LIBRARIA

Building Drugs from Knowledge

A Consortium Approach to Predicting Toxicity

A Step Towards Building a
Larger Consortium Among Competitors

Steven Muskal, Ph.D.
VP Informatics
Libraria, Inc.
1/2003

ADRs are a Major Cause of Death

Deaths Per Year	Cause
710,000	Heart Disease
550,000	Cancer
170,000	Stroke
120,000	Pulmonary
100,000	***Adverse Drug Reactions***
98,000	Accidents
69,000	Diabetes
65,000	Pneumonia/Flu
50,000	Alzheimers
37,000	Nephritis
31,000	Septicemia

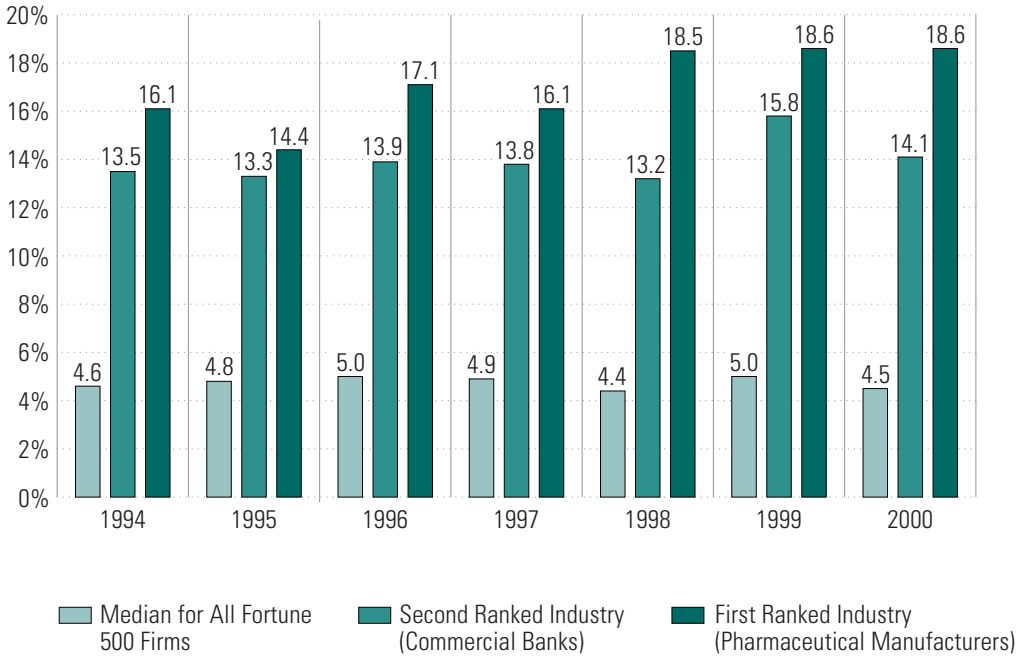
Refs: CDC Fastats estimated 2000 causes of death(http://www.cdc.gov/nchs/fastats/lcod.htm);
To Err Human, National Institute of Medicine, 1999; Bates et al., Incidence of adverse drug events and potential adverse drug events. JAMA 274:29, 1995;
Porter & Jick, Drug-related deaths among medical inpatients. JAMA 237:879-281, 1977.

See: http://www.drugintel.com/pharma/cause_of_death.htm

Profitability Among Pharmaceutical Manufacturers Compared to Other Industries, 1994–2000

exhibit

32



note

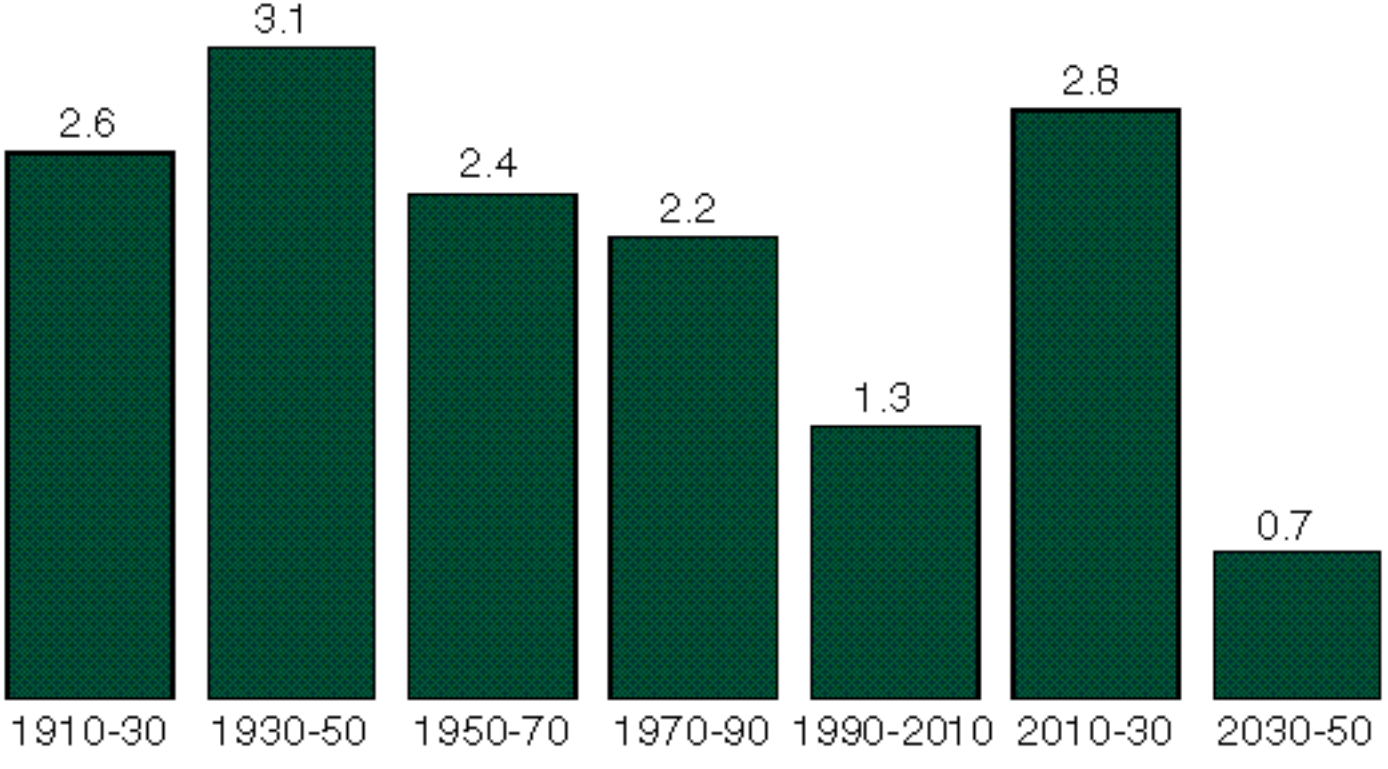
Percent shown is the median percent net profit after taxes as a percent of firm revenues for all firms in the industry. The second ranked industry each year was commercial banks.

source

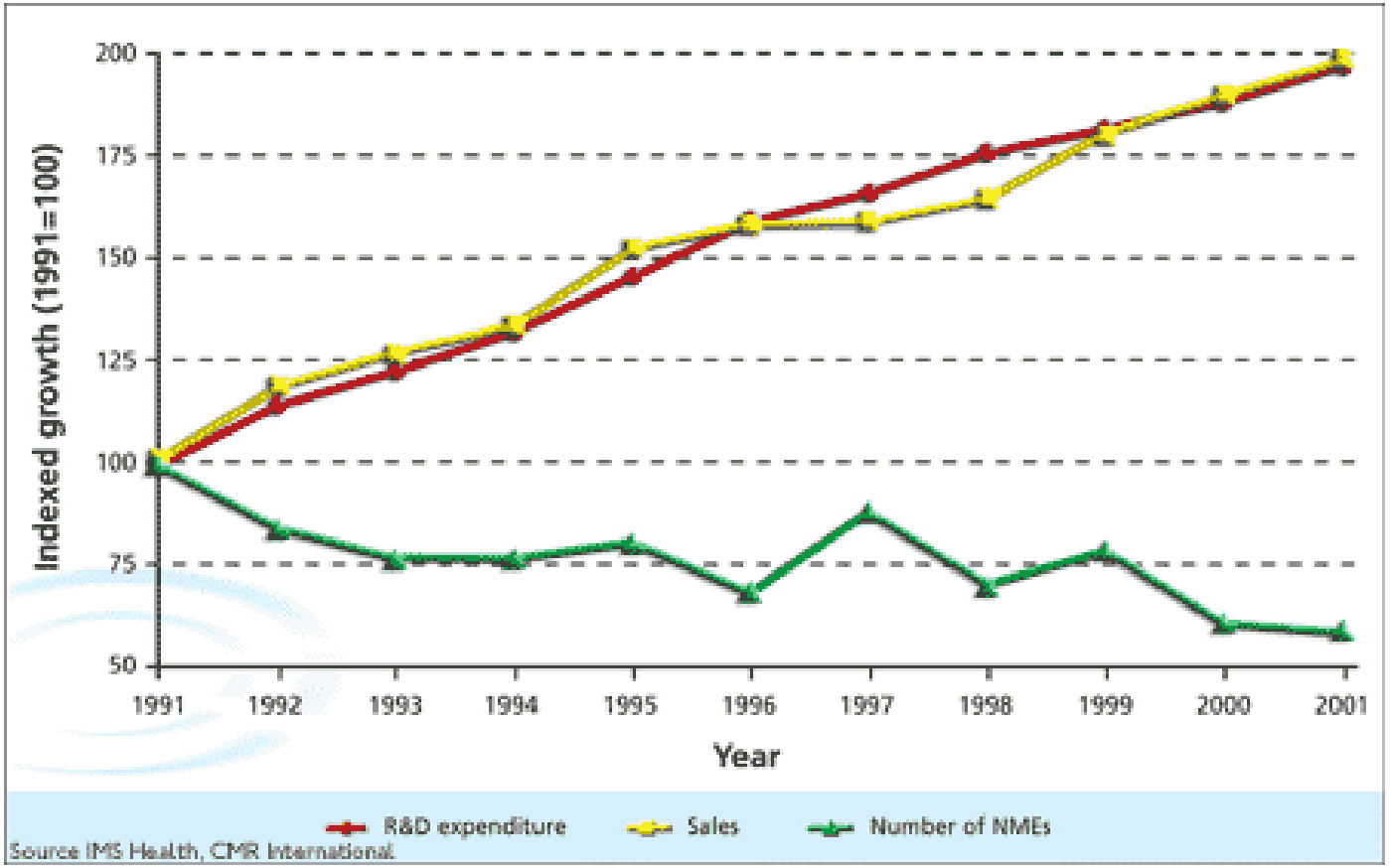
Fortune 500 Industry Rankings, *Fortune*, April issues, various years.

Fifteen Years From Now, Elderly Population Growth Will Explode

Average annual growth rate (in percent) of the elderly population: 1910-30 to 2030-50



Not a Pretty Picture



See: http://www.cmr.org/pdfs/springnews2002.pdf

Is Greed Always Good?

“It is not from the benevolence of the butcher, the brewer or the baker that we expect our dinner, but from their regard to their own self interest.... [Every individual] intends only his own security, only his own gain. And he is in this led by an invisible hand to promote an end which was no part of his intention. By pursuing his own interest, he frequently promotes that of society more effectually than when he really intends to promote it.” -Adam Smith, 1776

Tragedy Of The Commons

“Picture a pasture open to all. It is expected that each herdsman will try to keep as many cattle as possible on this commons. ... Therein is the tragedy. Each man is locked into a system that compels him to increase his herd without limit, in a world that is limited. Ruin is the destination toward which all men rush, each pursuing his own best interest in a society that believes in the freedom of the commons.” -Garrett Harding

A “Beautiful” Concept – Nash Equilibrium

DEFINITION: If there is a set of strategies with the property that no player can benefit by changing her strategy while the other players keep their strategies unchanged, then that set of strategies and the corresponding payoffs constitute the Nash Equilibrium.

A Nash Equilibrium* will be reached when each agent's actions begets a reaction by all the other agents which, in turn, begets the same initial action. In other words, the best responses of all players are in accordance with each other.

Can pharmaceutical players reach a Nash Equilibrium?!

Ref: "Equilibrium points in N-Person Games", Nash, JF. 1950, Proceedings of NAS.

Pharma is Living the Tragedy of the Commons

- Pharma is not more efficient despite increased R&D spending
- Mergers don't seem to help (Is in-licensing a form of denial?)

Can a business maintain *viability* when it needs to spend >$800M over 12 years* to develop a product?

Ref: Outlook 2002. Tufts Center for the Study of Drug Development

Enough of the Rant

While it might benefit all if Pharma conducts research “in accordance” with one another, many barriers remain.

Sharing toxicity Information can be a giant a “baby step”

Some precedence – The International Toxicology Information Center (ITIC) consortium (International Uniform Chemical Information Database (IUCLID)

Toxicity Prediction Software Packages are Available

- DEREK - (J. E. Ridings *et al.*, *Toxicology* 106, 267-79. (1996)
- CASETOX/MULTICASE (G. Klopman, *J Chem Inf Comput Sci* 38, 78-81. (1998)
- TOPKAT (http://www.accelrys.com/products/topkat/)
- CSGenoTox (http://www.chemsilico.com/)
- TOXSYS (http://www.scivision.com/ToxSys.html)
- HazardExpert/ToxAlert (http://www.compudrug.com/)
- OncoLogic – (http://www.logichem.com/)

A Simple Toxicity Assessment Strategy

- **Premise**: A compound's toxicity (e.g. pLD50) can be gauged based on the toxicities of other structurally similar compounds.

  $ refsim ref.tdt {p_}ref-propfield ref-namefield sim-cutoff #members {LOO} {numref} < qry.tdt

- **Algorithm**:
  - For each qry-mol in “qry.tdt,” identify ref-mol(s) in “ref.tdt” within “sim-cutoff” tanimoto similarity.
  - Calculate average (“ARP”) of “ref-propfield” for ref-mol(s).
  - Assign “ARP” to qry.mol, provided at least “#member” consortium members were used in calculation.
- **Requirement**: reasonably large, descriptive reference set

  Can we come together to build a predictive reference set?

  {What if only fingerprints/toxprops were “shared?” –J.Delaney.}

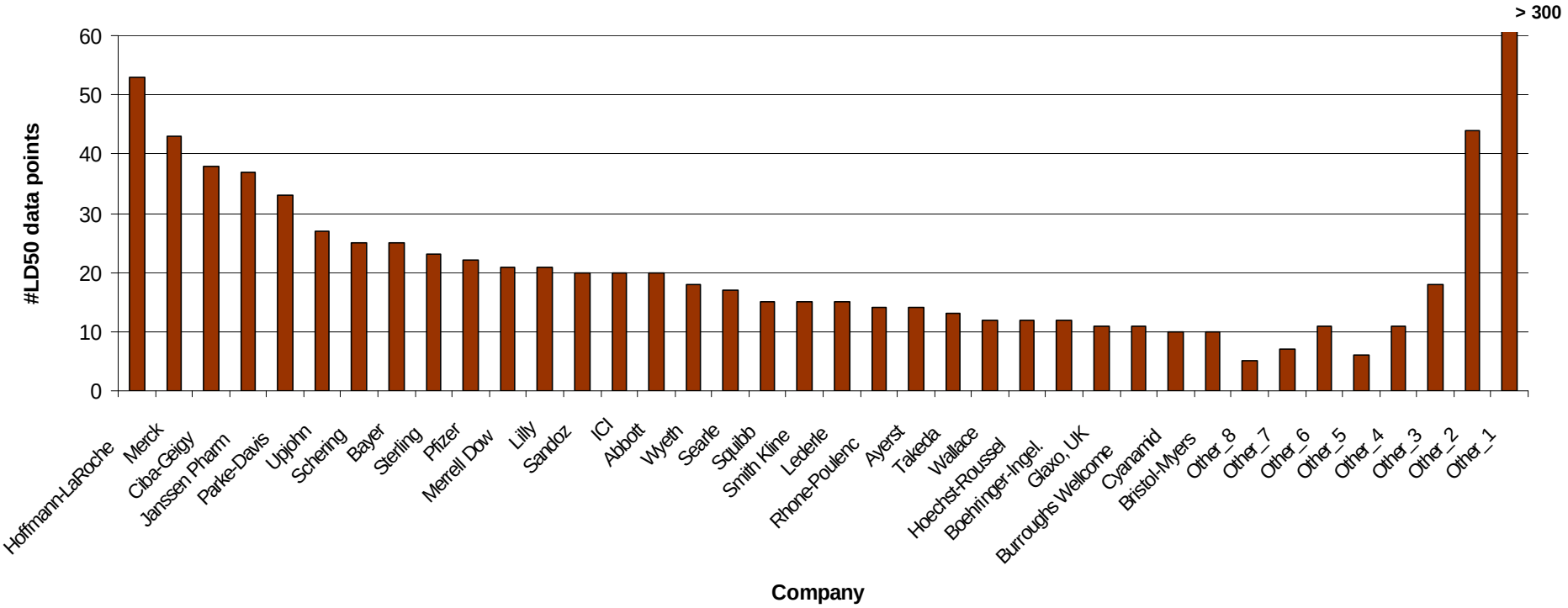
A Starter Reference Set – “RefSet”

- **Starter Source**: RTECS – over 133K compounds
  - http://www.nisc.com/factsheets/qrtc.htm
  - http://www.mdli.com/products/rtecs.html
- **Subset** - 13645 examples
  - Route: Oral Species: Rat EndPoint: LD50
- **MWT**: Avg: 304.09 Std: 183.78
- **ClogP**: Avg: 2.05 Std: 2.53
- **QPlogS**: Avg: -2.95 Std: 2.37
- **TPSA**: Avg: 65.28 Std: 70.39
- **RotBonds**: Avg: 5.32 Std: 5.5

Drug Subset- “DrugSet”

- Comprehensive Medicinal Chemistry – CMC 2002.1 (MDL)
- Approx. 8500 compounds tested in/on man
- “DrugSet” subset: 1781 “RefSet” compounds found in CMC
- **MWT**: Avg: 367.4 Std: 95.5
- **ClogP**: Avg: 2.87 Std: 2.7
- **QPlogS** Avg: -3.97 Std: 2.1
- **TPSA**: Avg: 72.24 Std: 48.3
- **RotBonds**: Avg: 5.96 Std: 4.2

DrugSet Source Companies

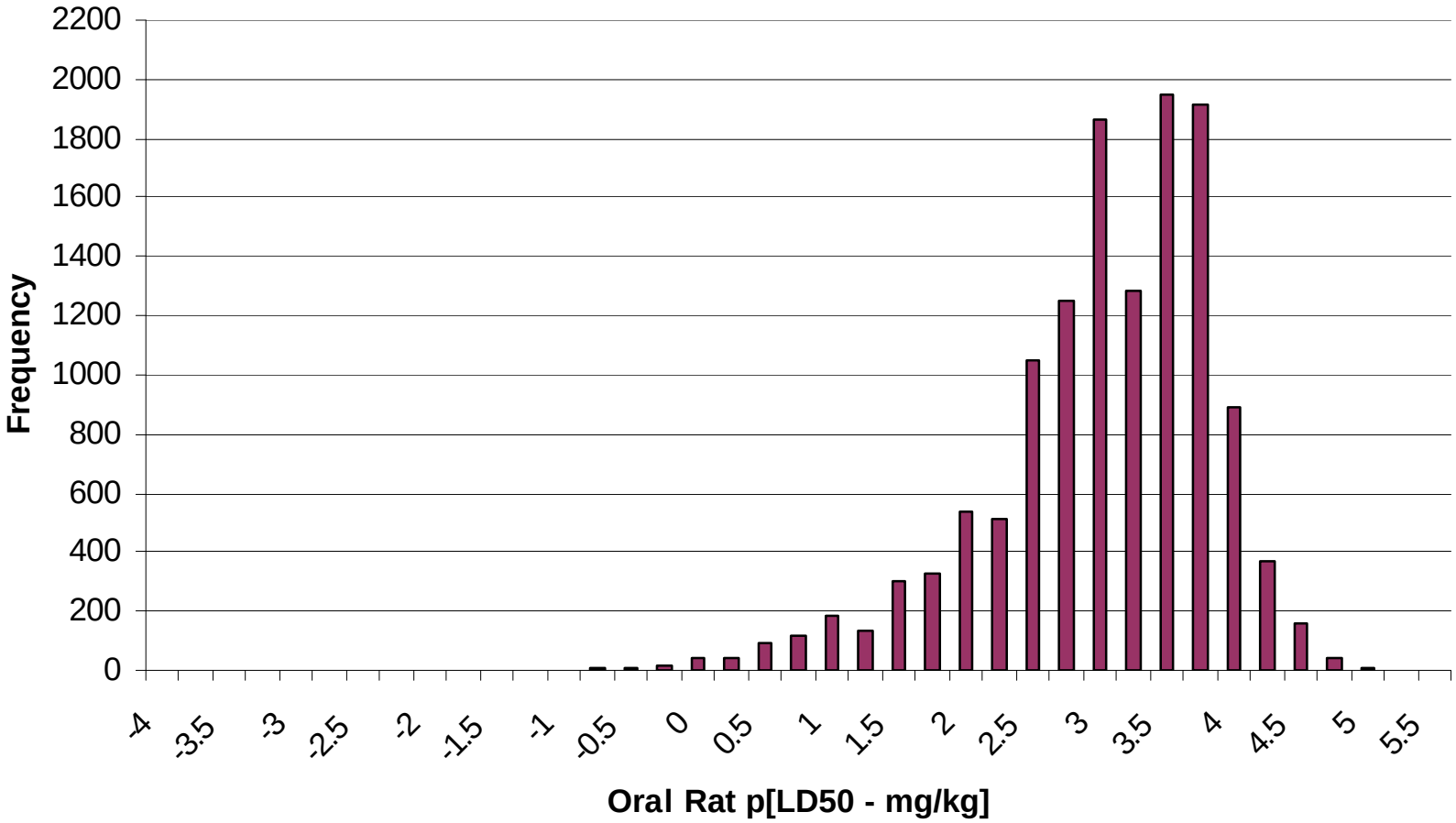


Notes: LD50 data points in DrugSet which have company source data in CMC 2002.1, legacy company-source names (e.g. Novartis/Sandoz/Ciba-Geigy, ICI/Zeneca, Wyeth/Ayerst, SmithKline/Glaxo/Wellcome, etc.) notwithstanding. Other_8, Other_7 represent companies with 8, 7, etc. data points. Over 300 source-company entries have 1 compound-LD50 data-point.

Reference Set Distribution

Oral Rat log(LD50) – pLD50

Range: [-3.85, 5.27] Avg: 2.92 SDev: 0.85

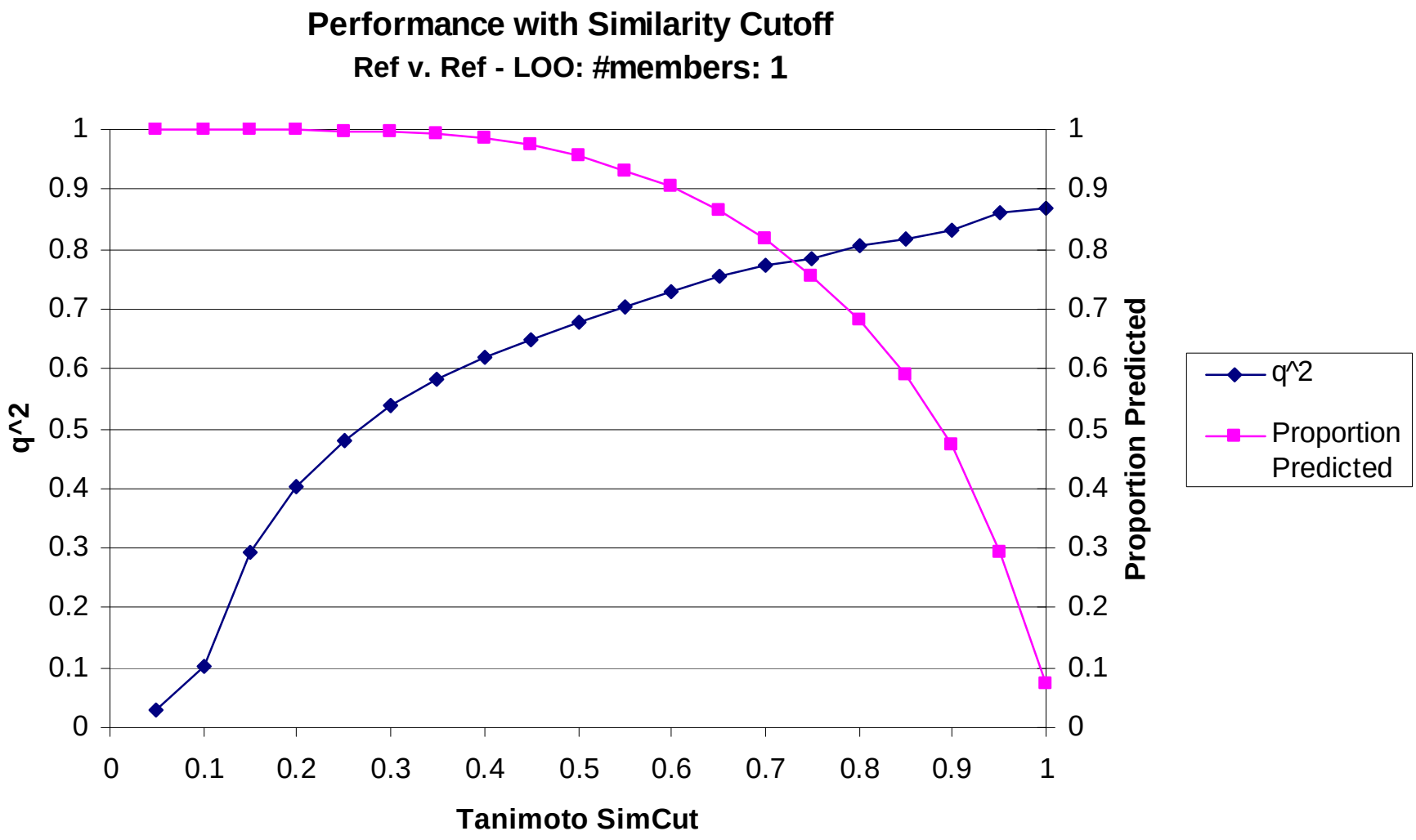


Did you know?

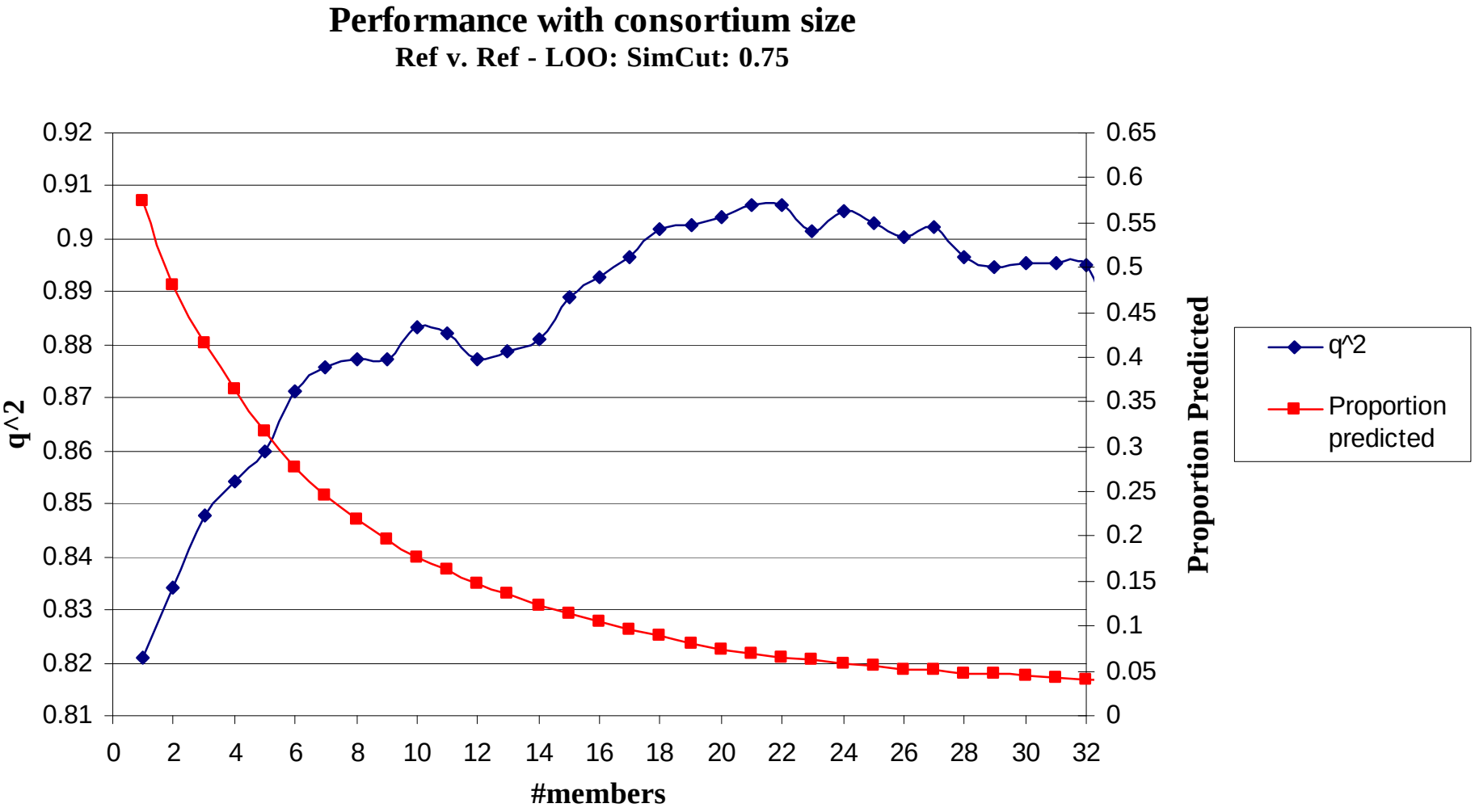
Compound	$LD50_{(oral/rat)}$	Reference
H2O	>90 mL/kg	Food Research 21,348,1956
NaCl	3000 mg/kg	Toxicology and Applied Pharmacology 20,57,1971
Sucrose	29700 mg/kg	Toxicology and Applied Pharmacology 7,609,1965
Caffeine	192 mg/kg	Journal of New Drugs 5,252,1965

Note: For a 170 Lb human (77.1 Kg):
LD50 Water: 6.9 L -> 234.7 fl-oz -> 29 x 8oz cups!
LD50 Caffeine: 14.8 g -> 110 strong expresso shots! A cup of coffee contains between 60-135 mg caffeine
LD50 Sucrose: 2289.9 g ->10x8oz cups of sugar!

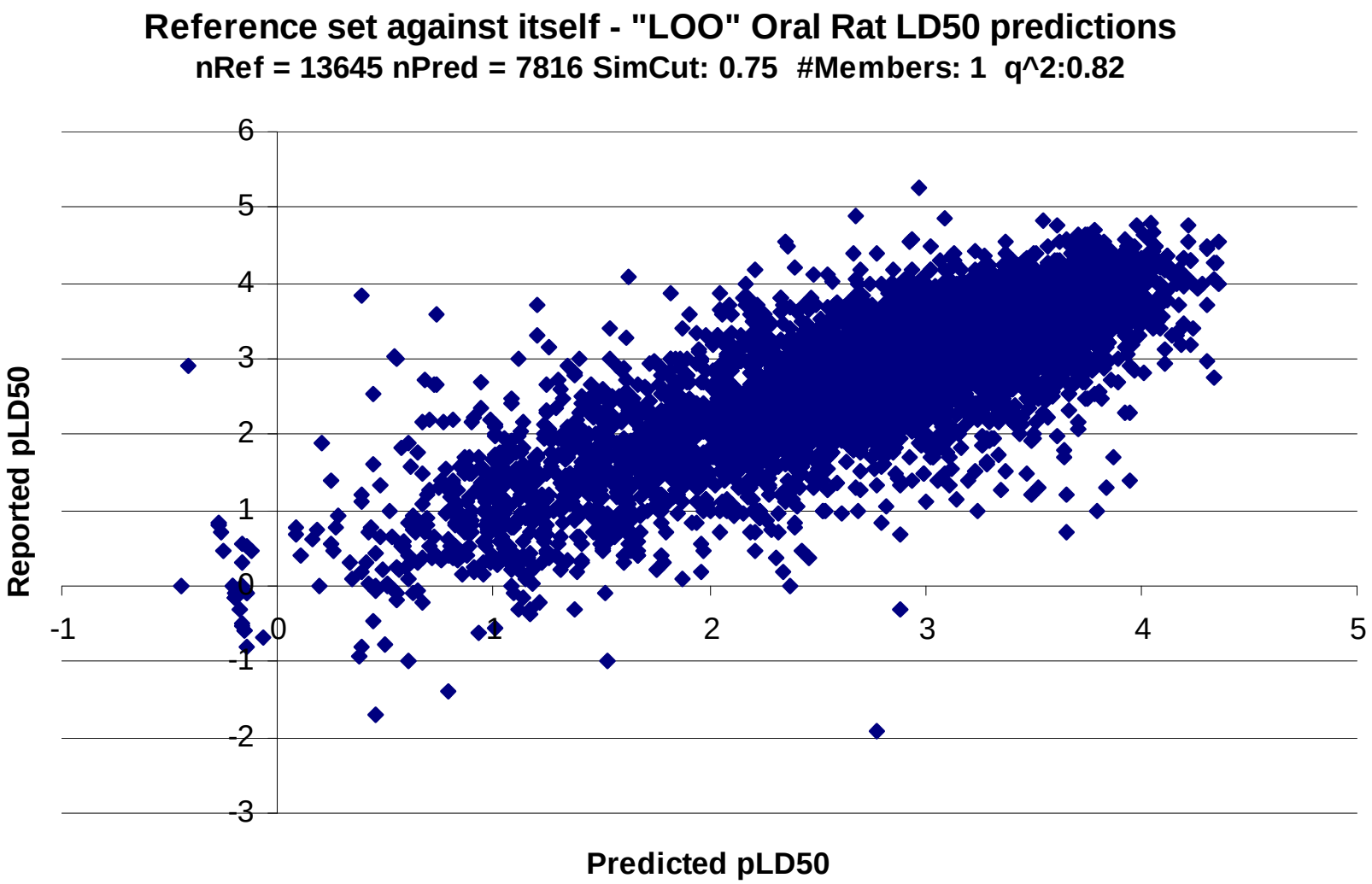
What is the Optimal Similarity Cutoff?



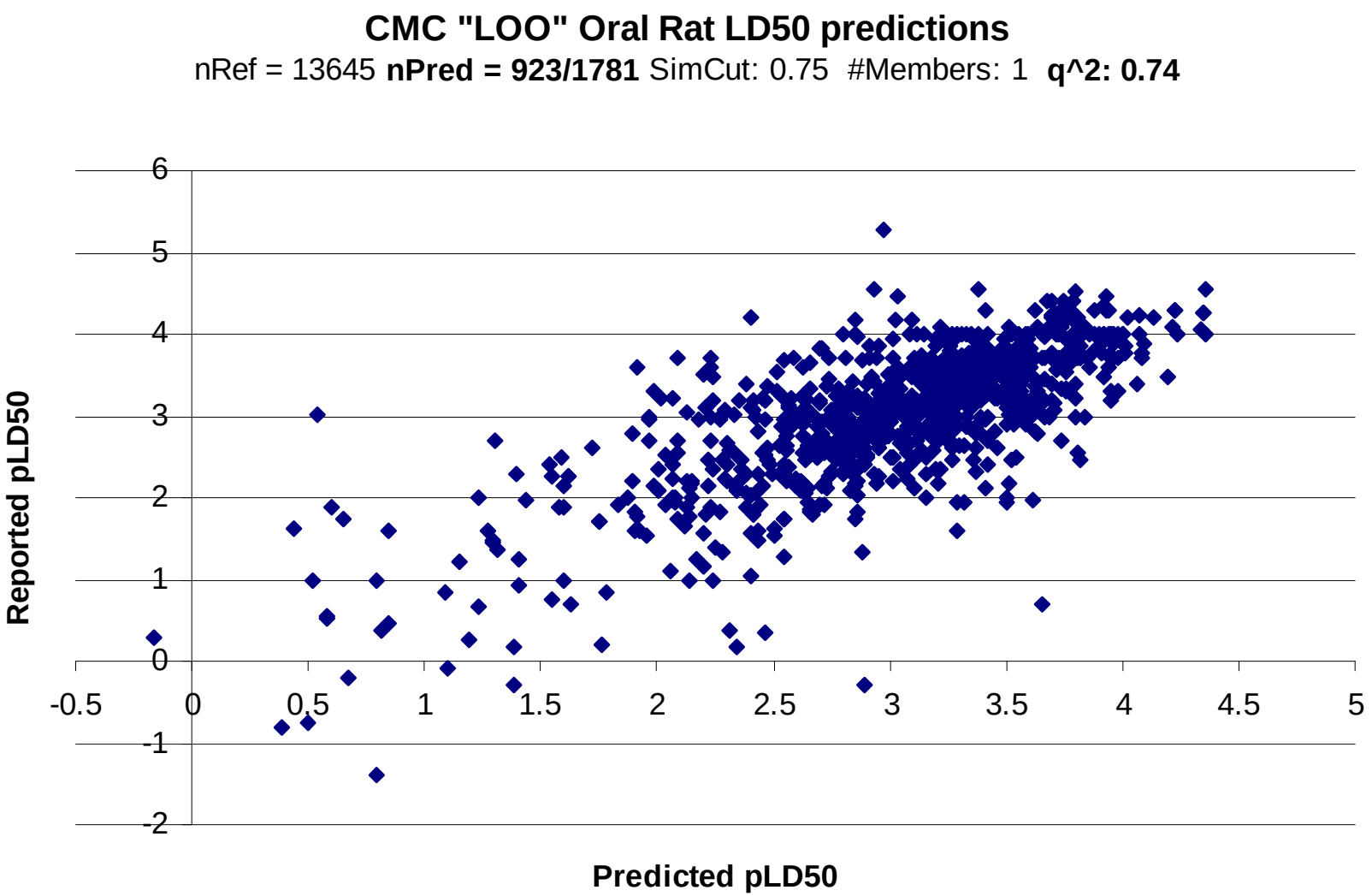
What is the Optimal Consortium Size?



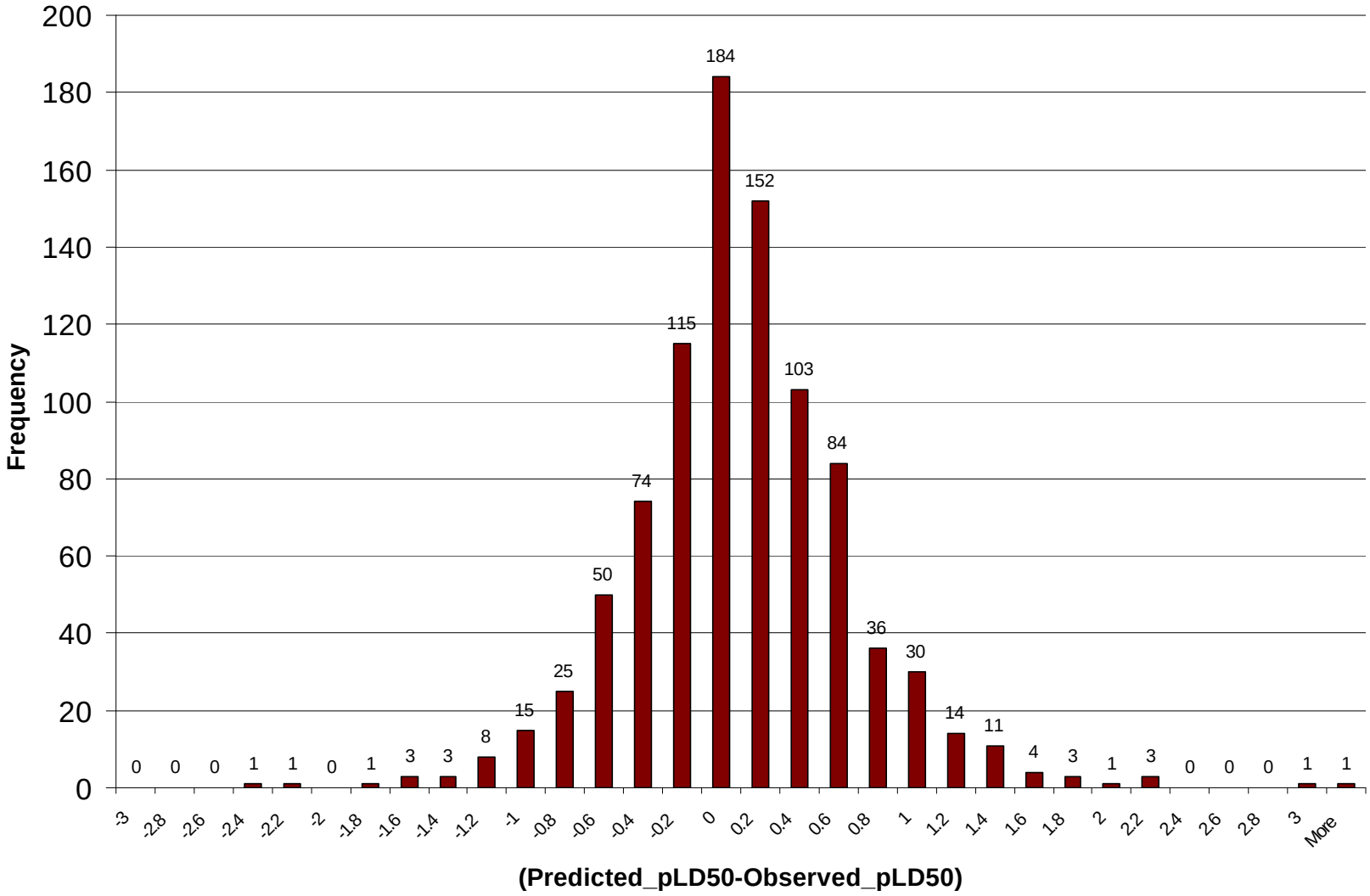
Evaluating RefSim – A Leave-One-Out Simulation



Predicting Drug Toxicity – “DrugSet”



DrugSet Prediction Errors



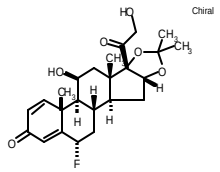
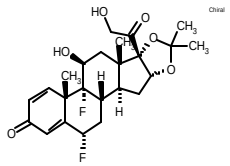
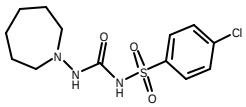
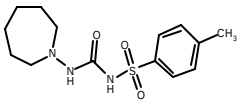
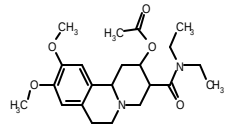
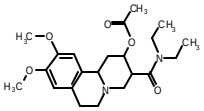
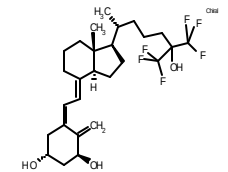
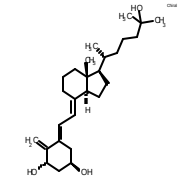
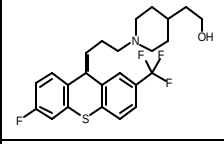
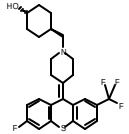
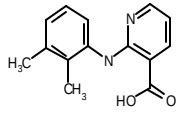
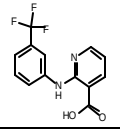
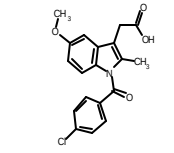
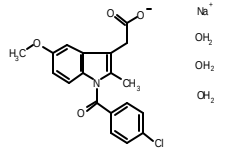
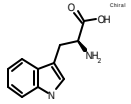
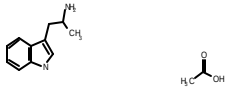
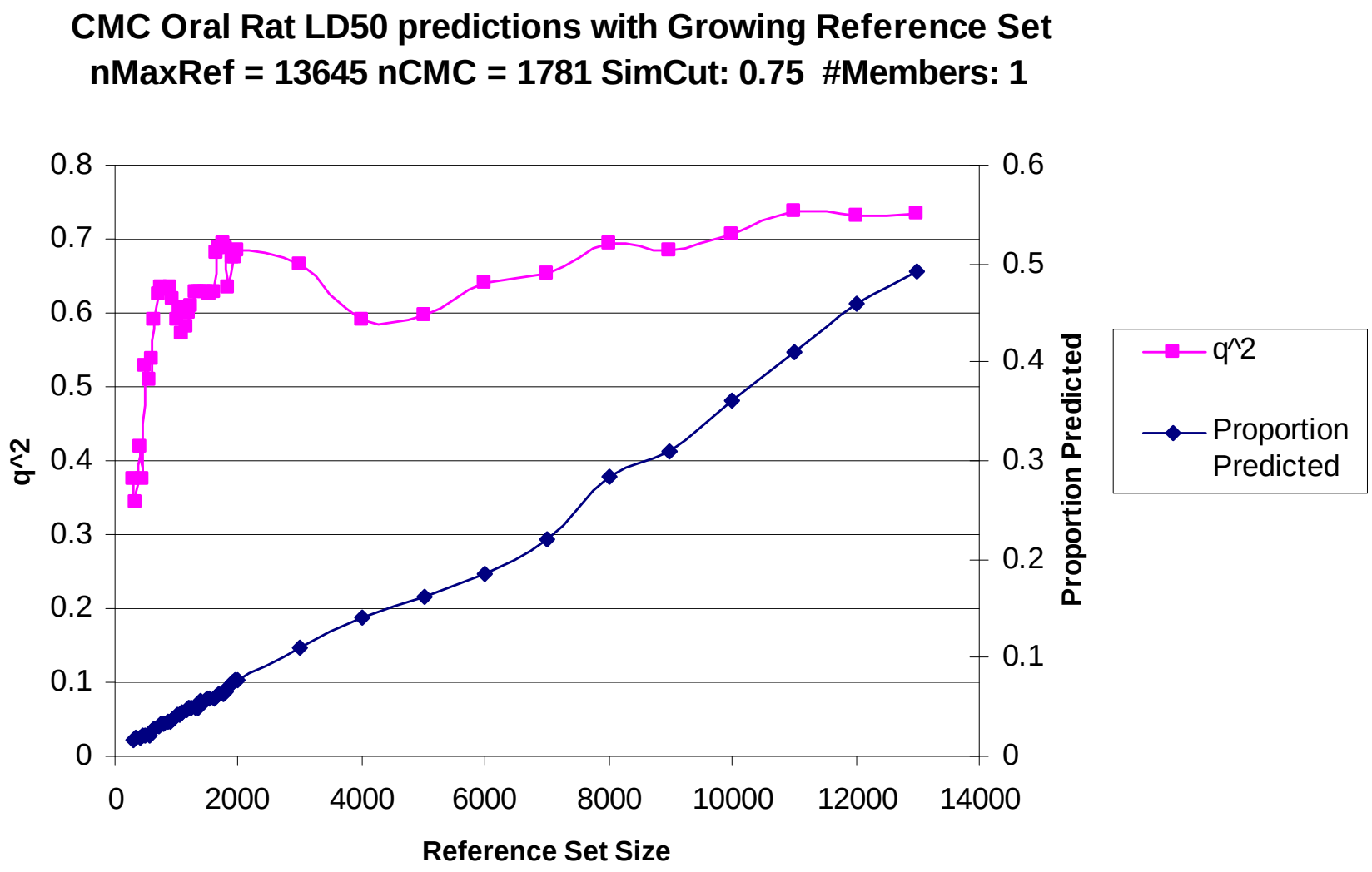
DrugSet Molecule Name	DrugSet Molecule (DSM)	DSM Observed LD50	DSM Reference	DSM Activity Class (CMC)	RefSet Most Similar Molecule (RSM)	RSM Observed LD50	RSM Reference	DSM Prediction Error: (Pred-Obs)	Num Ref Sim Mol	Tanimoto DSM&RSM
FLUNISOLIDE		>500 ug/kg	Gekkan Yakuji 26,501,1984	Glucocorticoid		>4 gm/kg	Drugs in Japan (Ethical Drugs) 6,694,1982	3.19	4	0.94
GLYPINAMIDE		>5 mg/kg	Patent, French Medicament Document #1087M	Antidiabetic		>5 gm/kg	Drugs in Japan (Ethical Drugs) 6,511,1982	2.96	2	0.87
BENZQUINAMIDE		1050 mg/kg	Psychotropic Drugs and Related Compounds - ,208,1972	Antiemetic		990 mg/kg	Toxicology and Applied Pharmacology 18,185,1971	-2.48	2	0.99
FLOCACITRIOL		41700 ng/kg	Kiso to Rinsho 30,2695,1996	Ca regulator		620 ug/kg	Patent, Japanese Kokai Tokyo Koho #94-247858	2.17	7	0.86
PIFLUTIXOLE		1500 ug/kg	Patent, United States Document #4309429	Neuroleptic		>60 mg/kg	Patent, United States Document #4309429	2.16	2	0.81
NIXYLIC ACID		2300 ug/kg	Therapie 22,157,1967	Antiinflammatory		250 mg/kg	Journal of Medicinal Chemistry 16,780,1973	2.10	2	0.94
INDOMETHACIN		2420 ug/kg	Arzneimittel-Forschung 25,1526,1975	Antiinflammatory		21 mg/kg	Gekkan Yakuji 37,952,1995	1.93	24	0.86
TRYPTOPHAN		>16 gm/kg	Iyakuhin Kenkyu 11,635,1980	Antidepressant		22 mg/kg	Toxicology and Applied Pharmacology 4,547,1962	-1.81	4	0.99

Table 2. DrugSet prediction outliers. The reported LD50's for the query DrugSet molecules (DSMs) and for their respective maximally similar molecules in the reference set (RSMs) show significant differences despite high Tanimoto similarity between DSM and RSM. This helps explain the fairly large absolute, signed DSM prediction error: pLD50(pred)-pLD50(obs).

How to Enhance Toxicity Predictions?



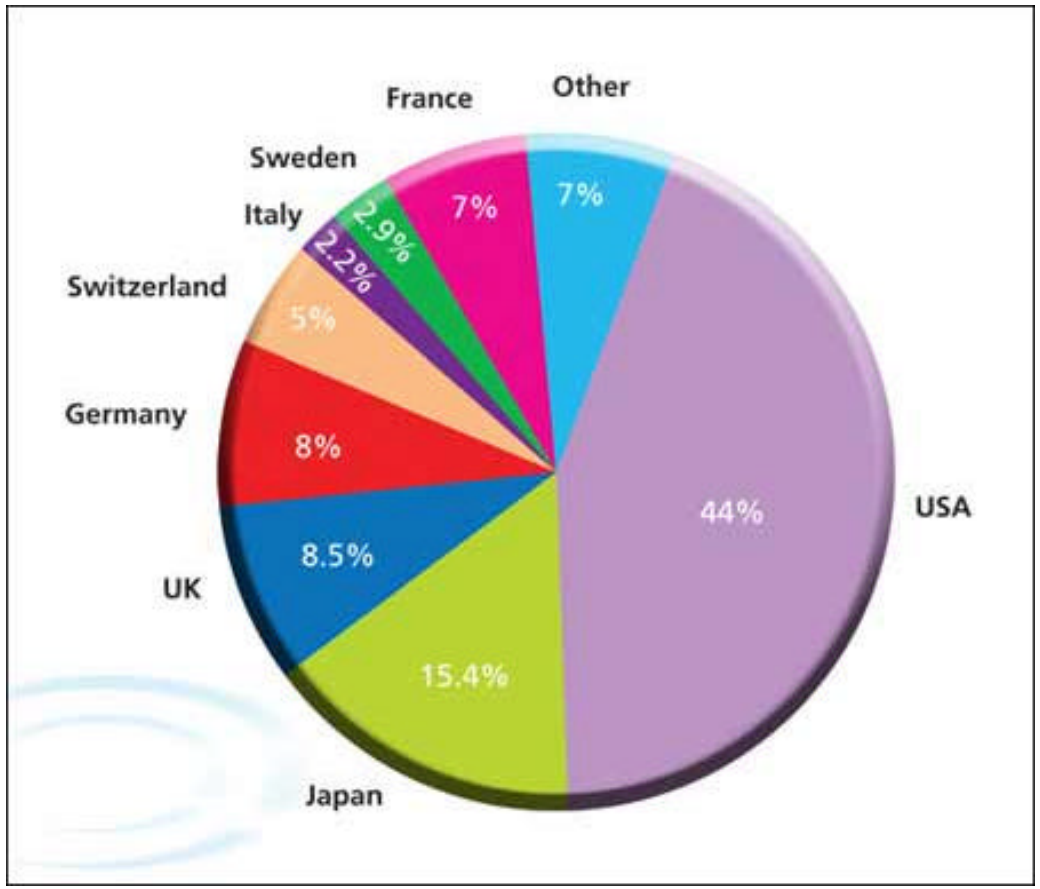
Summary & Conclusions

- Reasonably accurate and robust toxicity predictions can be achieved with a reference similarity approach
- Small to moderate consortia per compound-class may suffice to build a well rounded reference set
- An increase in reference set size is likely to improve both the quality and quantity of toxicity predictions
- Significant opportunity exists to enhance the “Starter” reference set, which only scratches the surface…
- Prediction technique can be readily incorporated into a high-throughput in silico eScreening strategy (e.g. compound prioritization, filtration, etc.)

There Can Be Safety In Numbers…

Supplementary Slides

Global Pharmaceutical R&D Expenditure by Country



Source:CMR International - http://www.cmr.org/pdfs/springnews2002.pdf

ADRs are a major cause of Death

Adverse Drug Reactions may be the fourth to sixth leading cause of death

Deaths Per Year	Cause
106,000	**Non-error, negative effects of drugs[1]**
80,000	**Infections in hospitals[4]**
45,000	**Other errors in hospitals[4]**
12,000	**Unnecessary surgery[2]**
7,000	**Medication errors in hospitals[3]**
250,000	**Total deaths per year from iatrogenic* causes**

* The term *iatrogenic* is defined as "induced in a patient by a physician's activity, manner, or therapy. Used especially to pertain to a complication of treatment."

1. Kohn L, ed., Corrigan J, ed., Donaldson M, ed. To Err Is Human: Building a Safer Health System. Washington, DC: National Academy Press, 1999.
2. Leape L. Unnecessary surgery. Annual Rev. Public Health. 1992; 13:363-383.
3. Phillips D, Christenfeld N, Glynn L. Increase in U.S. medication-error deaths between 1983 and 1993. Lancet, 1998; 351:643-644.
4.Lazarou J, Pomeranz B, Corey P. Incidence of adverse drug reactions in hospitalized patients. JAMA. 1998; 279:1200-1205.

See: http://www.chelationtherapyonline.com/articles/p62.htm

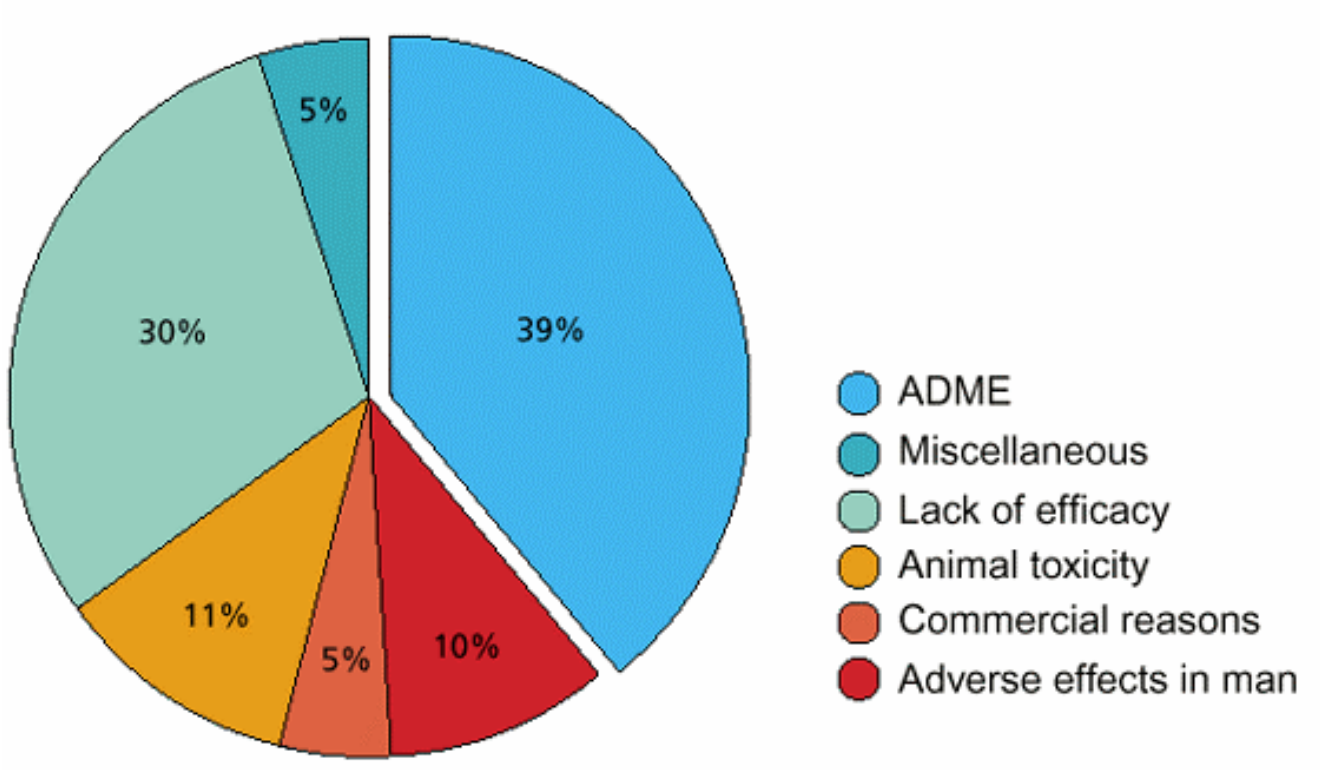
Attrition: The Reasons Have Been Cited…

Reasons for Failure	Reasons for Slow Down
Poor biopharmaceutical properties, 41%	Poor biopharmaceutical properties
Lack of efficacy, 31%	Low potency
Toxicity, 22%	Ambiguous toxicity finding
Market reasons, 6%	Inherently time-intensive target indication
	Synthetic complexity

• Prentis, R. A., Lis, Y. & Walker, S. R. Pharmaceutical innovation by the seven UK-owned pharmaceutical companies (1964-1985). *Br J Clin Pharmacol* 25, 387-96. (1988)

• Lipper, R.A. How can we optimize selection of drug development candidates from many compounds at the discovery stage? *Modern Drug Discovery,* 1999, 2 *(1)*, 55-60.

Causes of attrition during drug development



Source: Kennedy, T. *Drug Discovery Today* **2, 436–444, 1997.**