

Beyond filters: assessing ADMET risk for multi-objective drug development

Robert D. Clark, Jinhua Zhang, Robert Fraczekiewicz,
Marvin Waldman, Michael Bolger and Walter Woltoz

Simulations Plus, Inc.
Lancaster CA 93534 USA

The Four Rules of 5

(Lipinski *et al.*, *Adv. Drug Deliv. Rev.* **1997**, 23, 3-25)

- **Criteria**
 - $\text{OH} + \text{NH} > 5$
 - Molecular Weight (MW) > 500
 - $\log P > 5$ (MlogP > 4.15)
 - $\text{O} + \text{N} > 10$
- **Be wary of any compound that violates *two or more* rules**
 - *combination* of $\log P$ and MW is non-obvious but critical
 - there are big drug molecules and greasy drug molecules, but relatively few orally available big greasy drug molecules

The Four Rules of 5

- Shortcomings
 - not particularly applicable to actively transported drugs
 - only modestly predictive for passively transported drugs ($r^2 < 0.4$)
 - in general, OH + NH → the number of hydrogen bond donors
 - in general, O + N → the number of hydrogen bond acceptors
- Many additional relevant and reliably predictive PK and toxicity models have become available since 1997
 - accurate solubility & pK_a prediction
 - mutagenicity, hepatotoxicity and hERG inhibition
 - CYP inhibition → problematic drug-drug interactions
 - susceptibility to CYP oxidation
- Strengths
 - easy to remember
 - mechanistically based → seems convincing → widely “applied”
 - *though the “two or more part” has often been forgotten*

Outline

1. An alternative absorption risk score
2. Defining ADMET Risk™
3. Generating a new qualified subset of the WDI
4. Establishing thresholds & rules
5. Observed distribution of risks
6. Applicability domains
7. Independence of rules
8. An application
9. Acknowledgments

An Alternative Absorption Risk Score

- M.B. Bolger, R. Fraczekiewicz & V. Lukacova.
 - Simulations of Absorption, Metabolism and Bioavailability. In: *Drug Bioavailability*; H. Van de Waterbeemd & B. Testa, eds.; Wiley-VCH, Weinheim, 2009, pp. 453-495.
- Alternative rules with thresholds derived from data on human *Fa* data from Y.H. Zhao *et al. J. Pharmaceutical Sciences* **2001**, *90*, 749-784.

<u>rule</u>	<u>weight</u>	<u>criterion</u>	
– <i>ow</i>	1.0	S+logP	< -0.891
– <i>Pf</i>	1.0	S+Peff	< 0.120
– <i>pH</i>	1.0	S+pH	< 3.40
– <i>SA</i>	1.0	T_PSA	> 139.9
– <i>HD</i>	1.0	HBDch	> 1.13 (charge on donor atoms)
– <i>AQ</i>	1.0	ABSQ	> 5.11 (absolute charge)
– <i>PC</i>	2.0	FormalQ	<> 0

- Ro5 rules

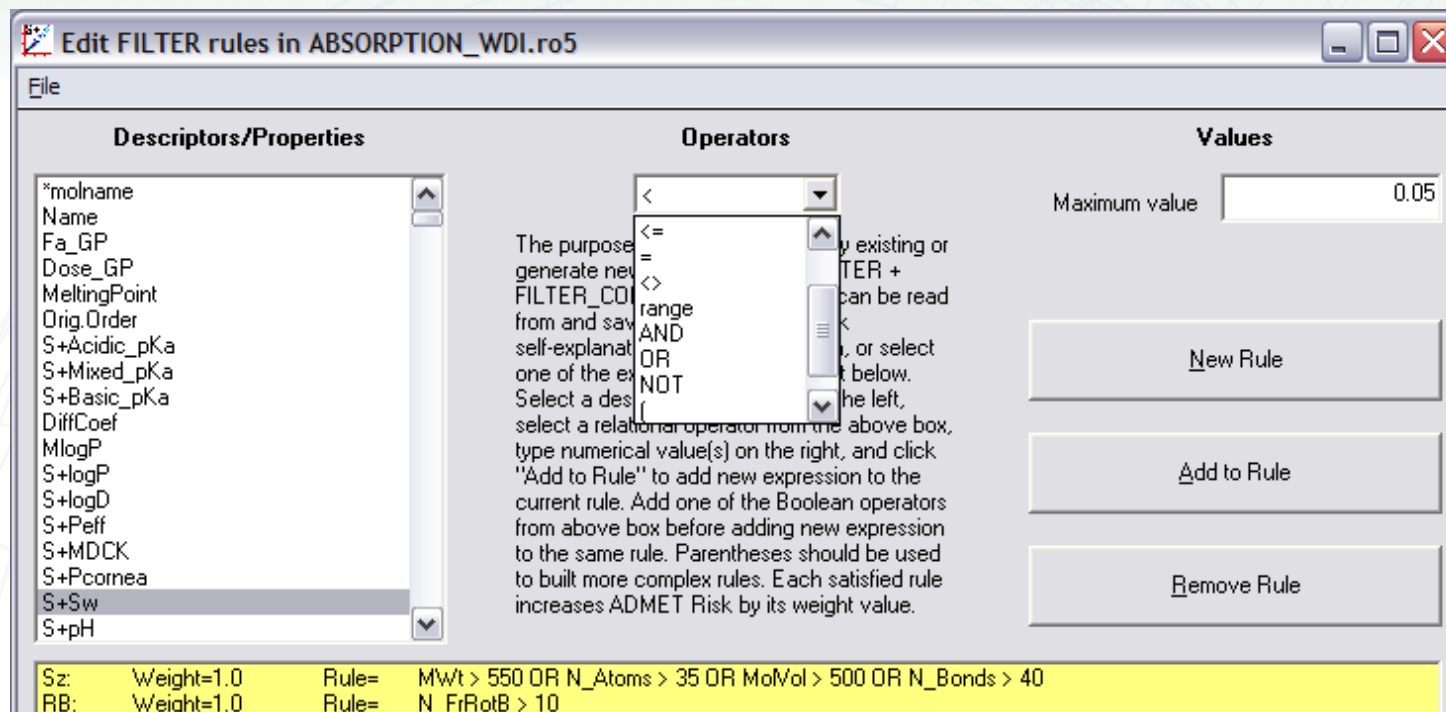
– <i>Mw</i>	1.0	MWt	> 500
– <i>LP</i>	1.0	MlogP	> 4.15 (Moriguchi logP)
– <i>Hb</i>	1.0	HBDH	> 5 (number of hydrogens on N or O)
– <i>NO</i>	1.0	M_NO	> 10 (number of N & O atoms)



What is ADMET Risk™?

- A flexible scoring system that encompasses many components of risk
 - absorption-related properties
 - pharmacokinetic properties
 - metabolism
 - toxicity
- A customizable set of (24 default) rules based on thresholds for ADMET Predictor™ property value estimates for a qualified subset of the World Drug Index (WDI)
 - thresholds set to pass 90% of commercial drugs
- An easy way for chemists to rapidly “see” in many dimensions by way of a single number
 - augmented by two-letter codes to identify which rules are violated
- A rapid way for chemists to compare and filter very large numbers of candidate molecules early in design and discovery

Defining S+Absn_Risk in ADMET Predictor™



Sz:	Weight=1.0	Rule=	MWt > 550 OR N_Atoms > 35 OR MoVol > 500 OR N_Bonds > 40
RB:	Weight=1.0	Rule=	N_FrRotB > 10
HD:	Weight=1.0	Rule=	HBDH > 4 AND HBDch > 1.8
HA:	Weight=1.0	Rule=	HBA > 9 AND HBAch < -5.8
ch:	Weight=1.0	Rule=	NPA_ABSQ > 21 OR T_PSA > 140
ow:	Weight=1.0	Rule=	S+logP > 5 OR S+logD > 4.1 OR MlogP > 4.1
Pf:	Weight=1.0	Rule=	S+Peff < 0.1 OR S+MDCK < 25
Sw:	Weight=1.0	Rule=	S+Sw < 0.005

numbers of votes "cast" when the rule is violated

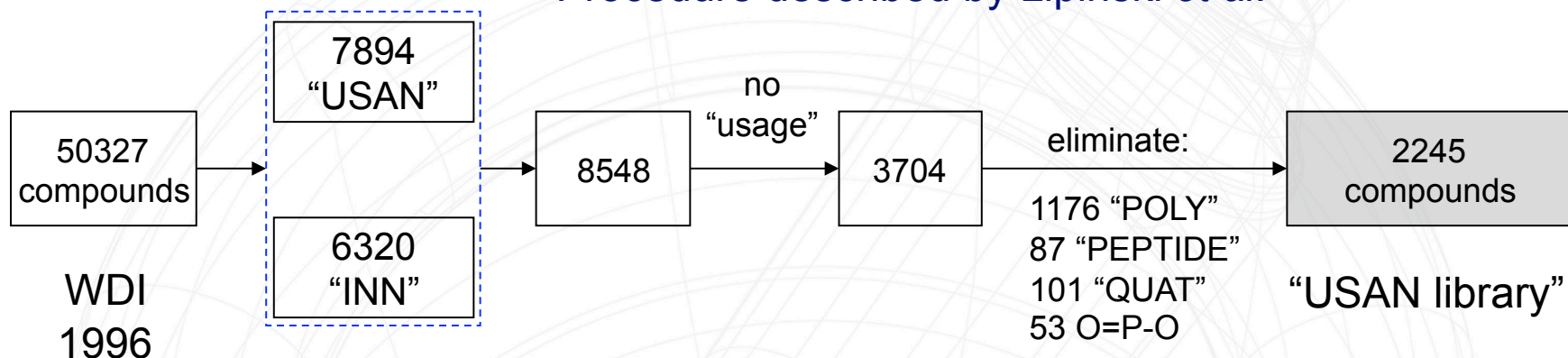
mnemonic to display when the rule is violated

simulationsplus,inc.

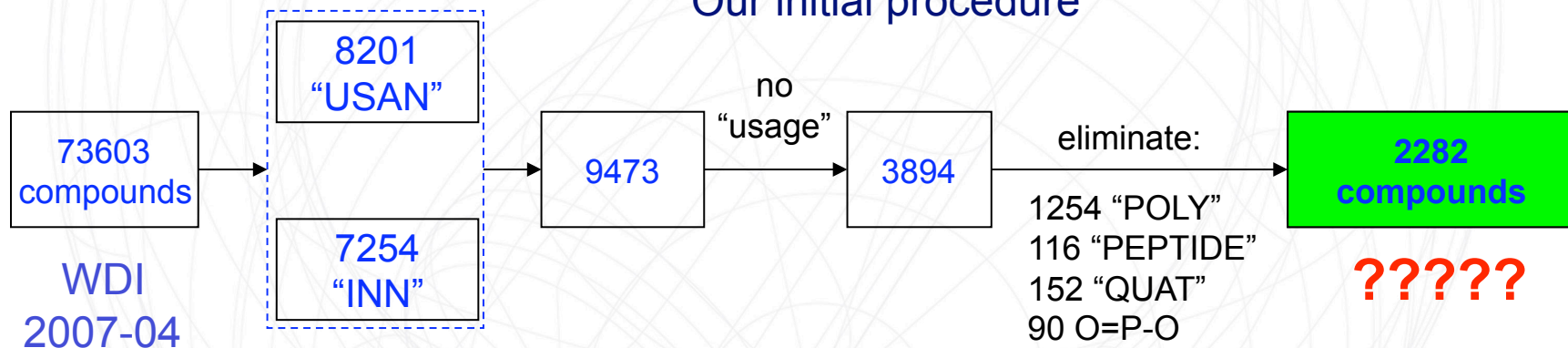


Qualifying a Subset of the World Drug Index

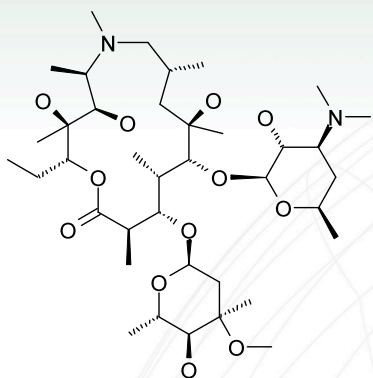
Procedure described by Lipinski *et al.*



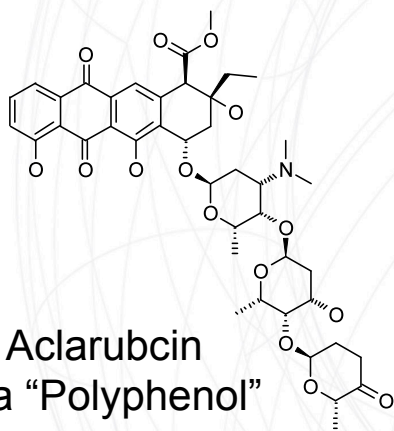
Our initial procedure



Entries Hit by “like %POLY%” Queries in the Substructure Field

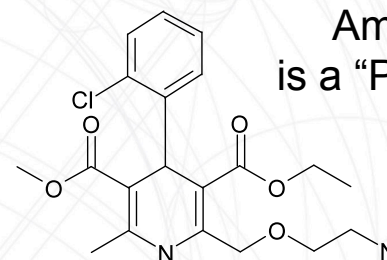


Azithromycin
is a “polyalcohol”

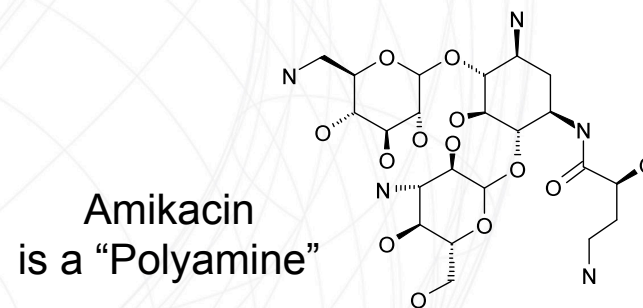


Aclarubcin
is a “Polyphenol”

- Polysaccharide (still dropped)
- Polyamine
- Poly-C-Acid
- Polyphenol
- Polyalcohol
- Polyolefin
- Polyketone
- ...



Amlodipine
is a “Poly-C-Acid”



Amikacin
is a “Polyamine”

Cannot be dropped
and still get ~10%
above MWt 500



Annotation evidently changed in ten years

simulationsplus,inc.



Inappropriate “Indicated Usages” That Were Dropped

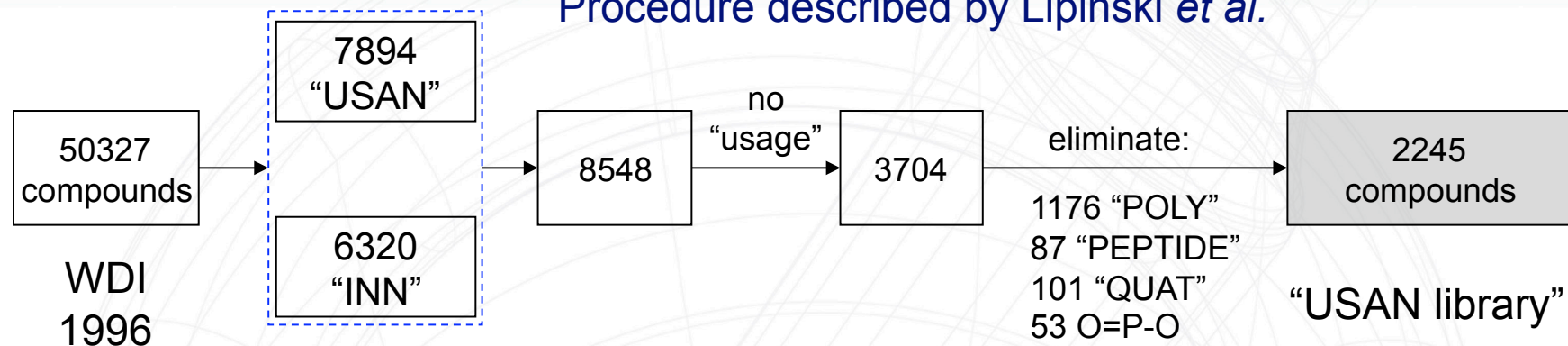
- Flavors
- Sweeteners
- Acidifiers
- Sedatives (CHBr_3)
- Imaging agents
- Preservatives
- Laxatives
- General anesthetics
($\text{ClCF}_2\text{CF}_2\text{Cl}$, cyclopropane)
- Laxatives
- Antioxidants (glycerol)
- Escharotics (KNO_3)
- Antacids
- Antiseptics
- Solvents
- Dietary supplements (arginine)
- Chelating agents
- Repellents
- Lubricants
- Dyes
- Emulsifiers
- Propellants
- Keratolytics (formaldehyde)
- Counterions from salts

Structural & Activity Classes That Were Removed

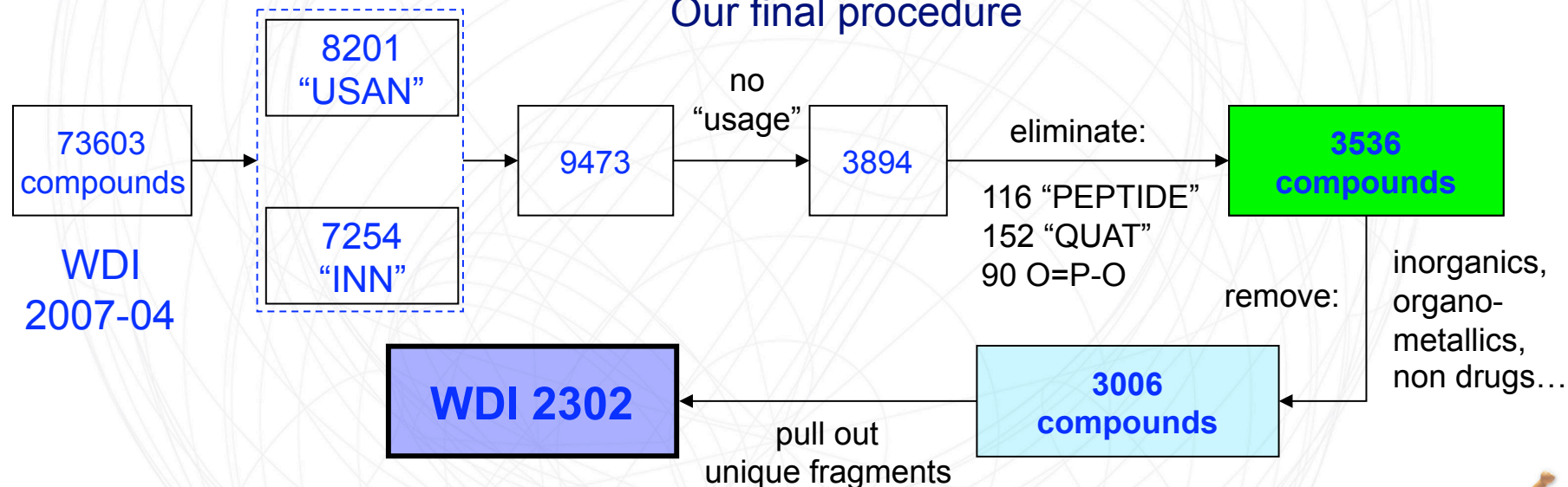
- Compounds with no C~C bond
 - Struct2D: sss: P~S
 - PT: like %INSECTICIDE%
 - PT: like %EMOLLIENT%
 - PT: like %ACIDIFIER%
 - PT: like %RUBEFACIENT%
 - PT: like %SWEETENER%
 - PT: like %CYTOSTATIC%
- AND MOL.WEIGHT <310
AND Struct2D: sss N[CH][CH][Cl,Br]

Qualifying a Subset of the World Drug Index

Procedure described by Lipinski *et al.*



Our final procedure



Rule of 5 Results: Lipinski's USAN Library vs. Our WDI Subsets

USAN dataset

- $N = 2245$
- Total N & O > 10: 12%
- MWt > 500: 11%
- MlogP > 4.15: 10%
- Total NH & OH > 5: 8%

WDI without duplicates

- $N = 2302$
- M_NO > 10: 8%
- MWt > 500: 11%
- MlogP > 4.15: 8.5%
- HBDH > 5: 4.7%

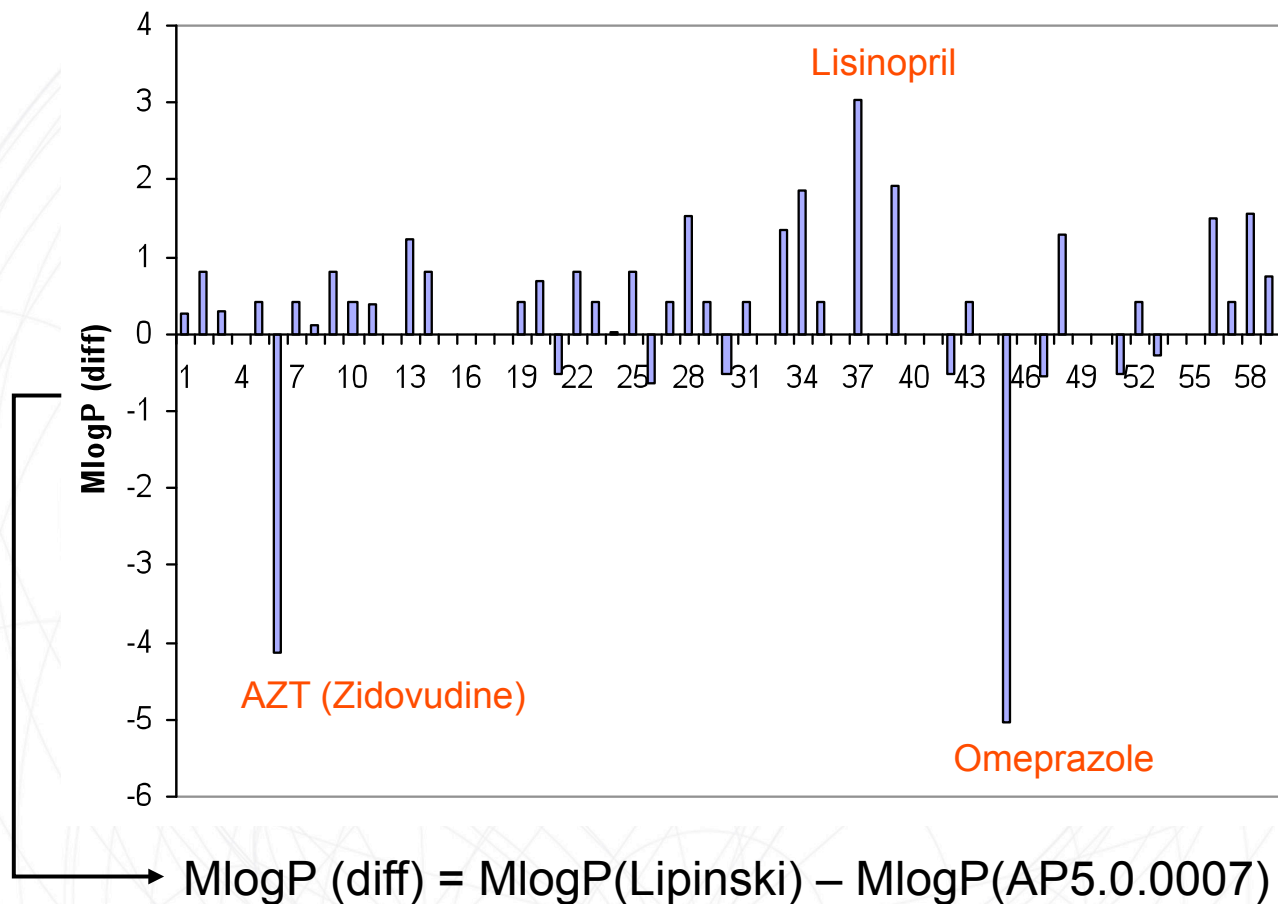
Why the Differences?

- *Not* counterion removal
- *Not* protonation/deprotonation
- *Not* redundancy
- *Maybe* MlogP calculation...

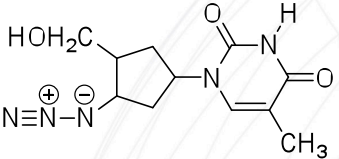
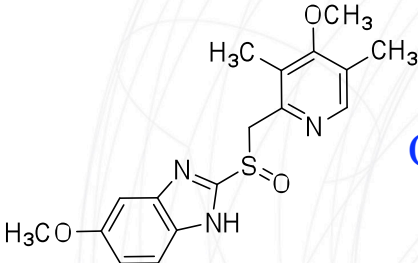
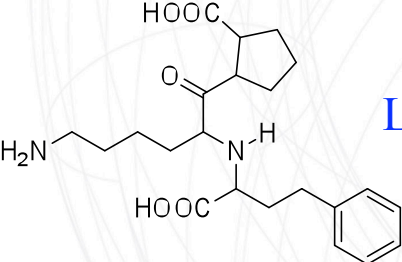
WDI with duplicates

- $N = 3006$
- M_NO > 10: 8%
- MWt > 500: 11%
- MlogP > 4.15: 8.2%
- HBDH > 5: 5.0%

Differences in Calculated MlogP Between Lipinski and ADMET Predictor™



Oops...

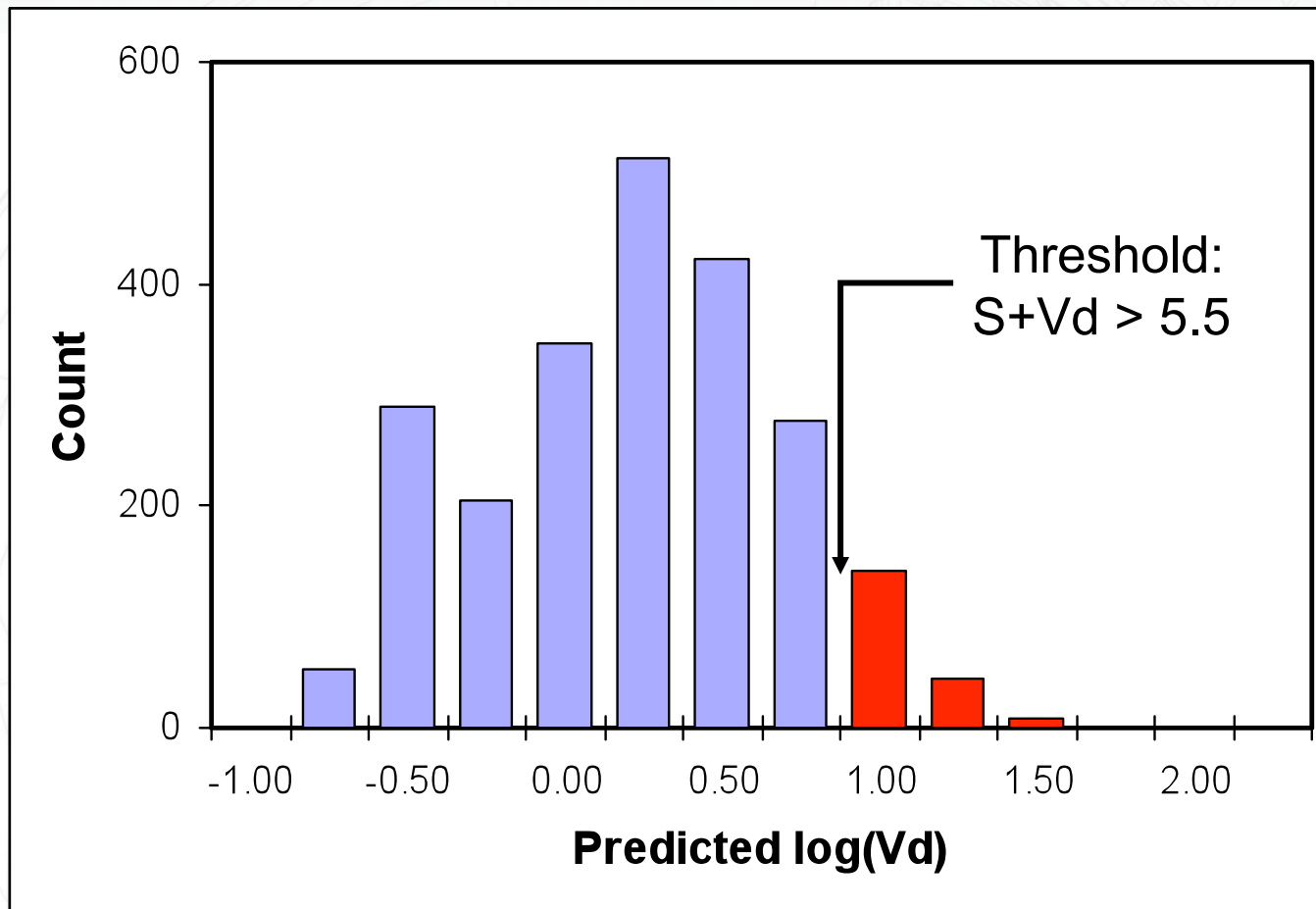
<u>Structure</u>	<u>Name</u>	<u>MlogP</u>		<u>Obsd</u> <u>logP</u>	<u>MWt</u>		
		<u>Lipinski</u>	<u>AP 5.5</u>		<u>Lipinski</u>	<u>AP 5.5</u>	<u>Actual</u>
	AZT	-4.38	0.38	0.05	267.3	267.3	267.3
	Omeprazole	-4.38	1.95	0.60	267.3	345.4	345.4
	Lisinopril	1.11	-1.92	-0.90	405.5	405.5	405.5



Curation, curation, curation...

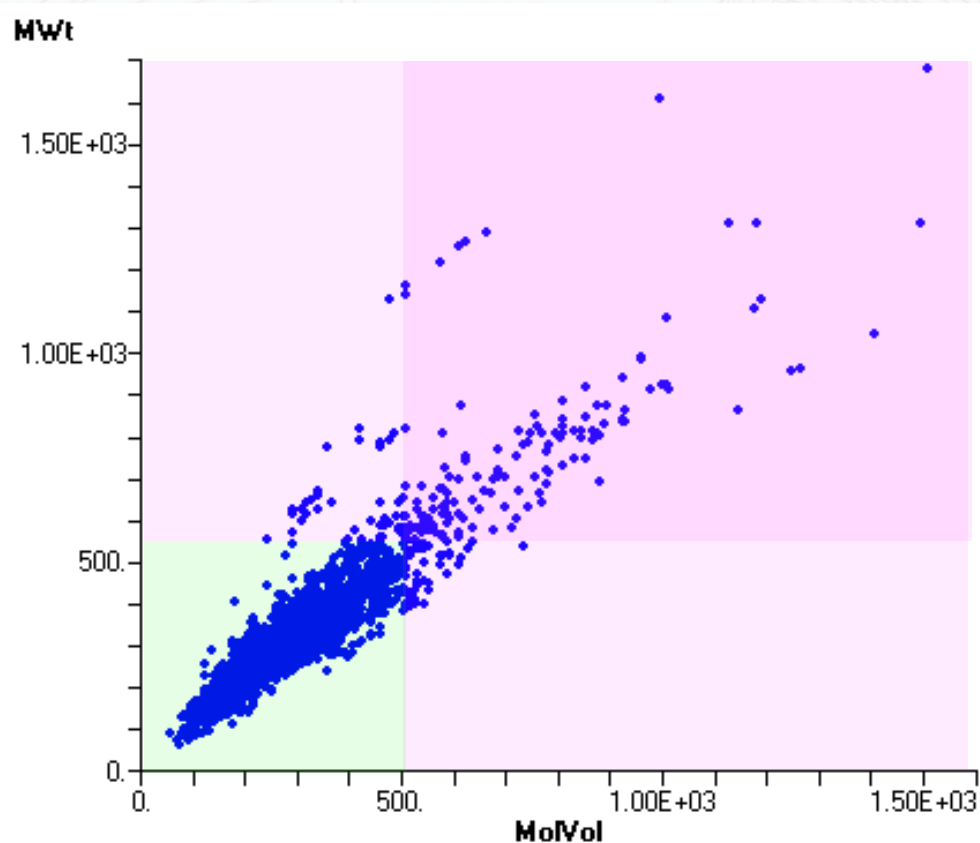


Setting an Individual Threshold



 Rule Vd: $S+Vd > 5.5$

Consolidating Correlated Properties



Rule Sz: MWt > 550 **OR** MoVol > 500 **OR** N_Atoms > 35 **OR** N_Bonds > 40



Default ADMET Risk Rules

S+Absn_Risk

- *Sz*: 1.0 MWt > 550 [Daltons] OR N_Atoms > 35 OR N_Bonds > 40
OR MolVol > 500 [cm³/mol]
- *RB* 1.0 N_FrRotB > 10
- *HD* 1.0 HBDH > 4 AND HBDch > 1.8
- *HA* 1.0 HBA > 9 AND HBACH < -5.8
- *ch* 1.0 NPA_ABSQ > 21 [sum] OR T_PSA > 140 [Å²]
- *ow* 1.0 S+logP > 5 OR S+logD > 4.1 OR MlogP > 4.1
- *Pf* 1.0 S+Peff < 0.1 [x10⁻⁴ cm/s] OR S+MDCK < 25 [x10⁻⁷ cm/s]
- *Sw* 1.0 S+Sw < 0.005 [mg/mL]

(pharmacokinetic risk rules)

- *fu* 1.0 S+PrUnbnd < 3.5 [%]
- *Vd* 1.0 S+Vd > 5.5 [L/kg]

Default ADMET Risk Rules (cont'd)

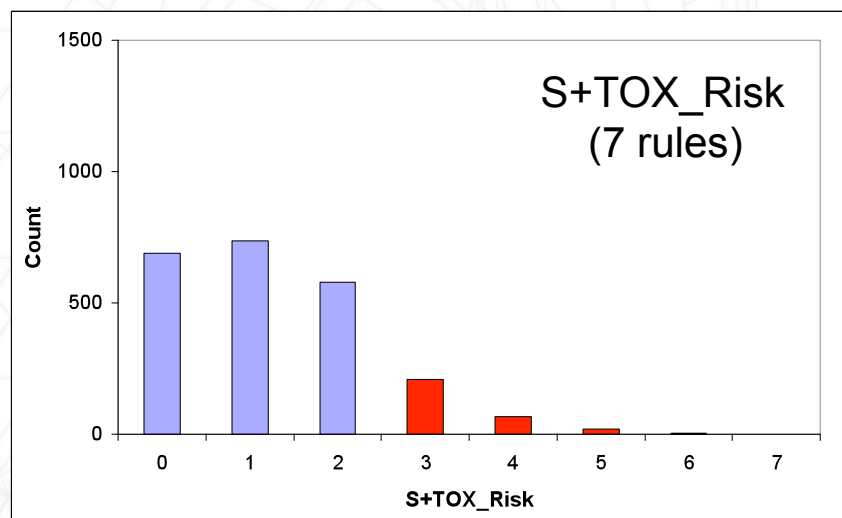
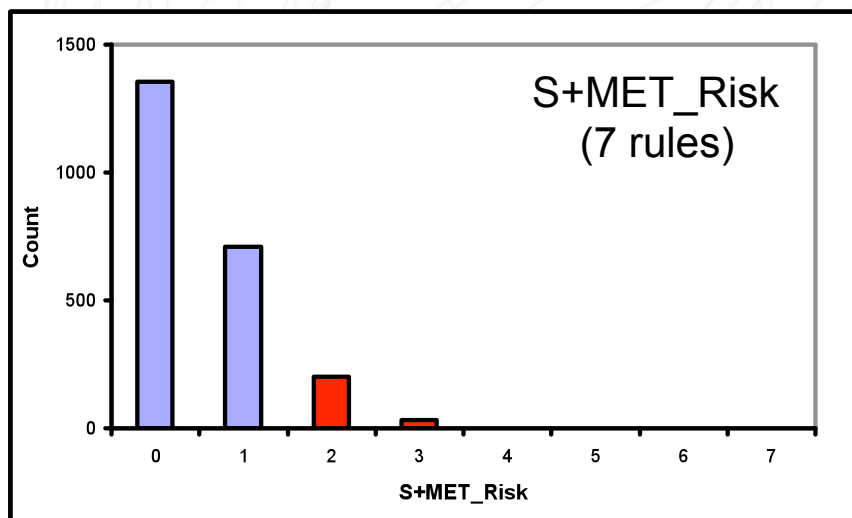
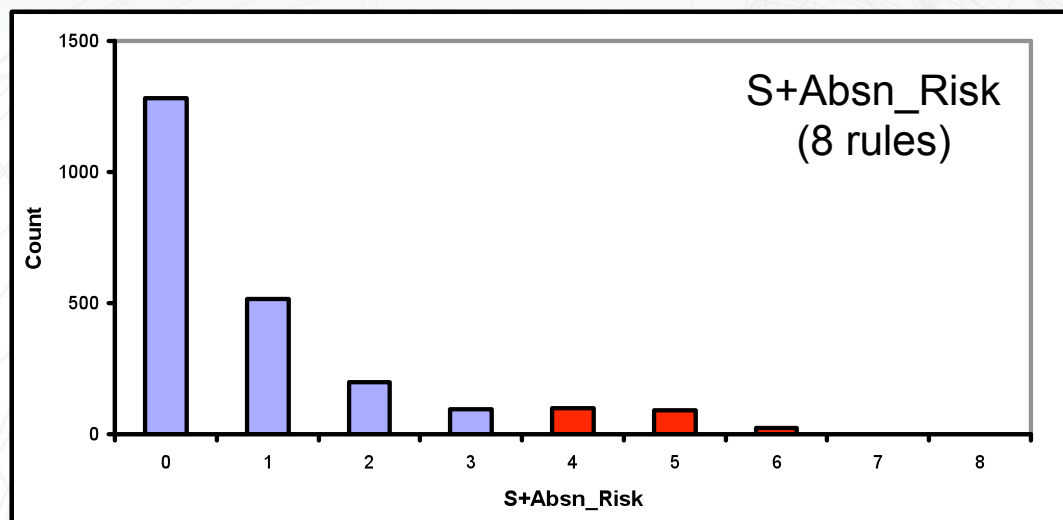
S+MET_Risk

- *IA* 1.0 CYP_1A2_Substr = Yes AND MET_1A2_CLint > 30 [$\mu\text{L}/\text{min}/\text{mg}$ protein]
- *I9* 1.0 CYP_2C19_Substr = Yes AND MET_2C19_CLint > 30 “ “
- *C9* 1.0 CYP_2C9_Substr = Yes AND MET_2C9_CLint > 30 “ “
- *D6* 1.0 CYP_2D6_Substr = Yes AND MET_2D6_CLint > 30 “ “
- *3A* 1.0 CYP_3A4_Substr = Yes AND MET_3A4_CLint > 30 “ “
- *mi* 1.0 MET_3A4_Ki_Mid < 1.5 [μM] AND
(MET_3A4_I_mid = Yes OR MET_3A4_Inh = Yes)
- *ti* 1.0 MET_3A4_Ki_tes < 1.0 [μM] AND
(MET_3A4_I_tes = Yes OR MET_3A4_Inh = Yes)

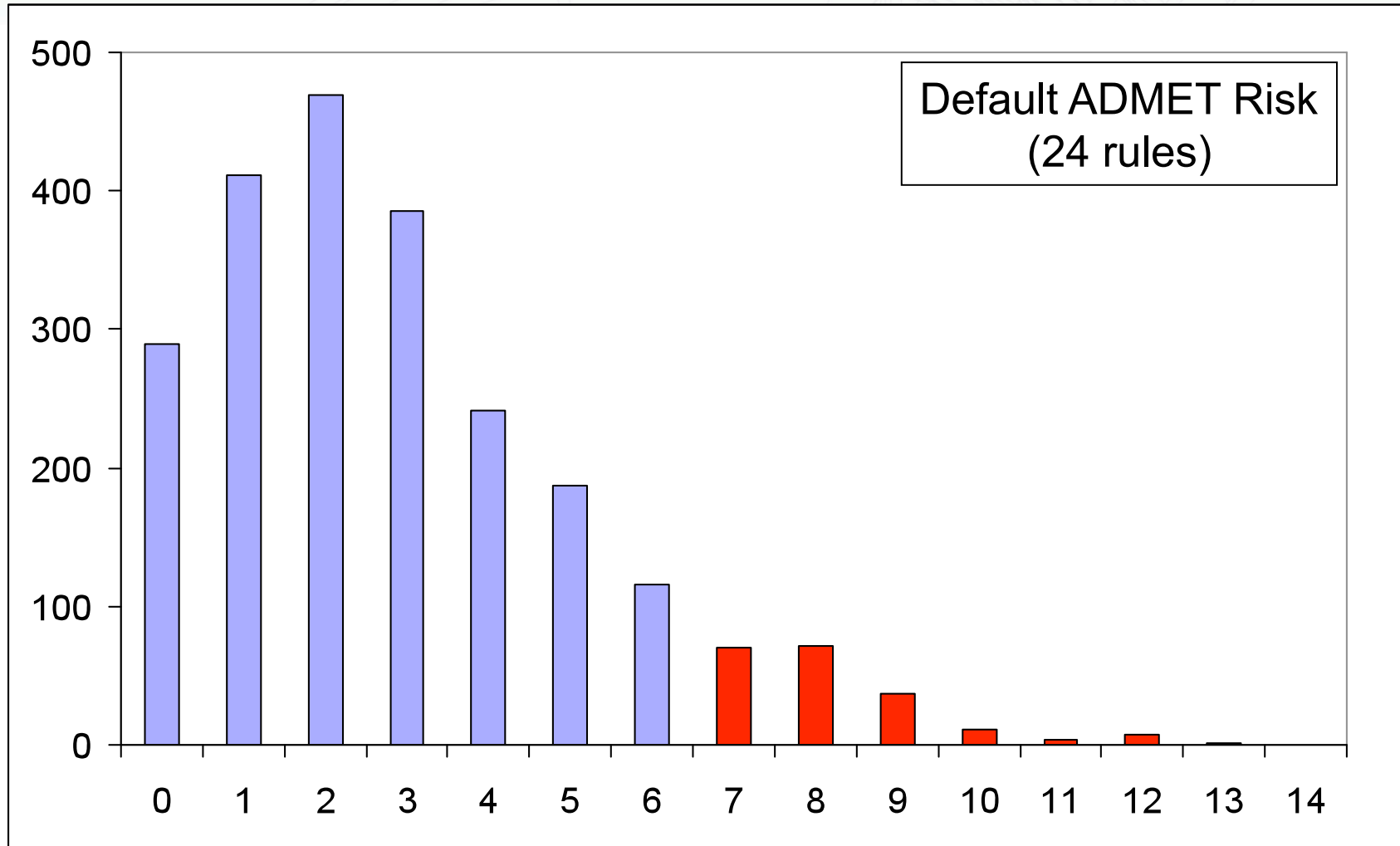
S+TOX_Risk

- *hE* 1.0 TOX_Herg > 6 [pIC50]
- *ra* 1.0 TOX_RAT < 300 [mg/kg oral]
- *Xr* 1.0 TOX_BRM_Rat < 4 [mg/kg/day]
- *Xm* 1.0 TOX_BRM_Mouse < 25 [mg/kg/day]
- *SG* 1.0 TOX_SGOT = Toxic AND TOX_SGPT = Toxic
- *Hp* 1.0 (TOX_AlkPhos = Toxic OR TOX_GGT = Toxic OR TOX_LDH = Toxic)
AND (TOX_SGOT = Toxic OR TOX_SGPT = Toxic)
- *Mu* 1.0 TOX_MUT_Risk > 2 [*in silico* Ames tests]

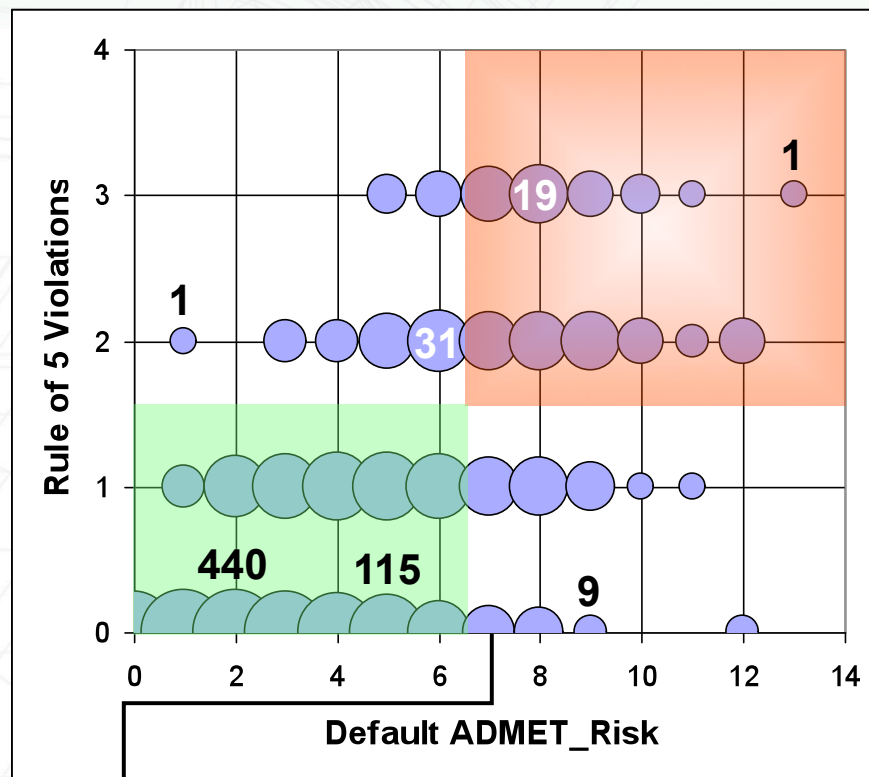
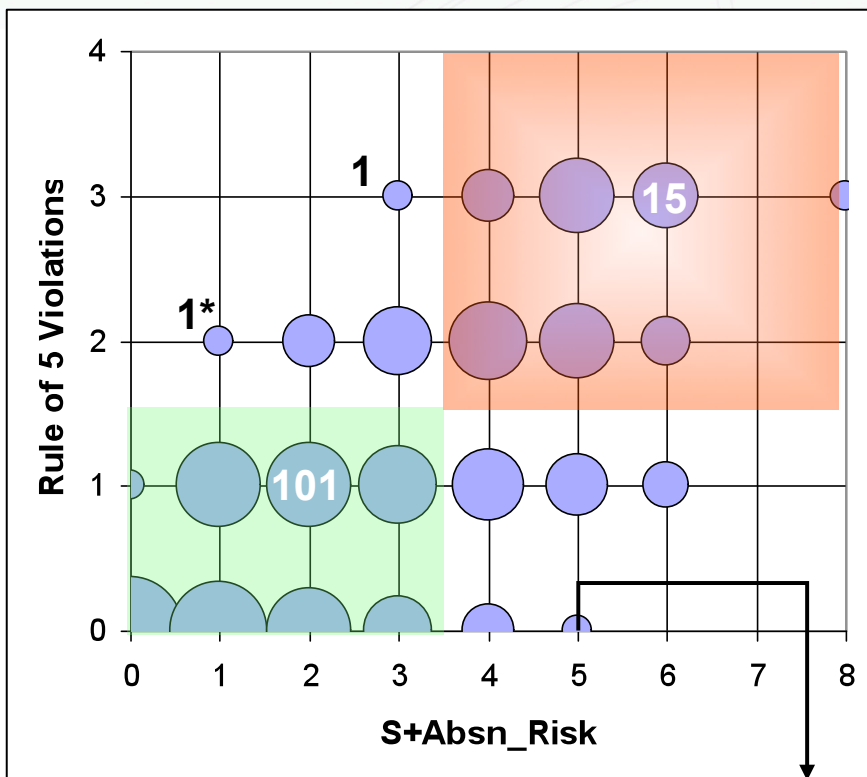
Distributions of ADMET Risk Across WDI 2302



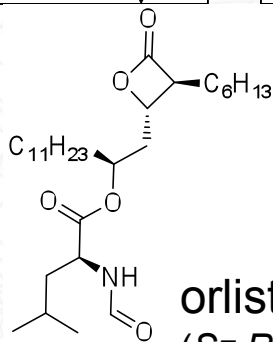
Distribution of Default ADMET Risk Across WDI 2302



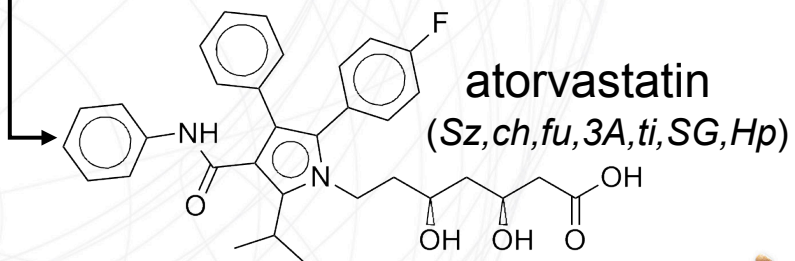
Relative Distributions of Rule Violations



*Symbol areas are scaled based on the logarithm of the frequency at the point at its center. Counts for a few points are included to show scale.



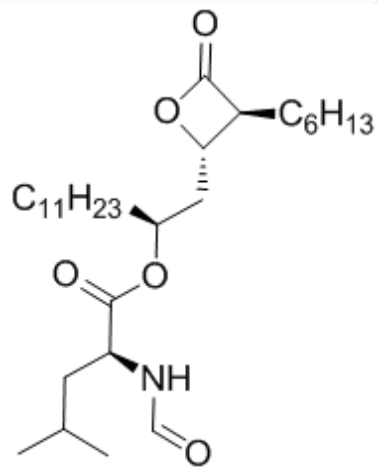
orlistat
(Sz, RB, ch, ow, Sw)



atorvastatin
(Sz, ch, fu, 3A, ti, SG, Hp)

What Kind of Drugs Are These?

orlistat violates **S+Abs_Risk** rules **Sz, RB, ch, ow & Sw**

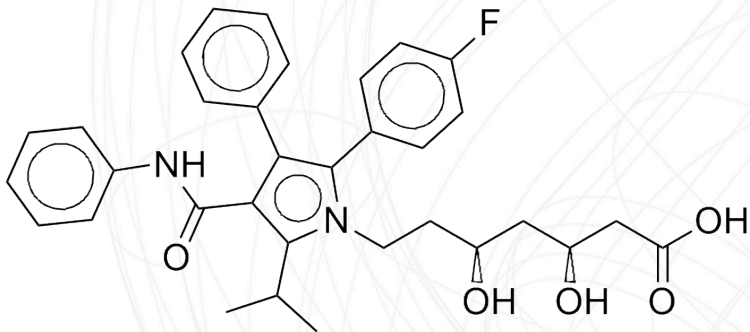


Orlistat is an obesity treatment that works by inhibiting lipases in the lumen of the stomach and small intestine, preventing hydrolysis of dietary fat into absorbable free fatty acids and monoglycerides. Undigested triglycerides are not absorbed.

“Systemic absorption of orlistat is minimal, however systemic absorption of the drug is not needed for activity.”

atorvastatin violates **ADMET_Risk** rules

Sz, ch, fu, 3A, ti, SG & Hp

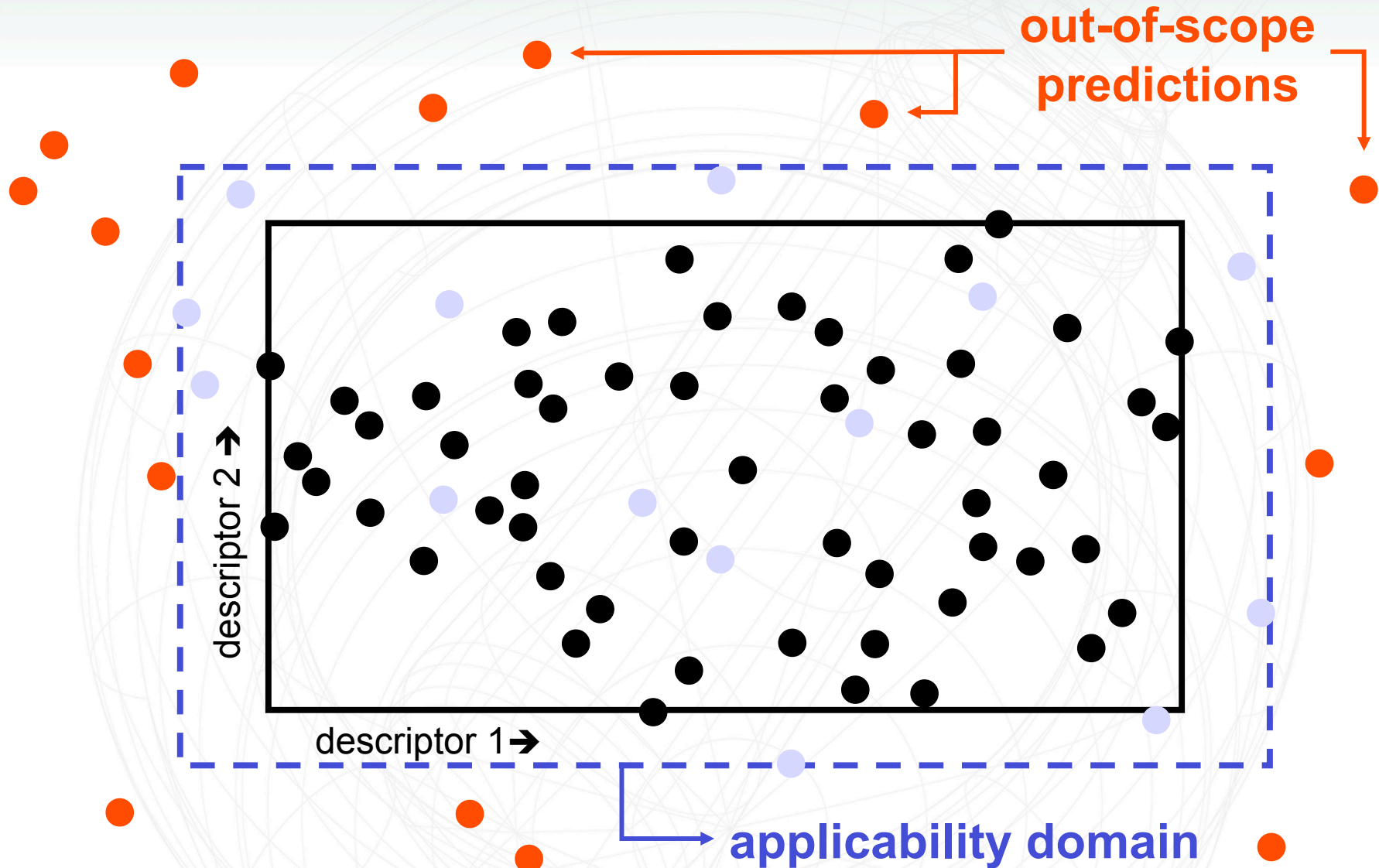


- Atorvastatin is 98% protein bound in plasma.
- It inhibits and is "extensively metabolized" by CYP 3A4 to metabolites that are themselves active
- “[P]ossible side effects include myotoxicity...and hepatotoxicity.”

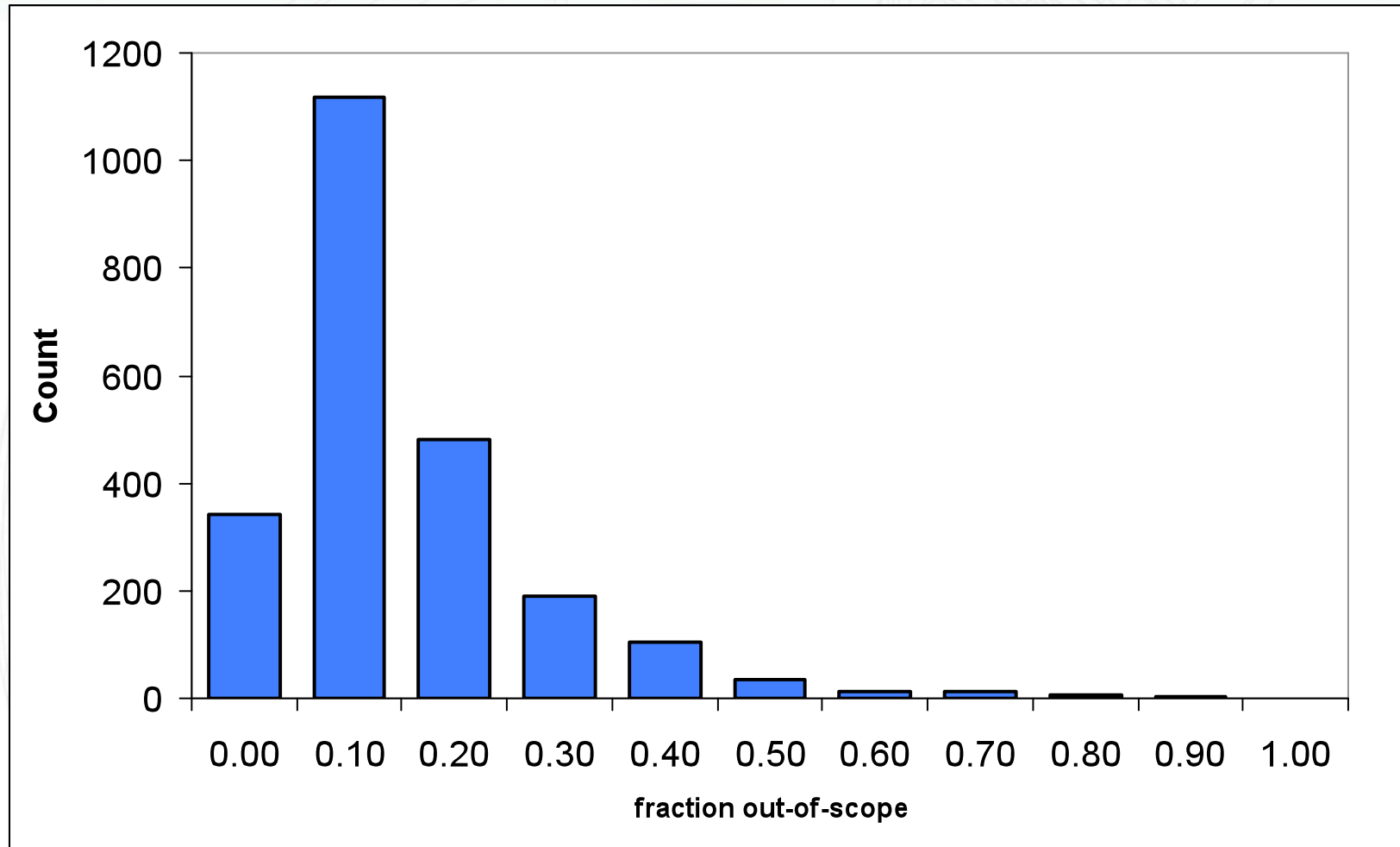
source: DrugBank <<http://www.drugbank.ca/drugs>>



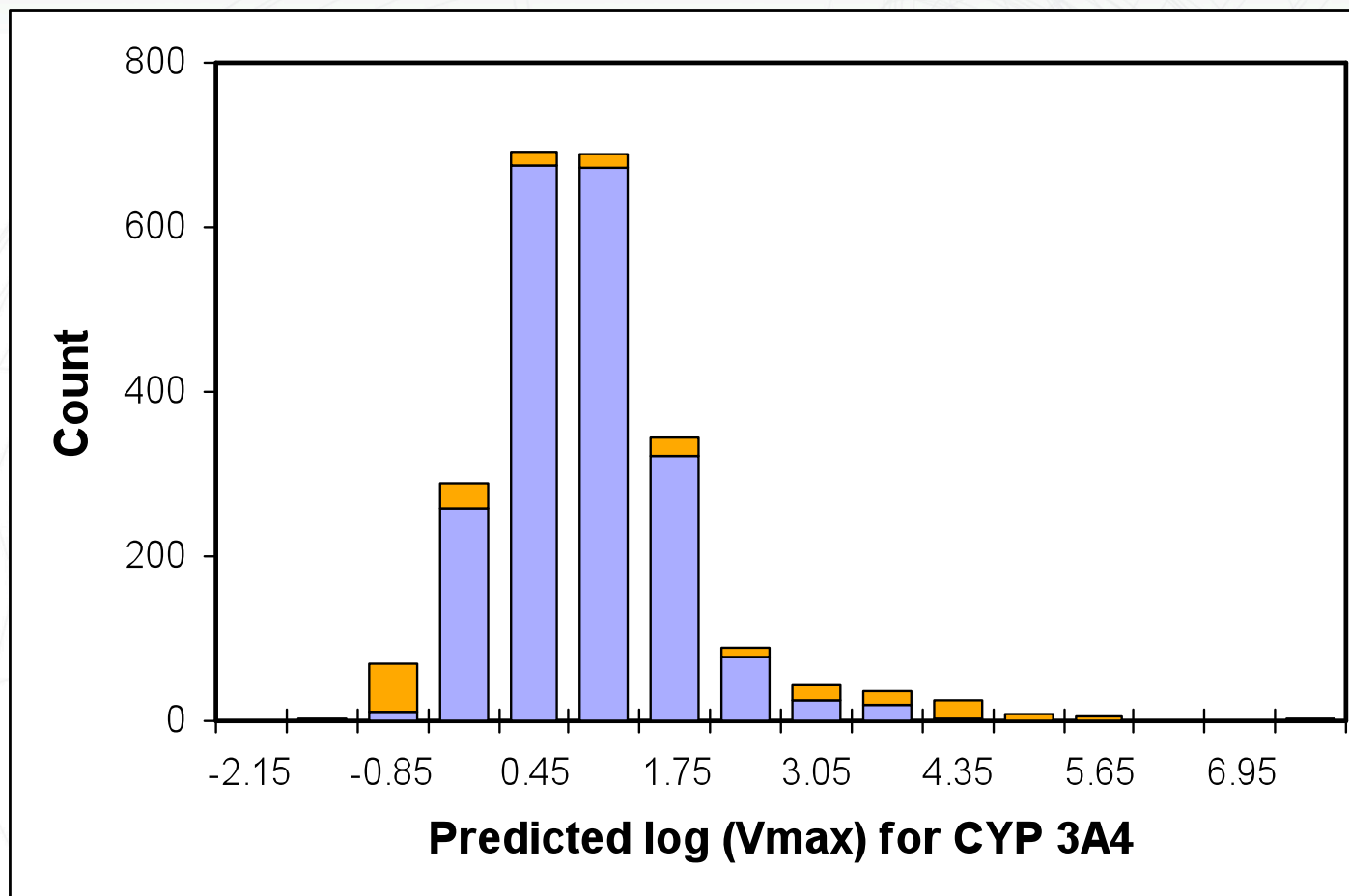
Applicability Domains



Out-of-Scope Predictions Are Relatively Uncommon



Applicability Domains: Typical Case



Model coverage: 90% with similar distributions

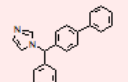
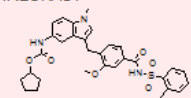
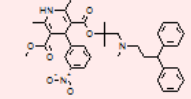
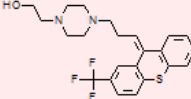
Some Drugs Really Do Have Bad Profiles: Among the Worst Offenders...

ADMET Predictor(TM) : GangOfFourBaddies.sdf (f:\admet risk)

File Batch Recalculate Options ADMET Modeler(TM) Help

Basic Modeler Settings Adv. Modeler Settings Ensemble Statistics Model Export

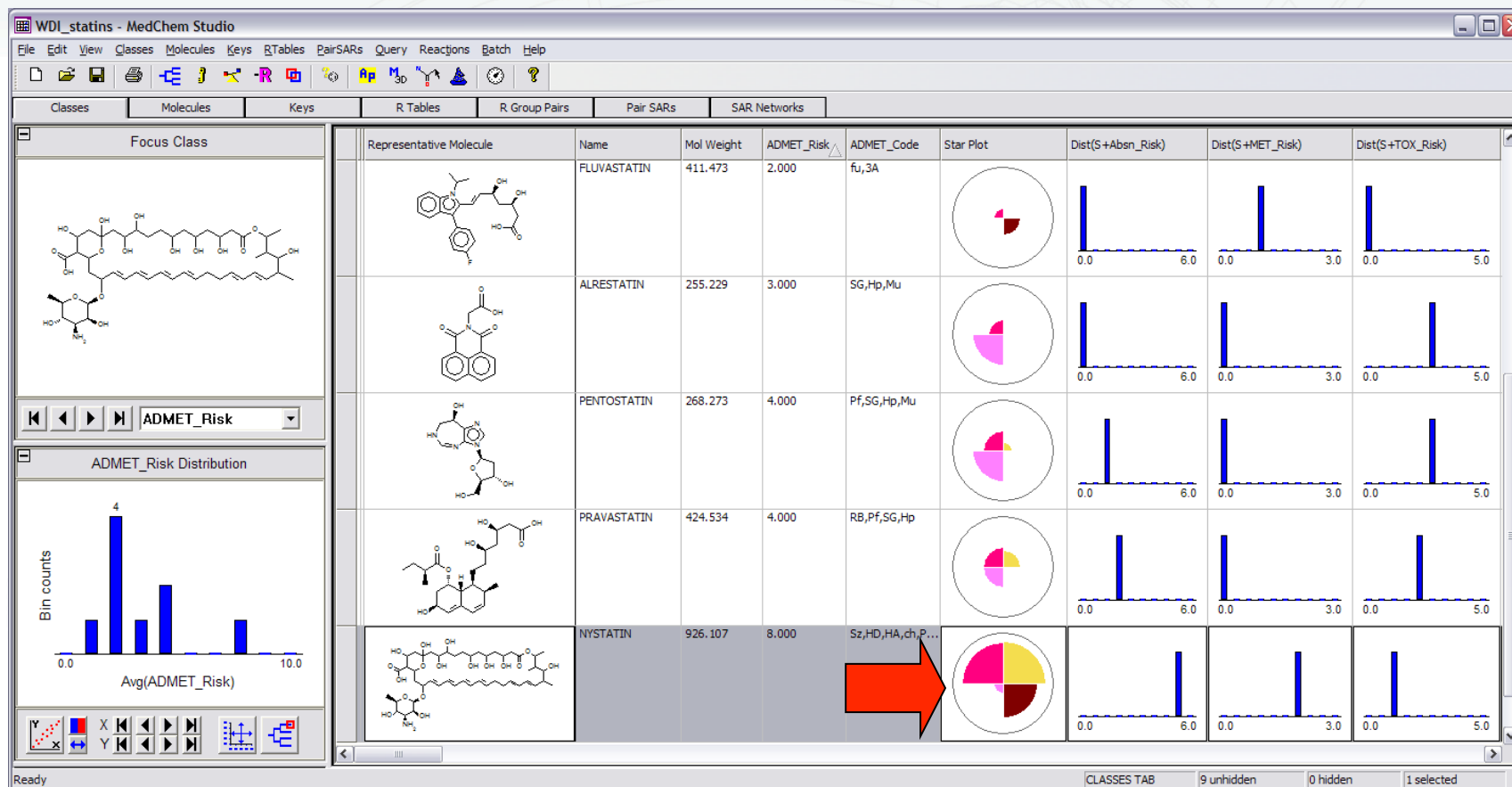
Molecular Data Prop./Desc. Histograms Prop./Desc. Correlations 4D Data Mining

*molname	RuleOf5	RuleOf5_Code	S+Absn_Risk	S+Absn_Code	S+MET_Risk	S+MET_Code	S+TOX_Risk	S+TOX_Code	ADMET_Risk	ADMET_Code
BIFONAZOLE 	0		2	ow,Sw	2	3A,ti	3	Xr,Xm,Hp	8	ow,Sw,fu,3A,ti,Xr,Xm,Hp ↑↑↑↑
ZAFIRLUKAST 	1	Mw	5	Sz,ch,ow,Pt,Sw	2	3A,ti	2	Hp,SG	10	Sz,ch,ow,Pt,Sw,fu,3A,ti,SG,Hp ↑↑↑↑↑↑↑↑
LERCANIDIPINE 	1	Mw	5	Sz,RB,ch,ow,Sw	3	3A,mi,ti	3	hE,Xr,Hp	11	Sz,RB,ch,ow,Sw,3A,mi,ti,hE,Xr,Hp ↑↑↑↑↑↑
FLUPENTIXOL 	0		1	Sw	3	1A,D6,3A	6	hE,ra,Xr,Xm,Hp,Mu	12	Sw,fu,Vd,1A,D6,3A,hE,ra,Xr,Xm,Hp,Mu ↑↑↑↑↑

Current Column = 0 4 reco **verified** 3M



“Statins” from qWDI-2302



Default ADMET_Risk

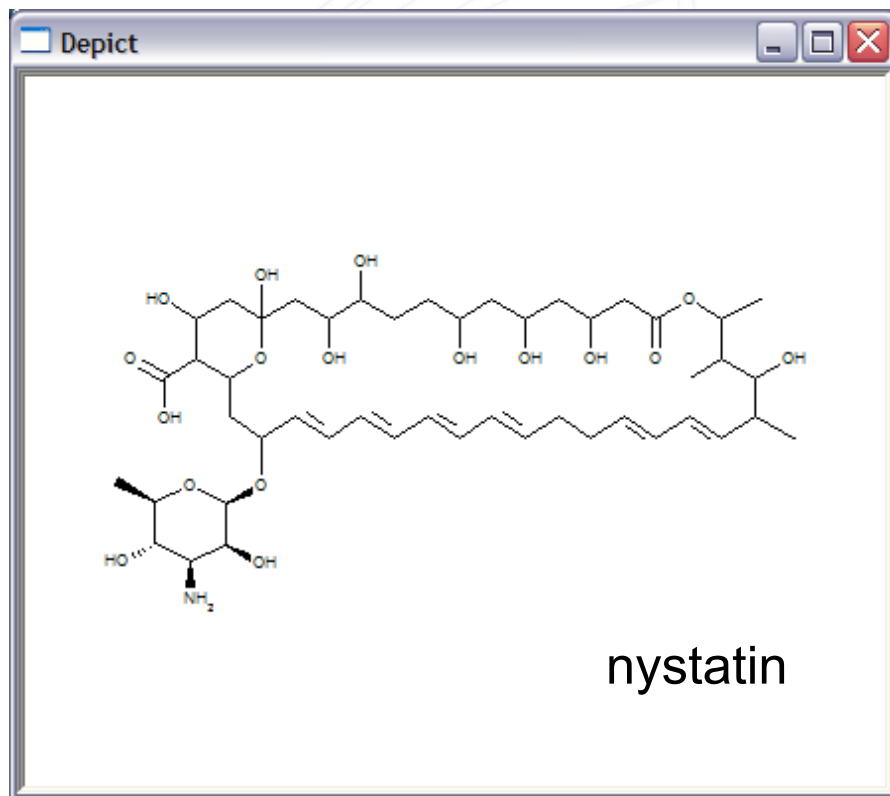
S+Absn_Risk

S+TOX_Risk

S+MET Risk



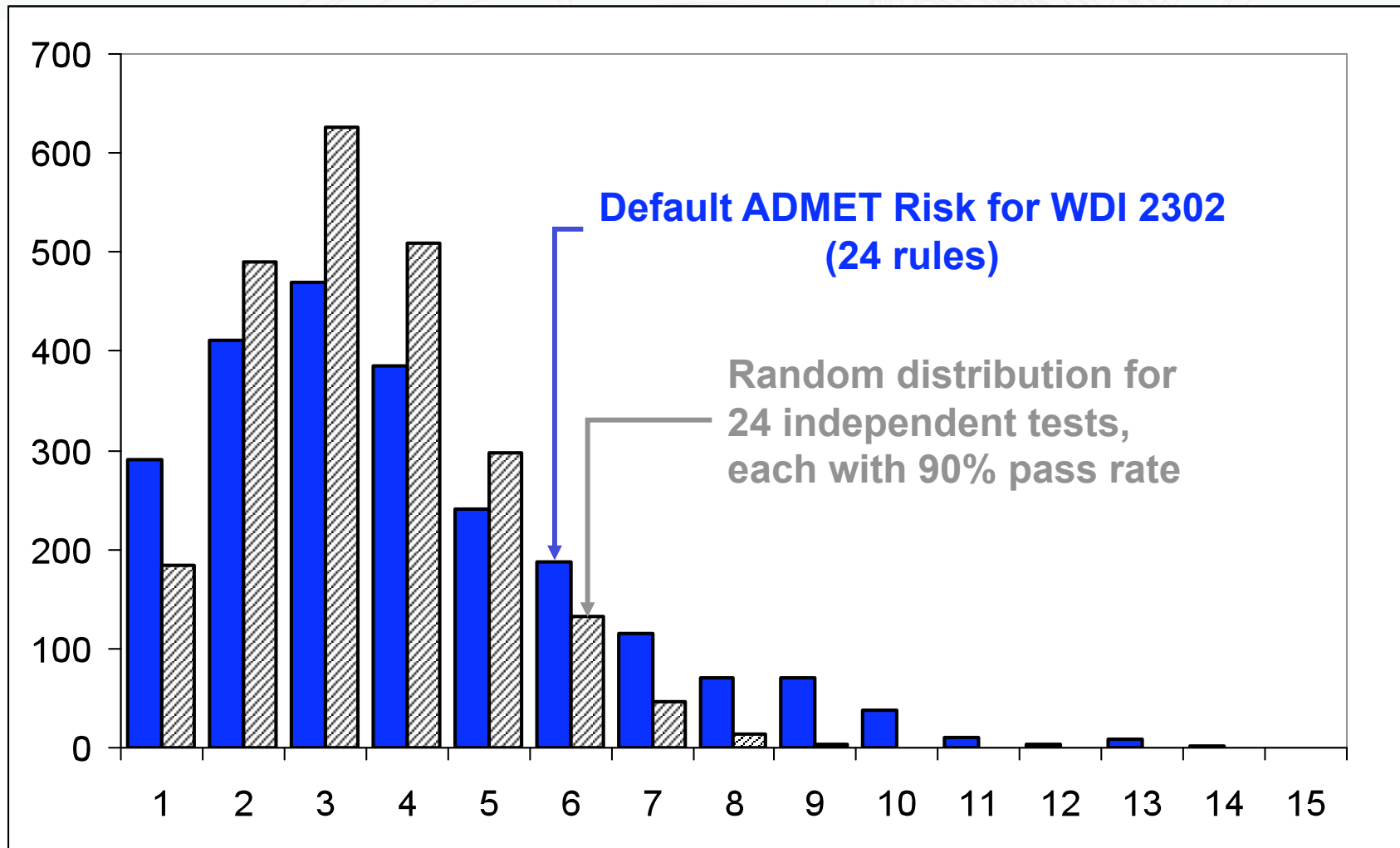
Nystatin Is Not Actually a Statin...



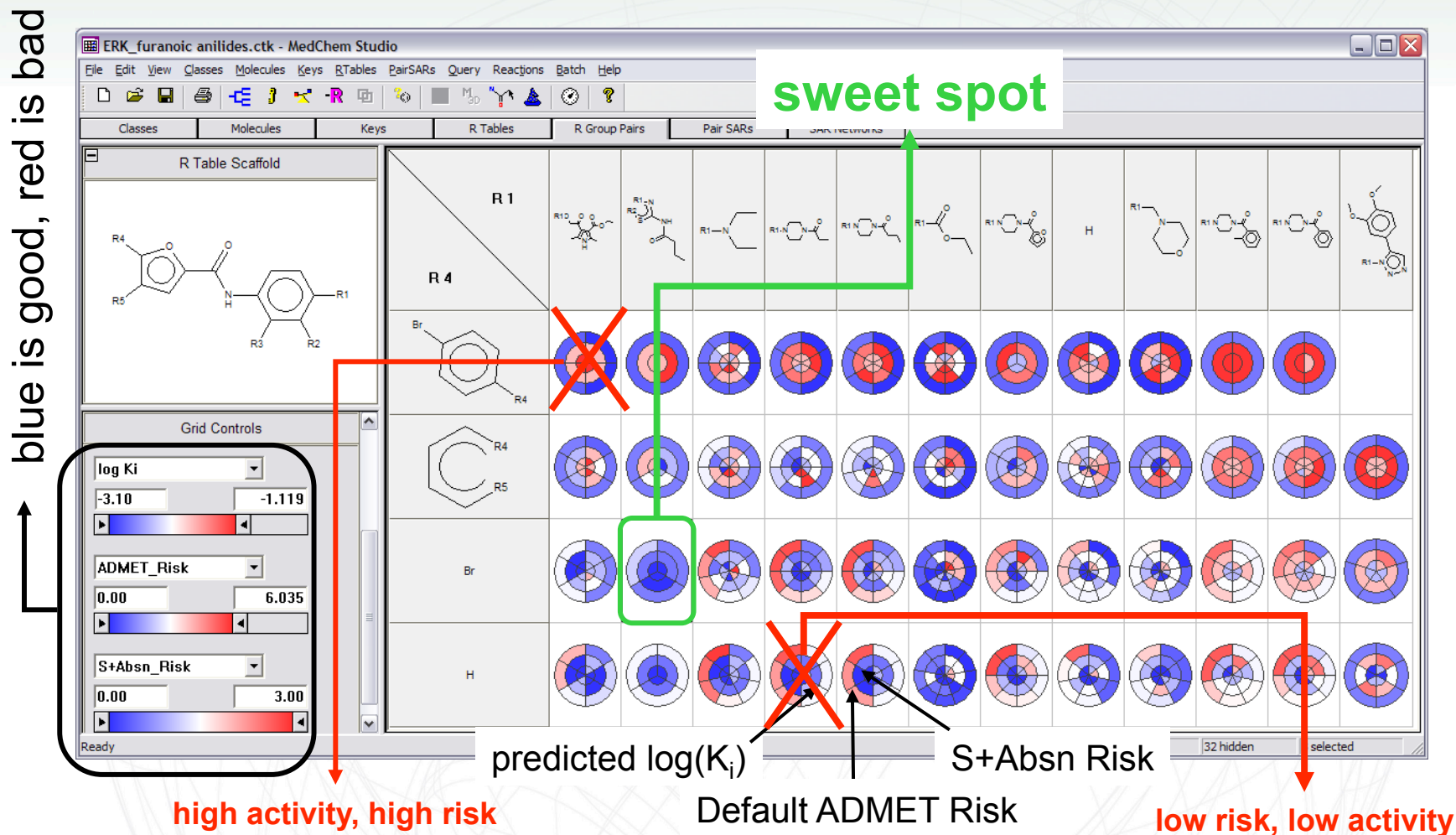
Wikipedia says:

“Due to its toxicity profile, there are currently no injectable formulations of [nystatin] on the US market. However, [it] may be safely given orally as well as applied topically due to its minimal absorption through mucocutaneous membranes such as the gut and the skin.”

Rules Are (Mostly) Independent



Applying ADMET Risk to Molecular Design*



*MedChem Studio™ R-Table Explosion based on PubChem Assay 995 for inhibition of ERK phosphorylation in whole cells.



Take-Home Messages

- The Rule of 5 is not the best indicator of oral absorption
- Accurate *in silico* ADMET predictions can be useful extensions to other filtering criteria that are associated only with oral absorption
 - ADMET Risk is a flexible scoring system easily adapted to your project
- Curation & attention to detail are as essential in constructing reference sets as in building QSAR models
- Model interdependencies & applicability domains need to be kept in mind when interpreting “flags” for potential problems
- Visualization of ADMET Risk distributions in sublibraries is a useful alternative to blind application of “hard” exclusion filters
- ADMET Risk codes are a useful, multi-dimensional way to help the synthesis chemist “see” beyond activity
 - *potential* problems with absorption, pharmacokinetics, metabolism, drug-drug interactions, and toxicities are readily highlighted
 - there are many hits, but it takes much more than potency to make a drug



Acknowledgments

- Jinhua Zhang
- Robert Fraczekiewicz
- Marvin Waldman
- Michael B. Bolger
- Walter S. Woltosz
- Dechuan Zhuang
- Adam Lee
- David Miller
- Aleksandra Mikosz
- Viera Lukacova...

Thank you...

...for your kind attention

<http://www.simulations-plus.com>

bob@simulations-plus.com