Challenges and needs to use existing data in drug development

Francois Pognan, PhD,
Discovery Investigative Safety
Basel, Switzerland

BioMed 21: A Human Pathways Approach to Disease Research
26-27 June 2017 – Bethesda, MD, USA

NOVARTIS

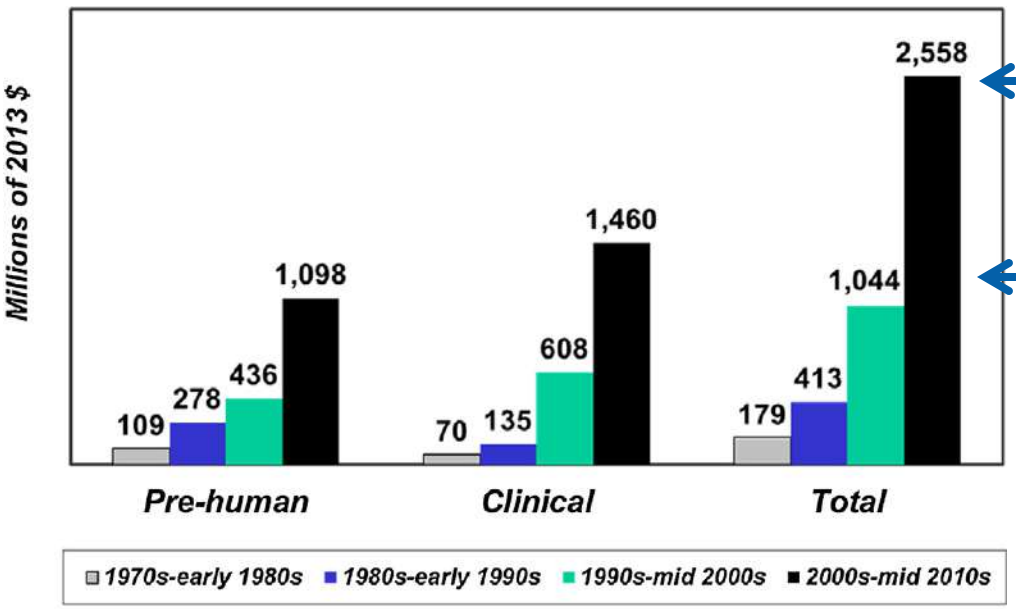
Challenges and needs to use existing data in drug development

Thoughts for today's debate

- Drug development cost and challenges
- Data waste
- Data mess with order
- New order needed for data recycling
- What is being done and will be done
- Examples of new knowledge out of old data
- Discussion: precompetitive data sharing

"Give, and you will receive"

Challenges and needs to use existing data in drug development



Sources: 1970s-early 1980s, Hansen (1979); 1980s-early 1990s, DiMasi et al. (1991); 1990s-mid 2000s, DiMasi et al. (2003); 2000s-mid 2010s, Current Study

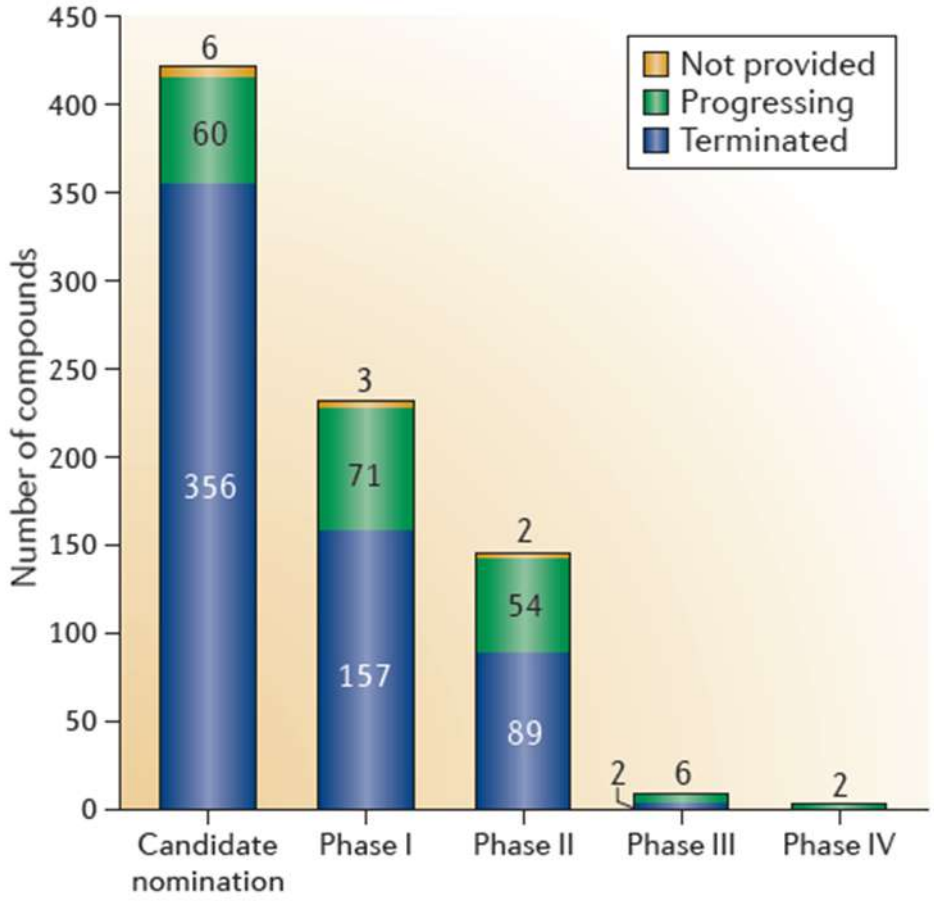
DiMasi, J. A., *et Al.. J Health Econ* 47, 20-3

Cost
of
drug development

One of the main reasons for the cancellation of the last three Apollo flights was the cost. In the time frame from 1969 to 1971 the cost of launching a Saturn V Apollo mission was US $185 to $189 million, (**equivalent to $1.22 billion in 2015**)

https://en.wikipedia.org/wiki/Saturn_V#Cost

Challenges and needs to use existing data in drug development



Waring, M. J. *et al.*. *Nature Reviews Drug Discovery* **14**, 475-486

An enormous amount of data has been generated, most of it for "nothing"

- Preclinically
- Clinically

*****

Or are barely used

- Pharmacovigilance
- National health care systems

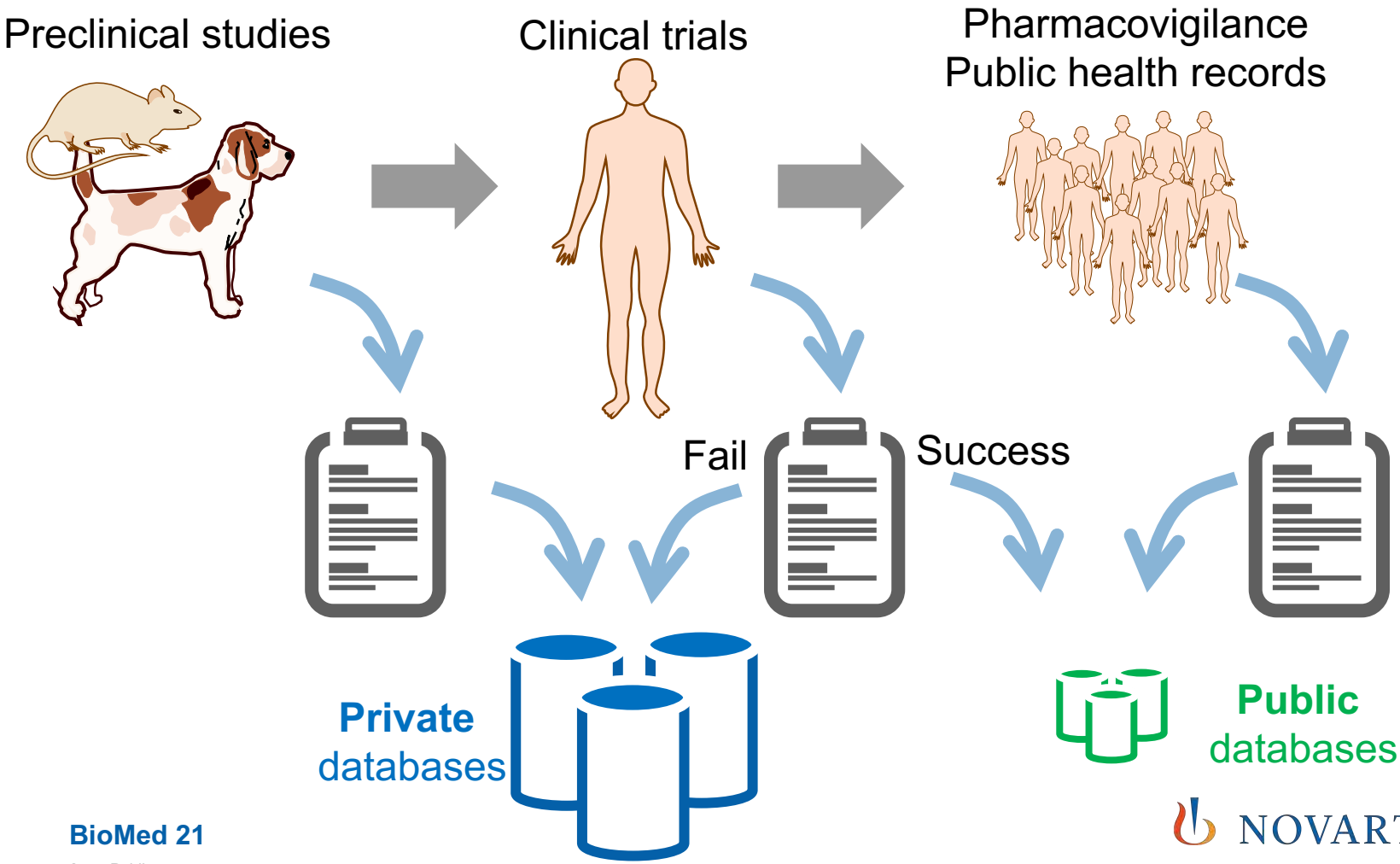
NOVARTIS

Challenges and needs to use existing data in drug development

- Collectively, the pharma industry has generated a gigantic amount of preclinical and clinical data
- These data are conscientiously kept in archives for retrieval when required by HAs
- Data are archives in a way that allows almost immediate retrieval, but do not permit cross or meta-analysis
- Pharma DBs contain data for compounds successfully developed to drug, and many logs more of unsuccessful molecules
- Data therein are mostly considered as competitive advantage
- These data are mostly sleeping despite containing an unmined wealth of data that could be turned into knowledge

Challenges and needs to use existing data in drug development

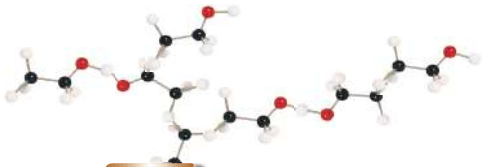
What data are where ?



Challenges and needs to use existing data in drug development

Mostly:
Rat
Dog
monkey
But also:
Mouse
Rabbit
Mini-pig...

Variety of preclinical data



Clinical Trials

Discovery | Phase 1 | Phase 2 | Phase 3 | Launch

IND | Full Development | NDA

Target liabilities
Efficacy models
Insilico & invitro safety screens

Dose ranging studies – 1- 2 wks.
Single dose toxicity

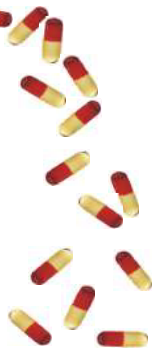
Regulatory tox package for FIH

Regulatory tox package for FIH
2 wk-1 month studies

Chronic tox
Fertility
Teratogenicity
Developmental tox
Abuse liability

Subchronic / chronic toxicity
3, 6, or 26/39 month studies

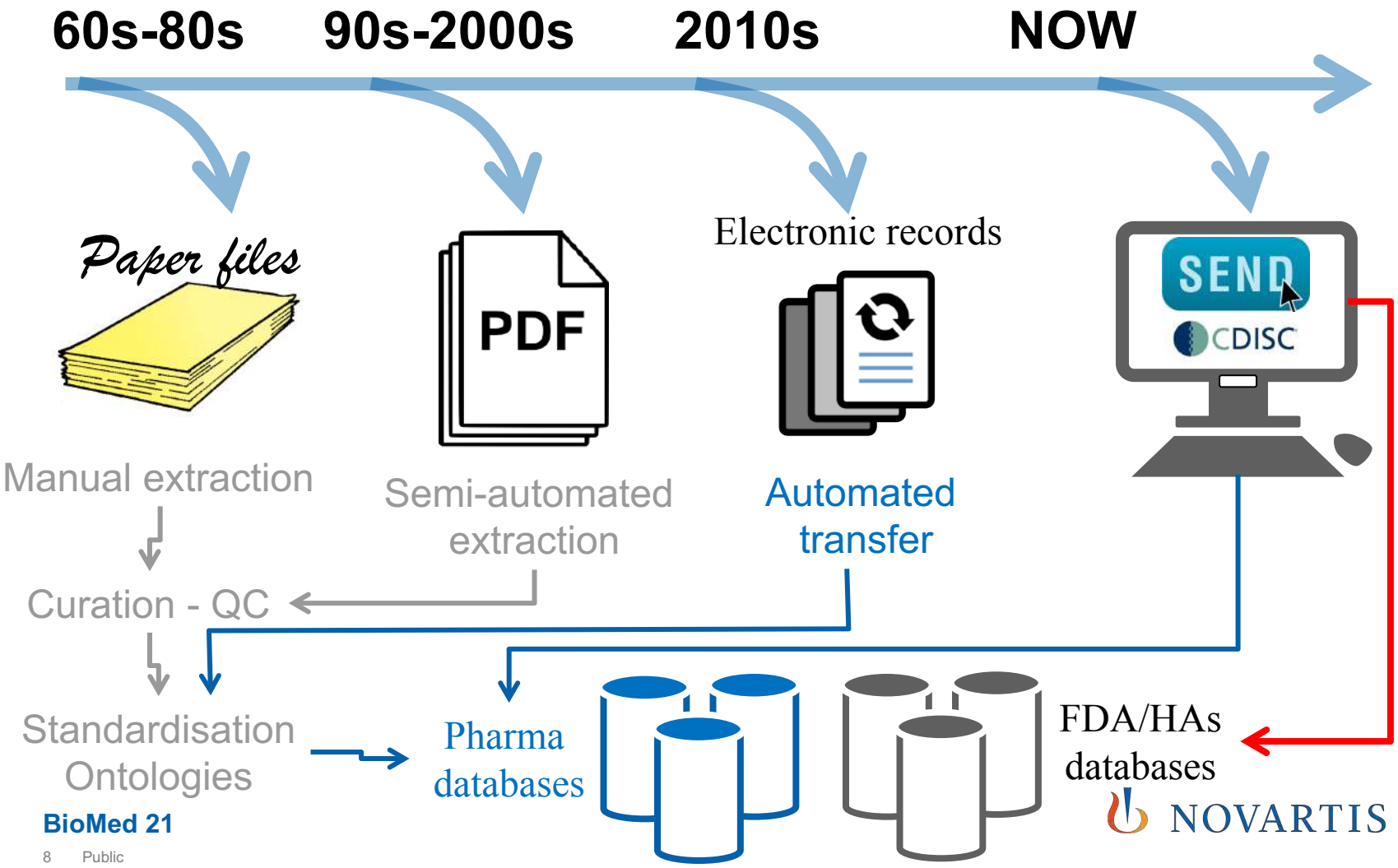
Carcinogenicity
PPND



NOVARTIS

Challenges and needs to use existing data in drug development

Heterogeneity of preclinical data supports



Challenges and needs to use existing data in drug development

- The eTOX initiative under the aegis of the European Union (IMI – Innovative Medicine Initiative)

Project vision: to develop innovative strategies and novel software tools to better predict the potential side-effects of new drug candidates on the basis of integrative approaches.

Data mining

In silico prediction

Challenges and needs to use existing data in drug development

From data to knowledge to wisdom

EFPIA data

Organised eTOX db

Challenges and needs to use existing data in drug development

From data to knowledge to wisdom

Organised eTOX db

Enabled ontology classifications - knowledge

The Art of Cleanup by Ursus Wehrli
https://www.brainpickings.org/2013/03/28/the-art-of-cleanup-ursus-wehrli/

Meta-analysis,
extra knowledge : wisdom

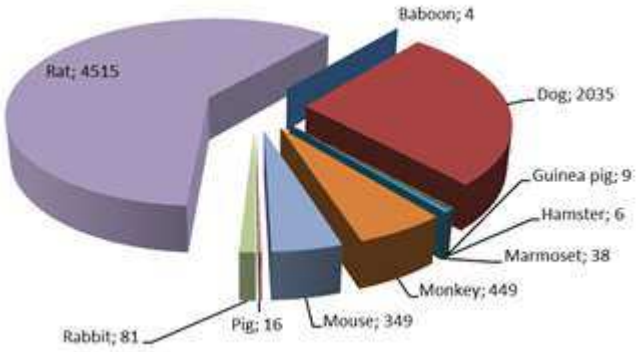
NOVARTIS

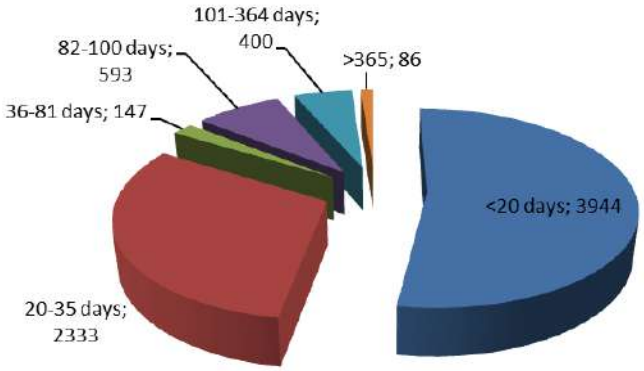
Challenges and needs to use existing data in drug development

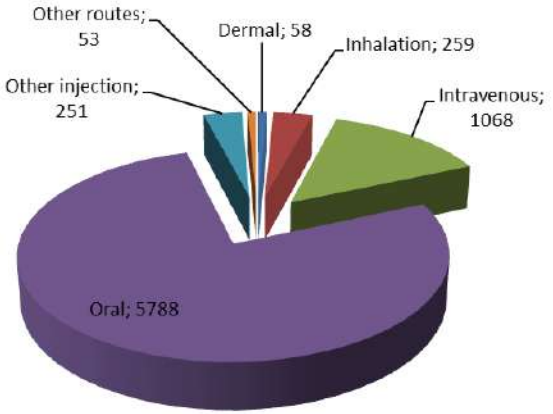
As of Jan 2017, the eTOX database is

Total number of structures: 2112 (476 confidential compounds)
Total number of studies: 8238
This represent about 20 millions entry lines

Private: company only and not shared







Challenges and needs to use existing data in drug development

ontologies | dealing with 80 000 terms

Current workflow allows efficient ontology maintenance and term mapping
More than **80 000 finding verbatim terms** out of **20 millions entry lines** mapped to the appropriate locations.

1. Ontobrowser collects verbatim terms from Vitic-Nexus when they are loaded (lines of data in scope).

2. Unknown verbatim terms are manually assigned as synonyms to **> 7 000 preferred terms**.

3. Pending synonyms are approved (or rejected) by a second curator.

4. The Vitic-Nexus database is loaded with preferred terms and synonyms during build.

Ontology/Codelist	Preferred Terms	Synonyms	Cross Refs*
anatomy	3141	18873	13%
effect level	8	49	0%
histopathology	1008	33310	38%
in-life observation	698	12294	0%
laboratory test name	1324	11301	99%
moa	465	0	0%
pk parameters	309	1868	95%
route of administration	123	411	100%
sex of participants	3	123	100%
species	27	32	93%
strain/substrain	84	1121	98%

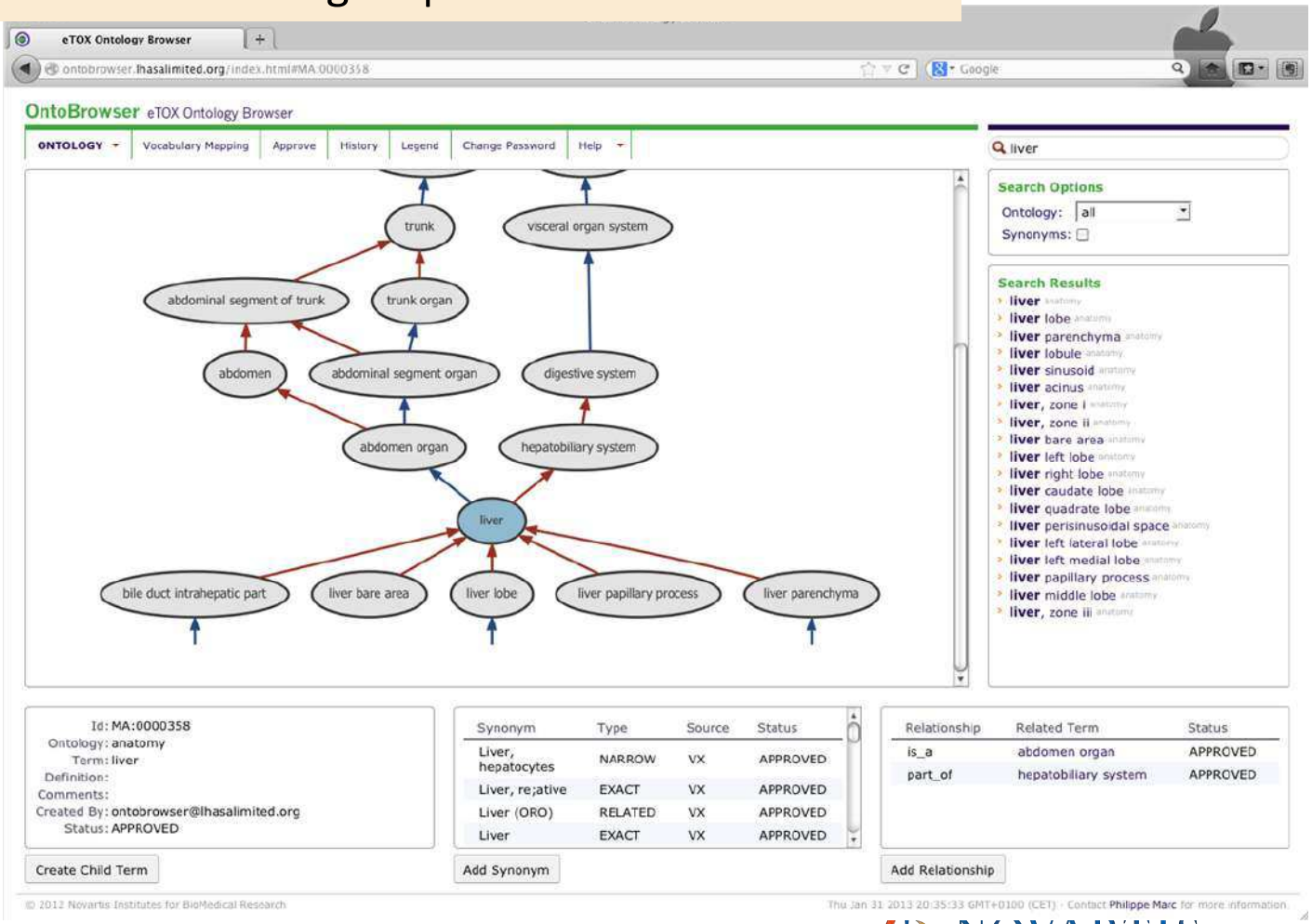
*Percentage of preferred terms with cross references public nomenclatures such as SEND/NCI, INHAND, MedDRA etc...

OntoBrowser http://opensource.nibr.com/projects/ontobrowser/ https://ontobrowser.lhasalimited.org/index.html#MC:0000301

Challenges and needs to use existing data in drug development

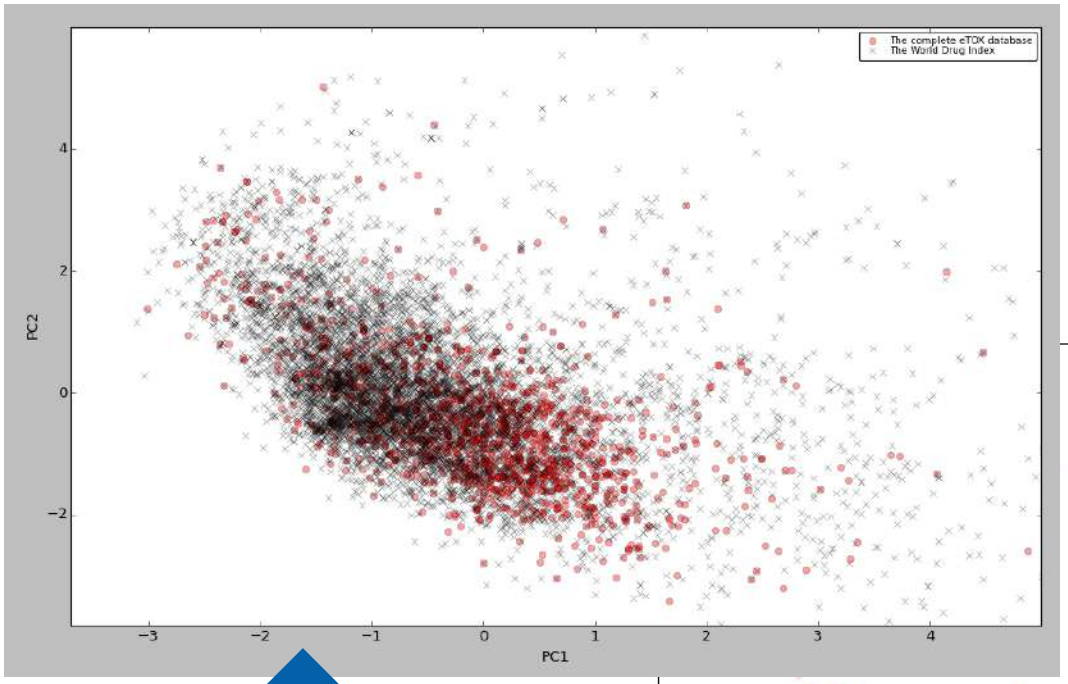
ontologies | what for ?

Collecting data is fine, but is a specific finding related to “liver”, hepato*, hepatic Kupffer, cholangiocyte, bile duct …???



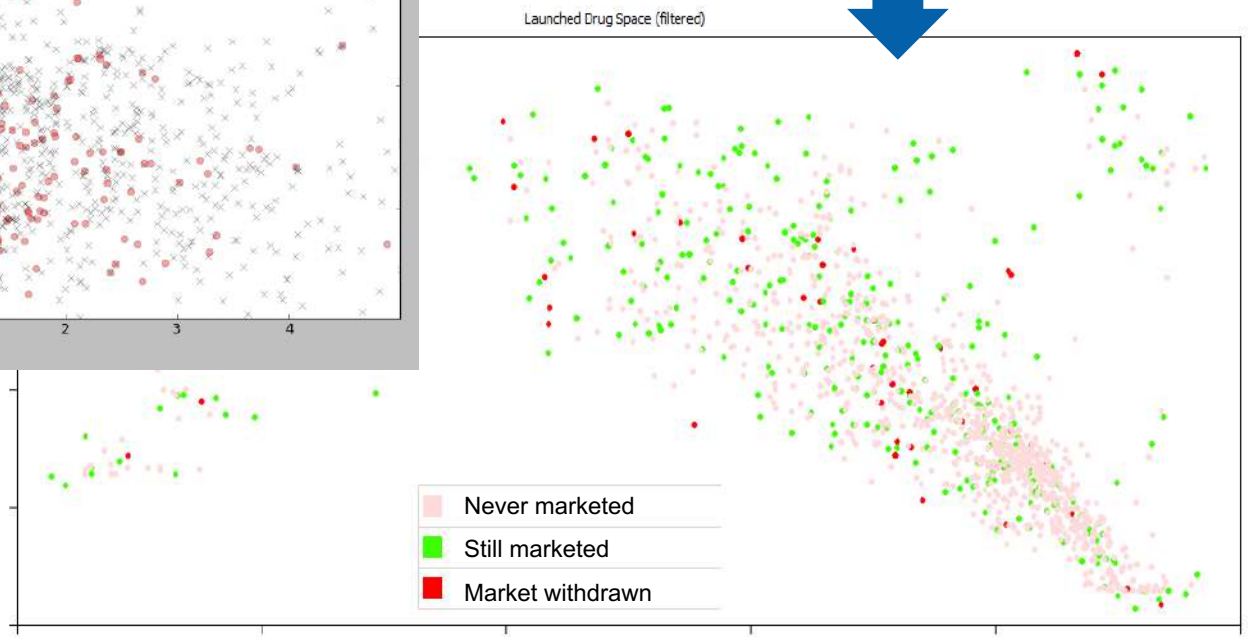
Challenges and needs to use existing data in drug development

eTOX chemical space and content



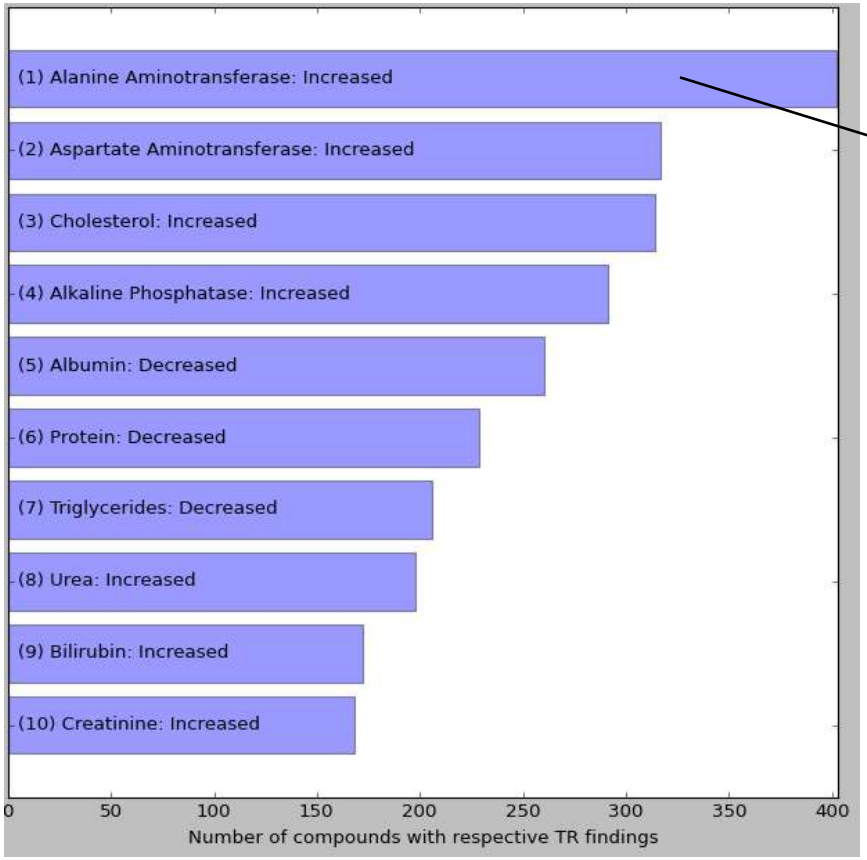
X grey crosses; WDI
O red dots; eTOX

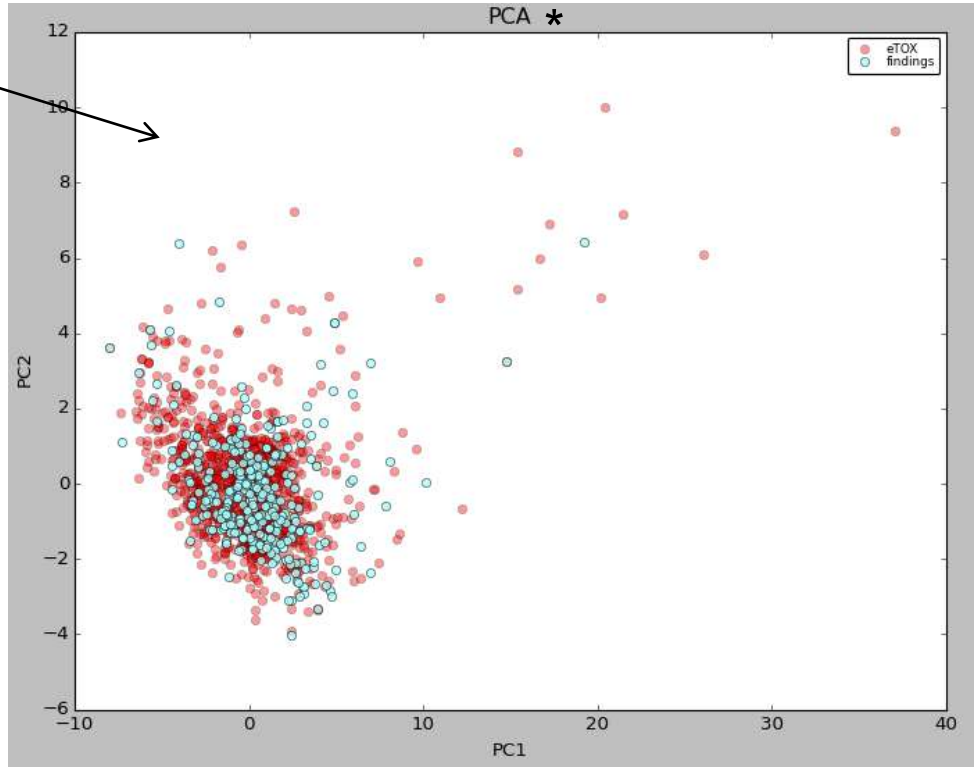
eTOX chemical space by PCA, showing the quantity of individual compounds, proportion of marketed drug still on market and those withdrawn for safety issues (cerivastatin, lumiracoxib, rosiglitazone, ximelagatran...).



Challenges and needs to use existing data in drug development

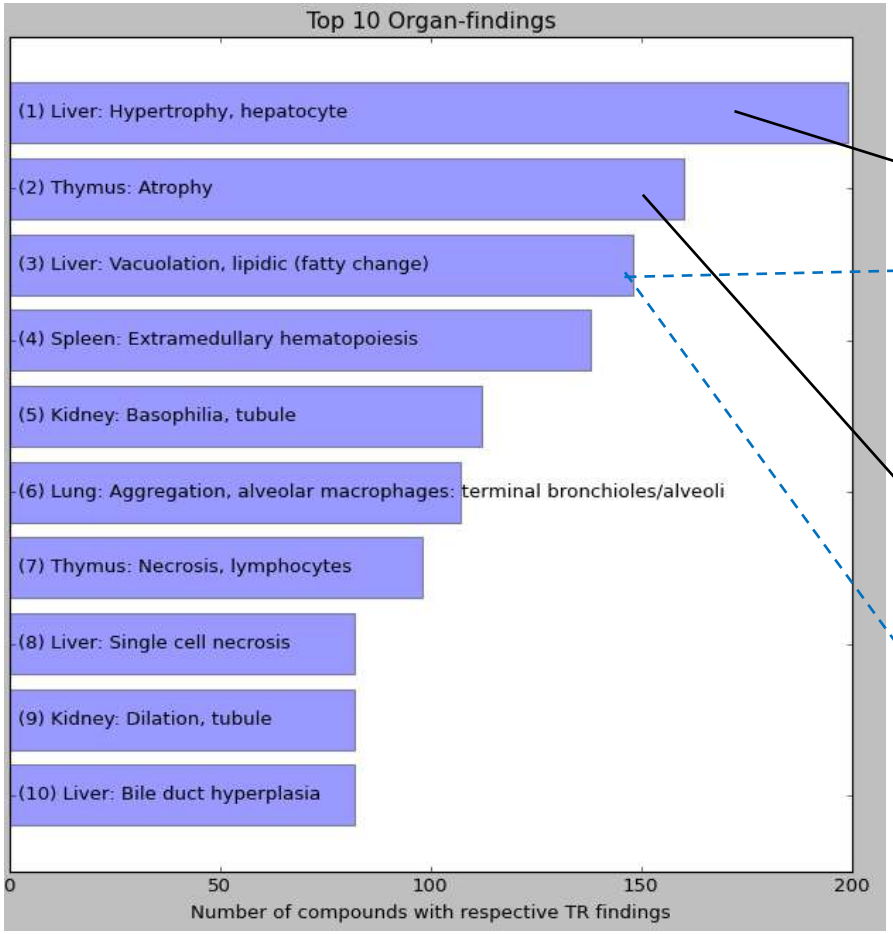
- Meta-analyses allow extraction of larger knowledge, not otherwise possible, eg:
  - Clin chem: Most common findings all species



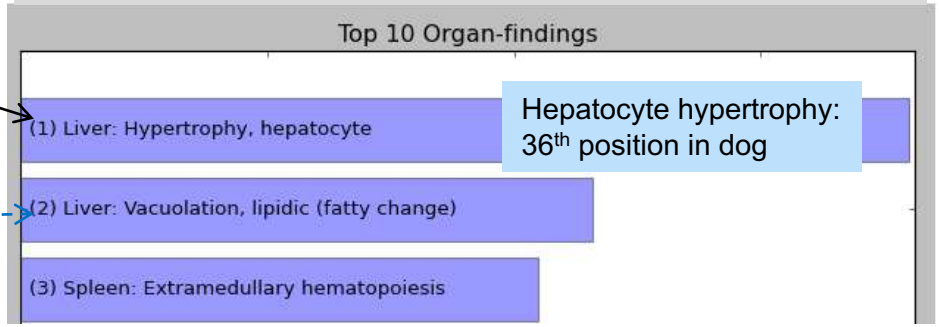


Challenges and needs to use existing data in drug development

Histopathological findings – All species

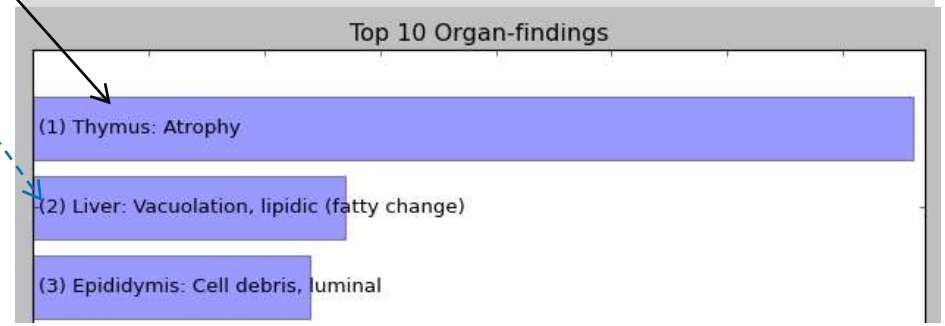


Rat – prone to hepatocyte hypertrophy (P450 induction)



Hepatocyte hypertrophy: 36th position in dog

Dog – prone to thymus atrophy (general health)



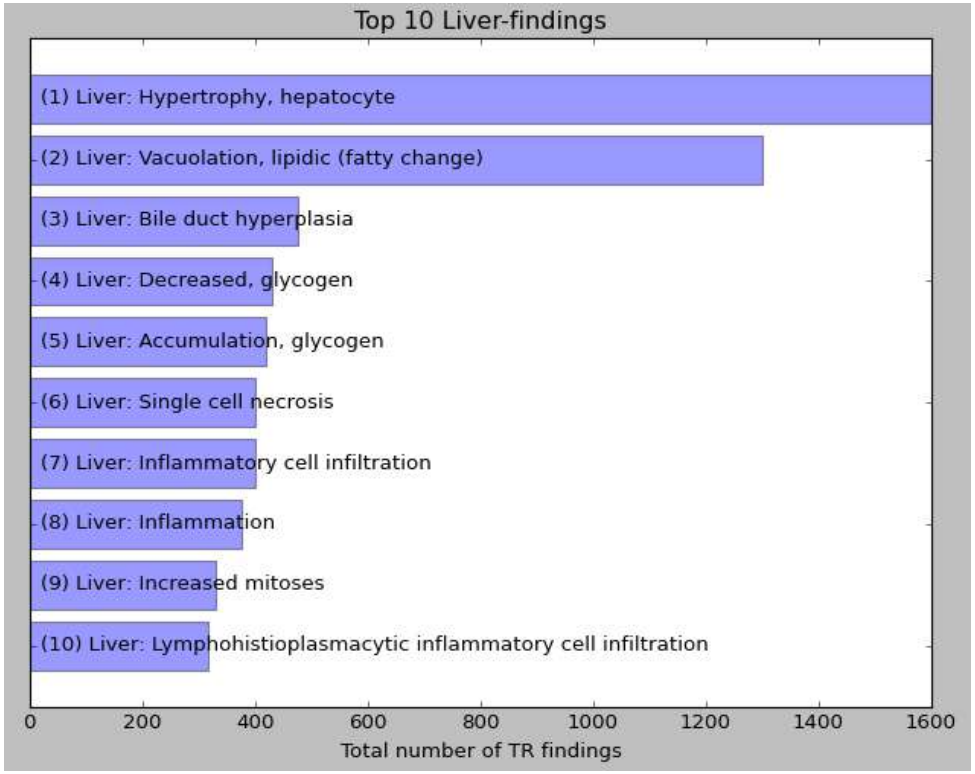
BioMed 21

NOVARTIS

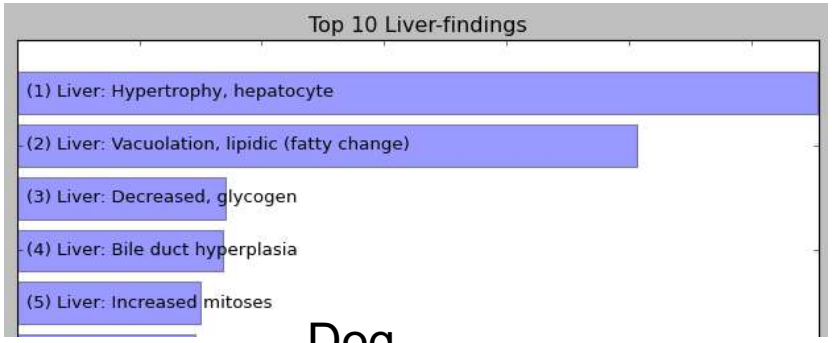
Challenges and needs to use existing data in drug development

Histopathological findings – liver only

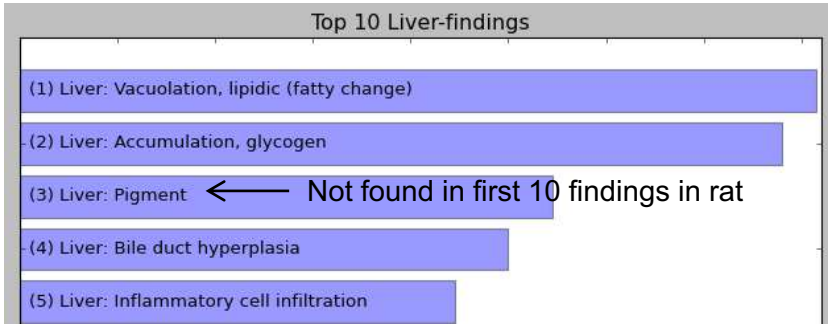
All species



Rat



Dog



NOVARTIS

Challenges and needs to use existing data in drug development

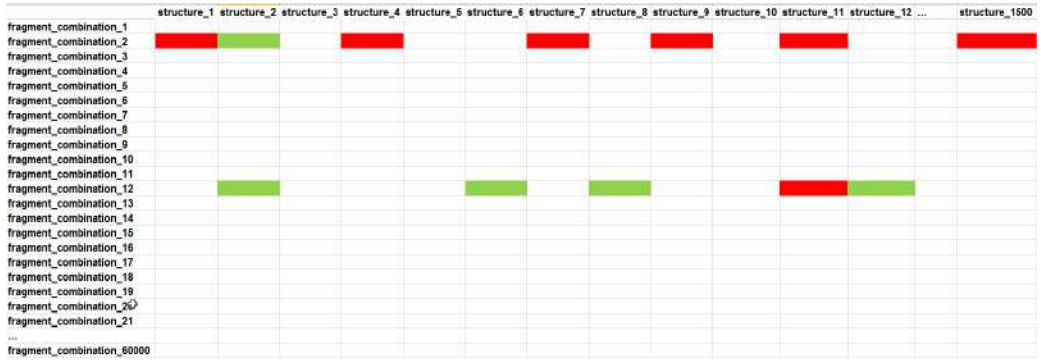
New structural alerts from the eTOX database (chemistry + toxicology)

Dataset

(structures flagged has positive or negative for a given tox. end-point)

Fragmentation

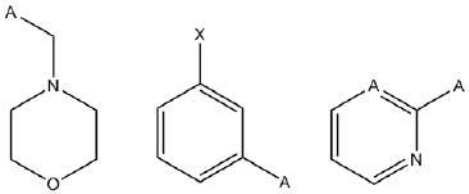
(generation of unique fragments from every dataset structures, using several fragmentation algorithms)



Matrix analysis

(Retrieval of structural alerts and scoring coming from the fragments that are the most frequently encounter in toxic structures)

e.g: fragment_combination_51263:



42 substructure-relationships with the dataset (40 with toxic and 2 with non toxic structures)

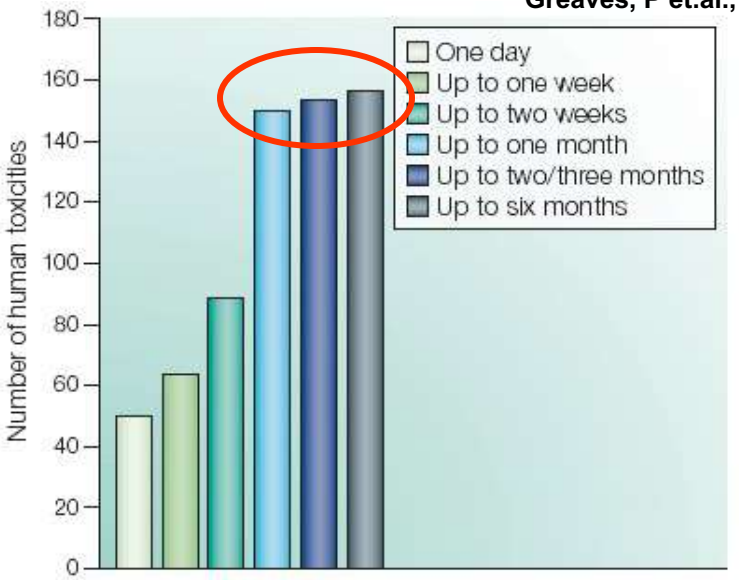
Matrix of substructure-relationship

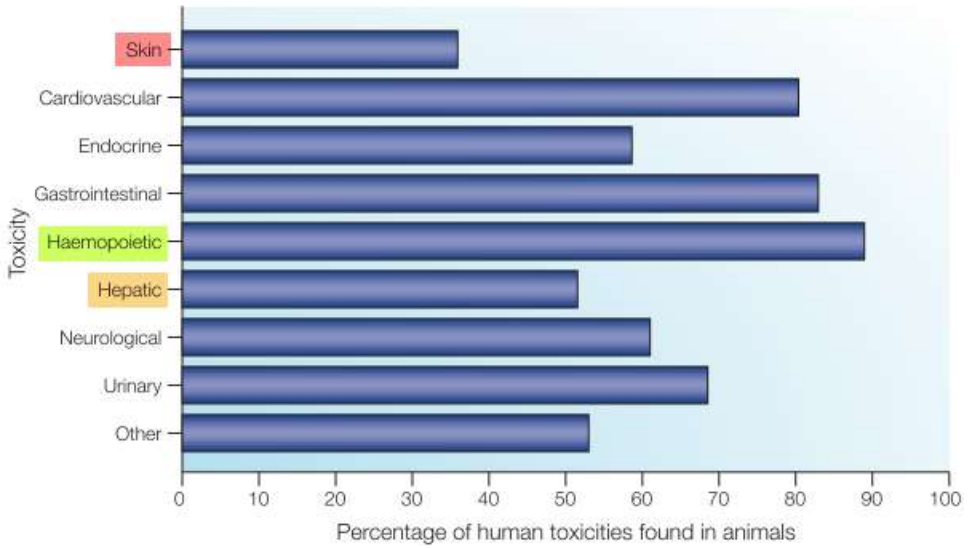
(example with 1.500 structures X 600.000 fragments where red shows substructure relationships with toxic structures and green shows substructure relationships with non-toxic structures)

Challenges and needs to use existing data in drug development

Animals studies predicting adverse effects in man

Greaves, P et.al., Nature Rev Drug Discovery 3 (2004) 226-236





With limited data:
94% of animal toxicities that correlate with human toxicities were detected in preclinical studies of ≤ 1 month duration.

If some parameters or studies are irrelevant to human, why running them, need better studies, better design, different approach
→ Impact on guidelines ?

Challenges and needs to use existing data in drug development

What is next after eTOX - eTRANSAFE (eTS)

- **Three objectives:**
  - exhaustive analysis of correspondence and validity of animal data for human safety
  - discovery of translational and reverse-translational biomarkers
  - predict animal toxicities
- **Four deliverables:**
  - preclinical data base with retrospective and prospective data (SEND) from multiple companies
  - mining and visualisation tools for cross-analysis with human data
  - in silico predictive platform (algorithms)
  - new translation and reverse-translation biomarkers
- **Impacts**
  - preclinical studies adapted to human outcome = increased safety, reduced attrition
  - 3Rs
  - preclinical knowledge management

Challenges and needs to use existing data in drug development

Outlook

- New knowledge extractable from old data
- Need to very large amount of data → consortia
- Need for quality and standard format (e.g., SEND)
- Need for large, comprehensive systematic and objective assessment of preclinical data for their validity of human safety prediction
- Impact on
  - 3Rs
  - Future design of preclinical studies
  - Better understanding of toxicology
  - Translational aspects of safety assessment
  - Perhaps on duration of development
  - Perhaps on overall cost of development

Challenges and needs to use existing data in drug development

Outlook

- Emergence and understanding of:
  - Concept of data sharing
  - “Pre-competitive” data
  - International consortia
  - The pharma industry has to collaborate in order to compete
  - Health Authorities, and academics have to join or even to push for initiation of collaborations

Data sharing is a gain, not a loss


Thank you
Q&A