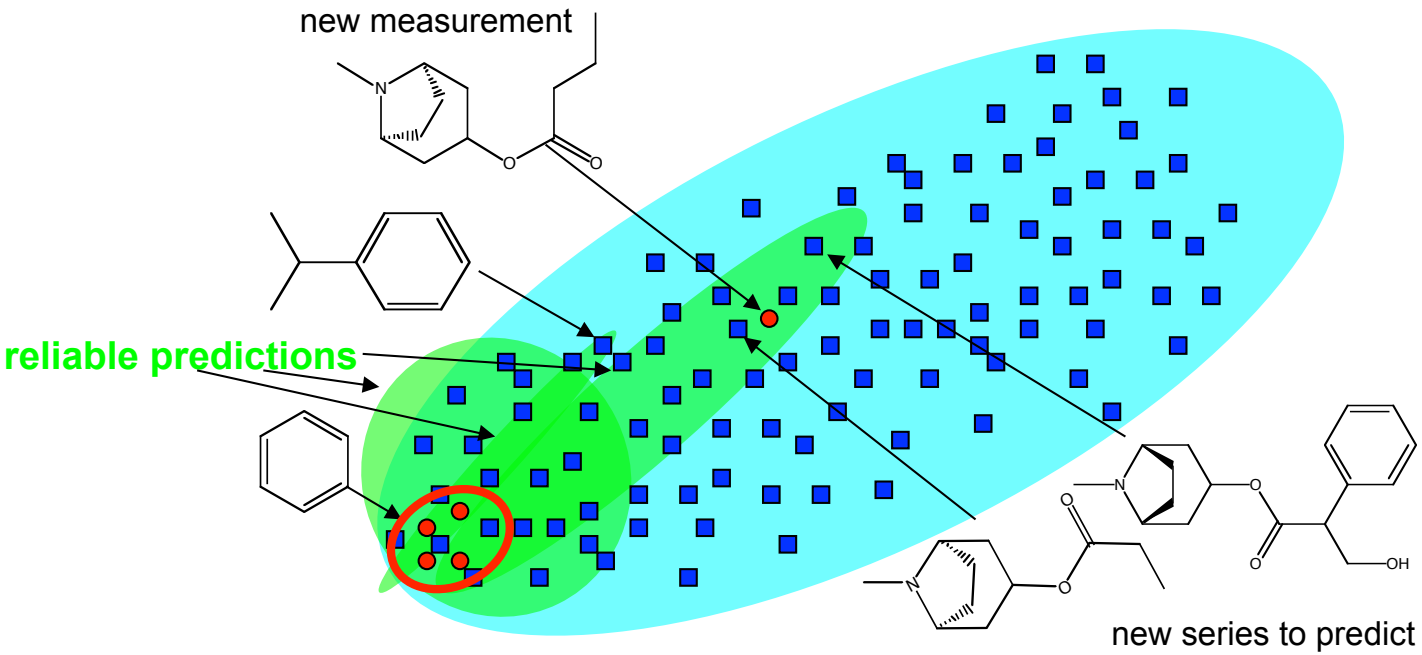
Results:

simultaneous prediction of several properties increased the accuracy of models



Varnek, et al, JCIM, 2009, 49(1):133-144.

ADMETox *in silico* challenges



ADMETox models should allows navigation in space of molecules with a confidence and:

- ✓ should reliably estimate which compounds can/can't be reliably predicted.
- ✓ provide experimental design and to minimize costs of new measurements.
- ✓ be easily interpretable for chemists

HelmholtzZentrum münchen
German Research Center for Environmental Health

Online CHEmical Modeling environment (OCHEM)

http://ochem.eu

HELMHOLTZ
ASSOCIATION

Motivation

Properties of molecules

- Data are lost after publication of an article
- The original sources of data are difficult to track
- The conditions of experiments are frequently not provided
- The conversion between different units is error prone
- Current databases do not allow community correction of errors
- The tracking of changes (by users) is required

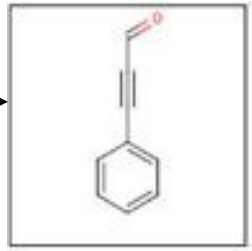
Models

- Most published models are never used
- Implementation can be as difficult as new model development
- Different implementations can produce different results*

Database schema

Simplified overview

Evidences



log(IGC50-1) = 2.02 -log (mmol/L) Temperature = 25.0

Zhu, H

Combinatorial QSAR modeling of chemical toxicants tested aga...

N: 445

Journal of chemical information and modeling **2008**; 48 (4) 766-84

2579-22-8 , phenylpropargyl aldehyde midnighter / itetko

Properties

log(IGC50-1)	(concentration)	1093 records
LogPsuv	(dimensionless)	21 records
LogPsuv(ion)	(dimensionless)	21 records
LogPl	(dimensionless)	35 records

Conditions

species	(dimensionless)
Temperature	(temperature)
dose	(concentration)
Concentration	(concentration)

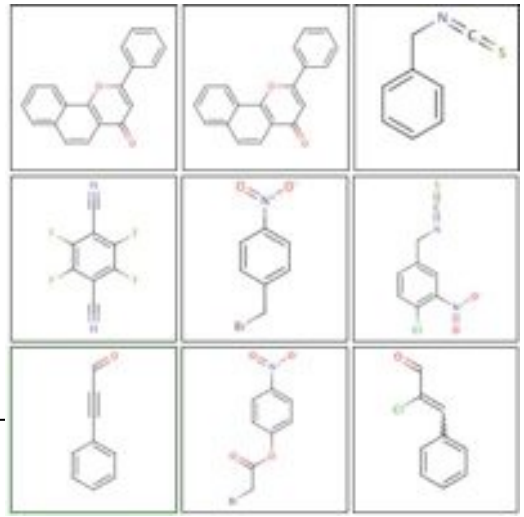
Units

log(mmol/L)	(concentration)
-log(mg/l)	(concentration)
nM	(concentration)
-log (mmol/L)	(concentration)

Molecules

Names

Users



Articles

Spink, DC;Spink, BC;Zhuo, X;Hussain, MM;Gierthy, JF;Ding, X;

NADPH- and hydroperoxide-supported 17beta-estradiol hydroxylation catalyzed by a variant form (432L, 453S) of human cytochrome P450 1B1.

The Journal of steroid biochemistry and molecular biology **2000**; 74 (1-2) 11-8

PubMed - ArticleID: Q1352

Zhang, L;Zhu, H;Oprea, TI;Golbraikh, A;Tropsha, A;

QSAR modeling of the blood-brain barrier permeability for diverse organic compounds.

Pharmaceutical research **2008**; 25 (8) 1902-14

DOI - PubMed - ArticleID: Q1577

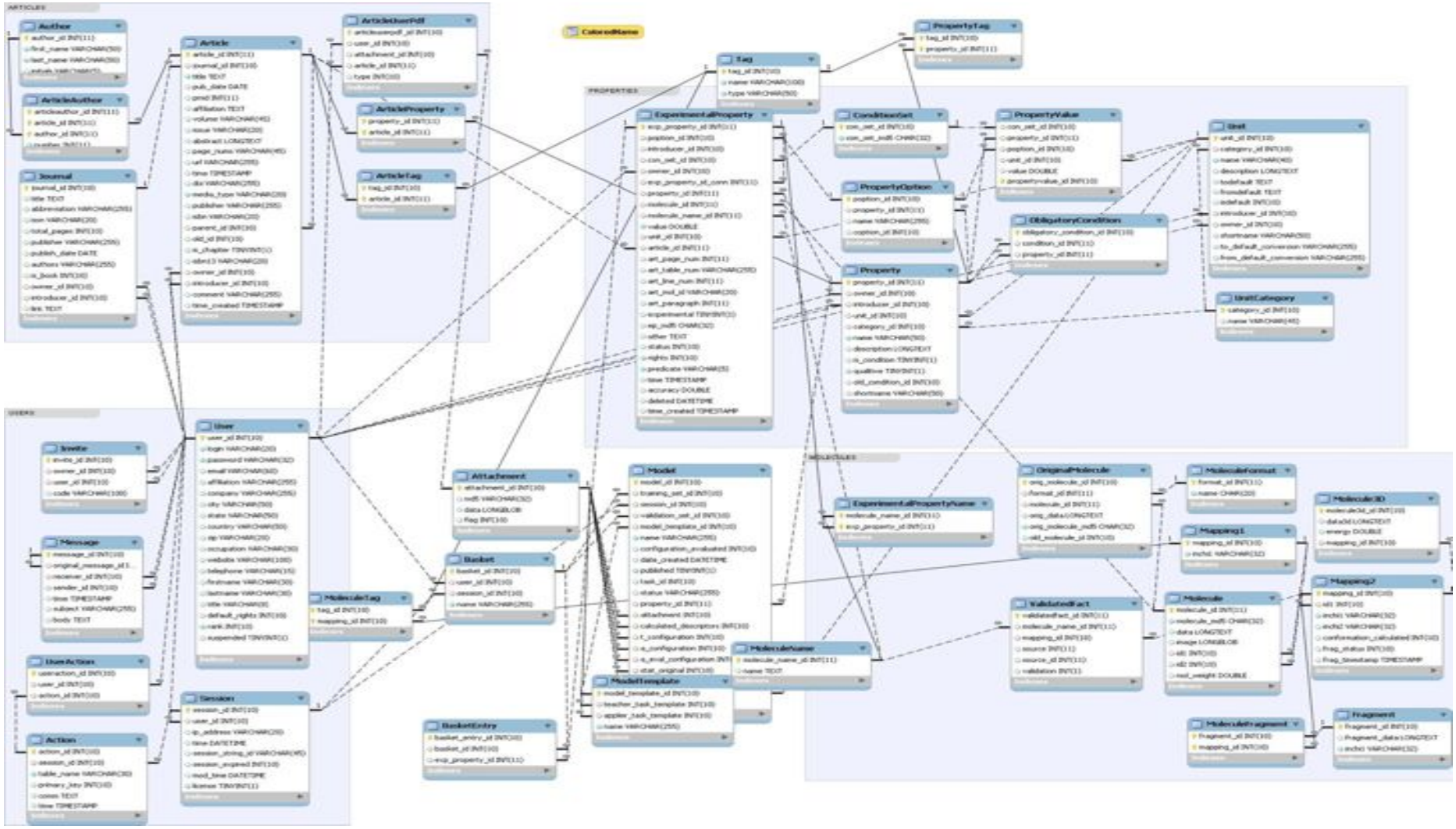
Zhu, H;Tropsha, A;Fourches, D;Varnek, A;Papa, E;Gramatica, P;Oberg, T;Dao, P;Cherkasov, A;Tetko, IV;

Combinatorial QSAR modeling of chemical toxicants tested against Tetrahymena pyriformis.

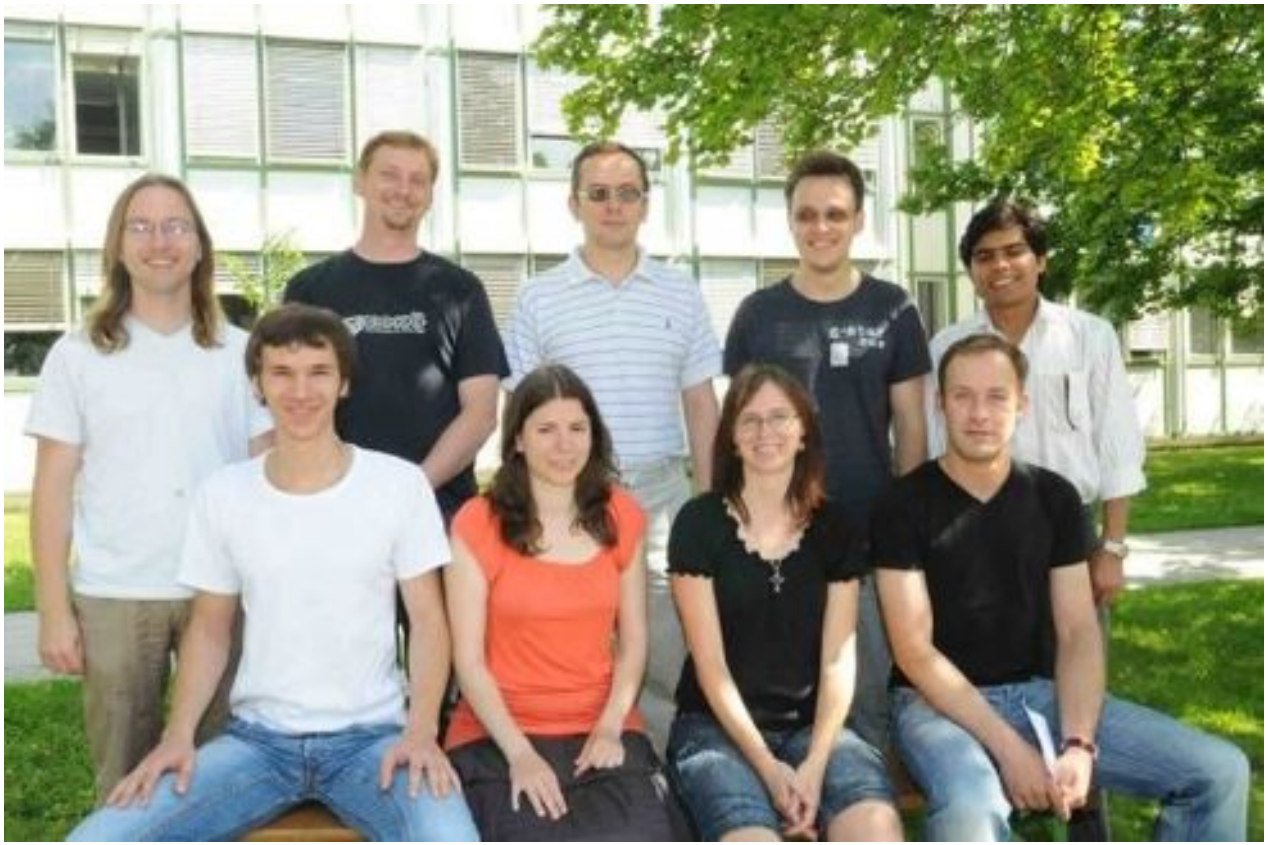
Journal of chemical information and modeling **2008**; 48 (4) 766-84

DOI - PubMed - PrePrint - ArticleID: Q1994

Data structure: behind the scene



Acknowledgements



Funding

GO-Bio BMBF http://qspr.eu
Germany-Ukraine grant UKR 08/006
FP7 Marie Curie ITN ECO http://eco-itn.eu
FP7 CADASTER http://www.cadaster.eu

My group

Iurii Sushko
Sergii Novotarskyi
Anil K. Pandey
Robert Körner
Stefan Brandmaier
Matthias Rupp
Vijyant Srivastava
Wolfram Teetz

Thanks to:

Dr. G. Poda (Pfizer)
Dr. C. Ostermann (Nycomed)
Dr. C. Höfer (DMPKore)
Prof. A. Tropsha (NC, USA)
Prof. T. Oprea (New Mexico, USA)
Prof. A. Varnek (Strasbourg, France)
Prof. R. Mannhold (Düsseldorf, Germany)
Prof. R. Todeschini (Milano, Italy)
+all coauthors of AMES
+many other colleagues

Interested in Short-Term stays (3-12 months)? Apply!!!

http://www.eco-itn.eu