

# Greater Data integration: Two Proposals for Discussion

Nawab Qizilbash MRCP MSc DPhil  
Oxon Epidemiology & Imperial  
College, London

# Disclaimer

- NQ is Head of OXON Epidemiology, a CRO that conducts epidemiological studies, patient registries, outcomes research, disproportionality analysis and meta-analysis for health sector companies

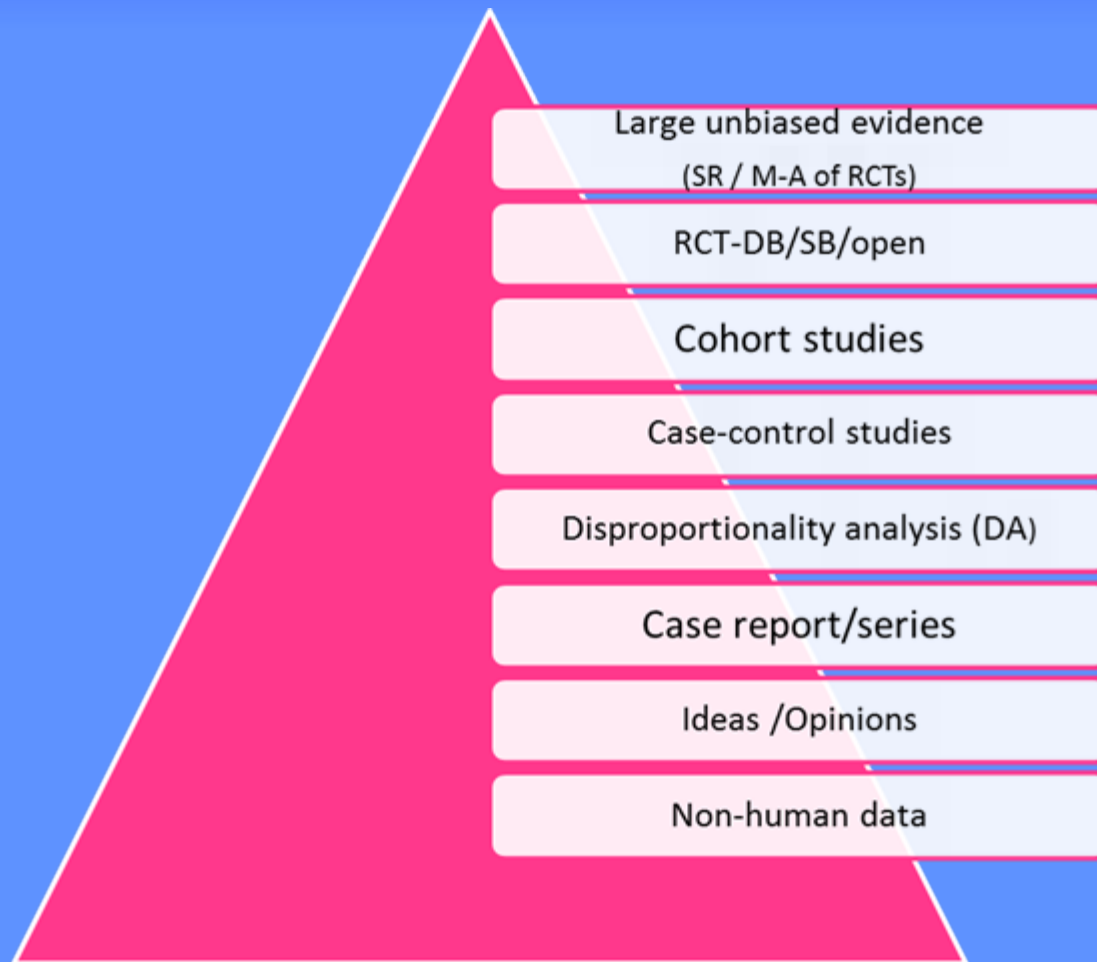
# Topics

- The need for meta-analysis(M-A), pooling & systematic reviews (SR) in safety
- Ideal scenarios
- The current situation
- Collaborative prospective pooling of epidemiological data (Miriam Sturkenboom)
- Desired developments?: Two proposals
- Discussion
- Next steps

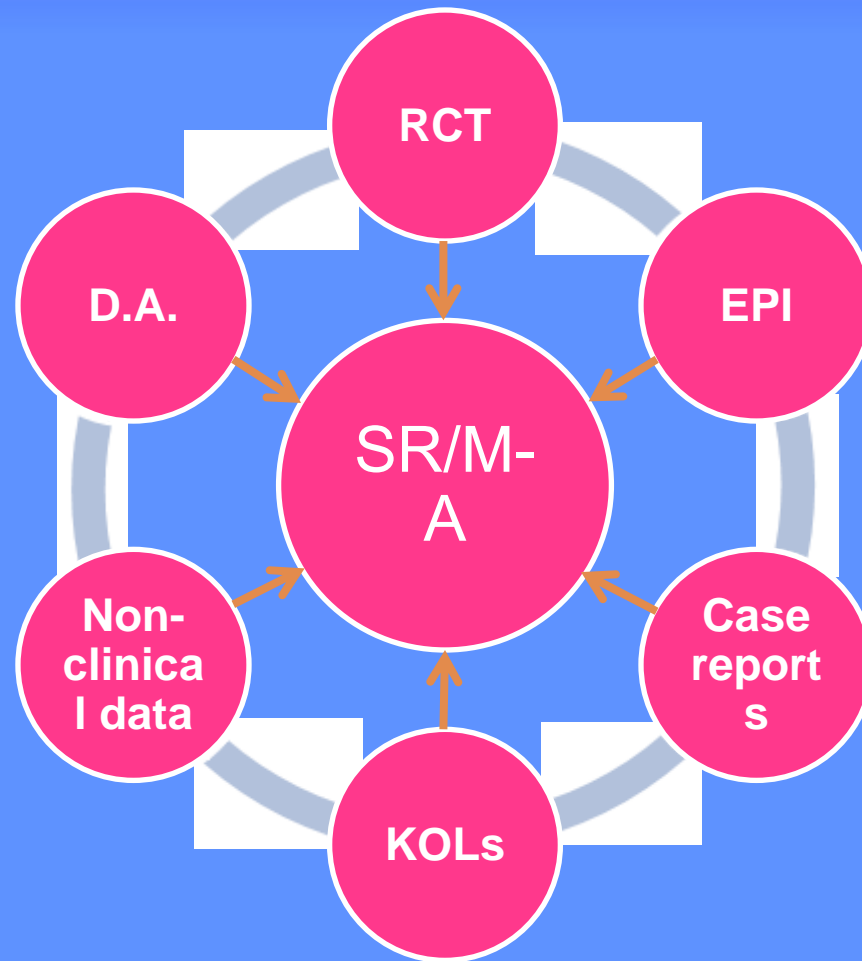
# Chief Question

Are safety data being assembled sufficiently quickly and comprehensively for issues that may go beyond one particular drug?

# Hierarchy of Evidence for Internal Validity of Interventions



# Hierarchy of Evidence for Safety



# Definitions

- **Systematic Review / Overview**  
A comprehensive collation of primary research studies with explicit objectives and methods, and conducted according to explicit and reproducible methods.
- **Meta-analysis**  
The quantitative synthesis of the data within the studies of a systematic review

# Why meta-analysis in safety?

- Quantitative safety assessment requires all relevant available information to be assessed in a timely and unbiased manner
- (Should be) required in all approvals
- Required for quantitative benefit-risk analysis
- (Should be) required in all important post-approval safety assessments

## Essential areas for M-A/SR in Safety:

- Non-common serious adverse events (SAE)
- Small or moderate relative risk increase of an SAE
- Identification of sub-groups at risk – person, therapy (dose, duration, formulation), co-therapy, timing, setting
- Where class effects may be relevant



# Why meta-analysis in safety?

- Increase precision
- Reduce bias
- Assess consistency
  - generalisability
  - sources of heterogeneity
  - Identify sub-groups at risk

# Technical Aims of M-A of controlled studies

	RCT	Cohort (incl. Nested CCS & SCCS)	Case-control
All eligible subjects ↓	All randomised	All selected	All selected
All relevant studies ↓	↓ study selection bias	↓ study selection bias	↓ study selection bias
Sufficient quality ↓	<ul style="list-style-type: none"> <li>• Allocation</li> <li>• Blindness</li> <li>• Assessment of outcome</li> <li>• Attrition</li> </ul>	<ul style="list-style-type: none"> <li>• Selection of cohorts</li> <li>• Comparability of cohorts</li> <li>• Assessment of outcome</li> <li>• Attrition</li> </ul>	<ul style="list-style-type: none"> <li>• Selection of cases /controls</li> <li>• Comparability of cases /controls</li> <li>• Ascertainment of exposure</li> </ul>
Analysed appropriately	Avoid/correct biases from: <ul style="list-style-type: none"> <li>•Allocation → Randomisation</li> <li>•Treatment → Blinding</li> <li>•Assessment → Blinding of intervention / assesor</li> <li>•Attrition → ITT / handle missing data</li> </ul>	Avoid/correct biases from: <ul style="list-style-type: none"> <li>•Allocation → Adjustment</li> <li>•Treatment → Adjustment</li> <li>•Assessment → Blinding of assesor</li> <li>•Attrition → 'ITT' / missing data</li> </ul>	Avoid/correct Biases from: <ul style="list-style-type: none"> <li>•Selection → Adjustment</li> <li>•Treatment → Adjustment</li> <li>•Assessment → Blinding of assesor</li> </ul>

# Types of Meta-analysis

*Systematic Review*

*Meta-analysis* → *Extract data from published reports*

*Collect aggregate data (AD)*

*Collect individual patient data (IPD)  
from conducted studies*

*Collaborative new studies*

*sharing of coefficients*

*sharing of data*

*sharing methodologies*

# Differences in Types of M-A

Type of M-A	Advantages	Disadvantages
Aggregate data	<ul style="list-style-type: none"> <li>• Greater collaboration</li> <li>• Less time &amp; resource</li> </ul>	Limited analysis by: <ul style="list-style-type: none"> <li>• Subgroups</li> <li>• Confounder adjustment</li> <li>• Analyses by time</li> </ul>
Individual patient data of conducted studies	<ul style="list-style-type: none"> <li>• Analysis of subgroups</li> <li>• Confounder adjustment</li> <li>• Analyses by time</li> </ul>	<ul style="list-style-type: none"> <li>• Unavailability of old data</li> <li>• Reluctance to collaborate /share</li> <li>• Time, resource, people &amp; structure</li> <li>• Heterogeneity in methods remains</li> </ul>
Collaborative studies	<ul style="list-style-type: none"> <li>• Analysis of subgroups</li> <li>• Confounder adjustment</li> <li>• Analyses by time</li> <li>• Less heterogeneity from design/definitions/analyses</li> </ul>	<ul style="list-style-type: none"> <li>• How to share?</li> <li>• Time, resource, people &amp; structure</li> <li>• How to keep centers 'involved'/engaged?</li> </ul>

Publication bias – an issue for all types of MA

# Major Sources of Heterogeneity in Controlled Clinical Safety Studies

- Study design - RCT, Cohort (includes nested CCS & SCCS), CCS
- Surveillance/detection methods
- Diagnoses & dictionaries
- Population (& setting)
- Intervention – dose, duration, titration, regime, formulation
- Comparison group
- Co-therapies – interactions
- Analysis – confounders, effect modifiers, loss to FU, time, effect measures, statistical models, multiple testing

# Heterogeneity in EU-ADR study (common methods/definitions)

Table 3. IRRs of UGIB during NSAID use

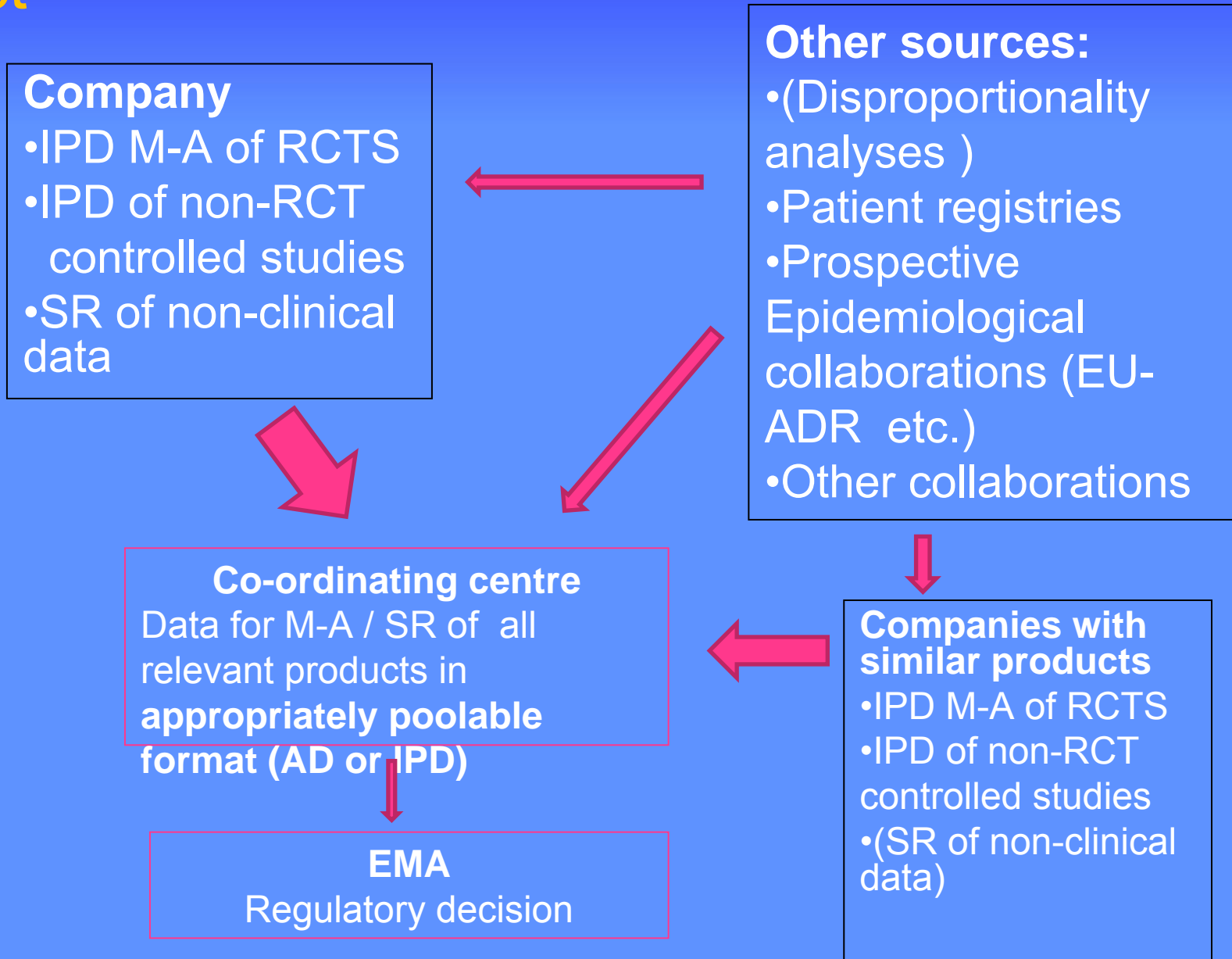
Country	Database	No. of events	Exposure*	Incidence rate <sup>†</sup>	Rate ratio <sup>‡</sup> (95%CI)
ITA	HSD	250	81 734	3.1	2.0 (1.7–2.2)
	Lombardy	991	314 852	3.1	2.9 (2.7–3.1)
	Tuscany	698	205 012	3.4	2.4 (2.3–2.6)
NL	IPCI	116	26 780	4.3	4.0 (3.3–4.9)
	PHARMO	342	177 698	1.9	2.8 (2.5–3.2)
UK	QRESEARCH	467	158 783	2.9	2.4 (2.2–2.6)
DK	Aarhus	2070	316 348	6.5	4.3 (4.1–4.5)
Total		4934	1 281 207	3.9	

\*In person-years.

<sup>†</sup>Per 1000 PYs.

<sup>‡</sup>Age and gender-adjusted; non-NSAID use as comparator; *p* value <<<0.01.

# Ideal Scenario: New/RMP ADR in a New/Old Product



# (Current) Ideal: New ADR in Old Drug



(Rossebø et al, NEJM 2008): Adding ezetimibe to statin versus placebo in aortic stenosis (SEAS trial) → increases cancer RR of 1.55 ( (1.13 to 2.12; P=0.01; 105 vs. 70)

(Peto et al, NEJM 2008): cancer only data from two larger on-going trials: SHARP & IMPROVE-IT  
→ Risk ratio of 0.96 (0.82 to 1.12; 313 vs. 326)  
No significant excess at any particular site  
No trend in cancer incidence/death with FU

→ No change to label

→ Additional studies: FU of large pragmatic trials & observational studies

→ (SHARP, Lancet 2011: Median FU 4.9 yrs, 438 vs 439 cancers)



# Current Situation(FDA): Statins and amyotrophic lateral sclerosis

Pre-approval

New

Mid-cycle

Late cycle

Generics

- AERS disproportionality analysis → EBGM 8.5 to 1.6.
- Aggregate data from 41 statin clinical trials (duration 0.5-5 yrs), 64,000 randomised, 400 000 p-yrs , mean duration of treatment 3.3 yrs), 9 cases of ALS with statins and 10 cases in placebo.
- 4.2 cases per 100,000 p-yrs on statins
- 5.0 cases per 100,000 p-yrs on placebo  
(Colman E, et al. *Pharmacoepidemiol Drug Saf* 2008)
- "continued study of this issue is warranted,"
- CCS in Kaiser Permanente
- Mean duration of statin prescriptions 34.6 days in 2006, (Verispan)

# Other Potential Sources of Data were not used

Cholesterol Treatment Trialists' Collaboration:

IPD Meta-analysis of 21 trials with at least 1000 participants  $\geq 2$  yrs treatment statin versus placebo

	CTTC	FDA analysis
Type of data	IPD	Aggregate
No. trials	21	41
No. patients	129,526	64,000
Median FU	4.8 yrs	3.3 yrs

# Meta-analyses on different levels: the SOS experience

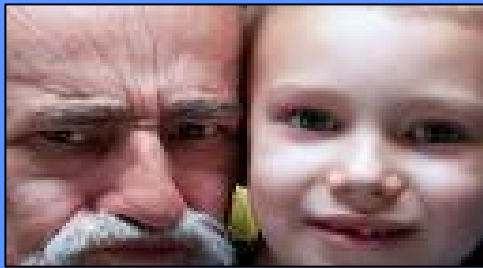
**Miriam Sturkenboom,**  
Erasmus University Medical Center, The  
Netherlands





The overall objective of the SOS project is to assess and compare the risk of cardiovascular events and gastrointestinal events in users of any type of tNSAIDs or COXIB.

The ultimate goal of this project is to provide evidence for regulatory and treatment decision making.



**For adults and children**



However CV is not relevant for children and other safety issues were added for them upon consultation of pediatricians (Reye, liver failure, Stevens Johnson, asthma exacerbation, renal failure, anaphylaxis)

# Example: SOS

## Safety of NSAIDs (funded by FP7 on request of EMA)

- **Partners:** Erasmus MC, RTI, University of Bordeaux, PHARMO, PEDIANET, University of Milano Bicocca, University of Bremen, McGill, ASL di Cremona, University Hospital Padova, FIMIM, University of Nottingham)
- Do Meta-analyses of literature (RTI, Bordeaux)
- Do IPD meta-analysis of conducted studies (McGill)
- Do collaborative multi-database study (8 databases)
- Purpose: to assist physician and regulatory decision making



# Meta-analysis of literature

- Clinical trials and meta-analyses of NSAIDs (University of Bordeaux, RTI, University hospital Padova)
  - *De Salvo et al. CPT 2011:*
    - Conclusion: very few RCTs with enough information on UGIC/CV safety prior to coxib era
    - Observational studies are necessary to complement information from trials
- Observational studies (RTI)
  - Very few studies that allow for assessment of stroke risk
  - Very few studies that allow for assessment of duration and dose effects (meta-analysis results available from website)



# Example of MA based on IPD of conducted studies

- NSAIDs and MI: McGill University (courtesy J. Brophy, McGill)
  - Aim: IPD meta-analysis of 8 conducted observational studies (focus: subgroup analyses, duration relationships)
  - Only 4 studies could share IPD data
  - Took several years and a lot of effort to share data, despite willingness of investigators



# Advantages of collaborative observational studies

- Complementary to meta-analysis on heterogeneously defined outcomes, exposures, designs as in regular meta-analysis
- Possibility to have common protocol, common outcome and exposure definitions and common and shared analysis plan
  - Reduce heterogeneity





# How do SOS partners collaborate in multi -database studies?

- Common protocol
- Common outcome definitions
- Systematic exposure assessment and drug utilization analyses
- Common software for standardized distributed data elaboration on common data models (Jerboa as in EU-ADR)
- Common secure remote research environment
- Distributed analyses and PI ship



# Databases



SISR

9,000,000

general population

ICD-9



OSSIFF

3,000,000

general population

ICD-9



Pedianet

160,000

children, general population

ICD-9, free text



IPCI

1,000,000

general population

ICPC, free text



PHARMO

3,000,000

general population

ICD-9



BIPS

13,600,000

general population

ICD-10-GM



THIN  
QRESEARCH

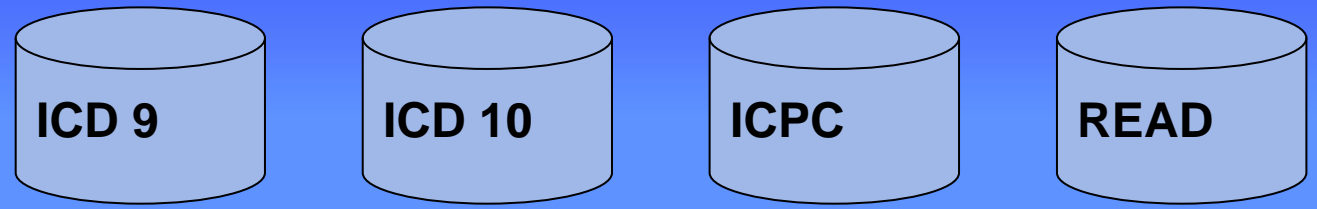
3,600,000

6,000,000

general population

READ, free text

# Terminology Mapping



Choice of the events  
**4 outcomes**  
**41 confounders**  
**1 exclusion**

**Common semantic base – WP 6.2**  
**Terminology mapping**

**6526 UMLS different concepts**  
(concepts existing in at least one of the 4 terminologies)

Number of corresponding codes  
According to terminology

**2406**

**1614**

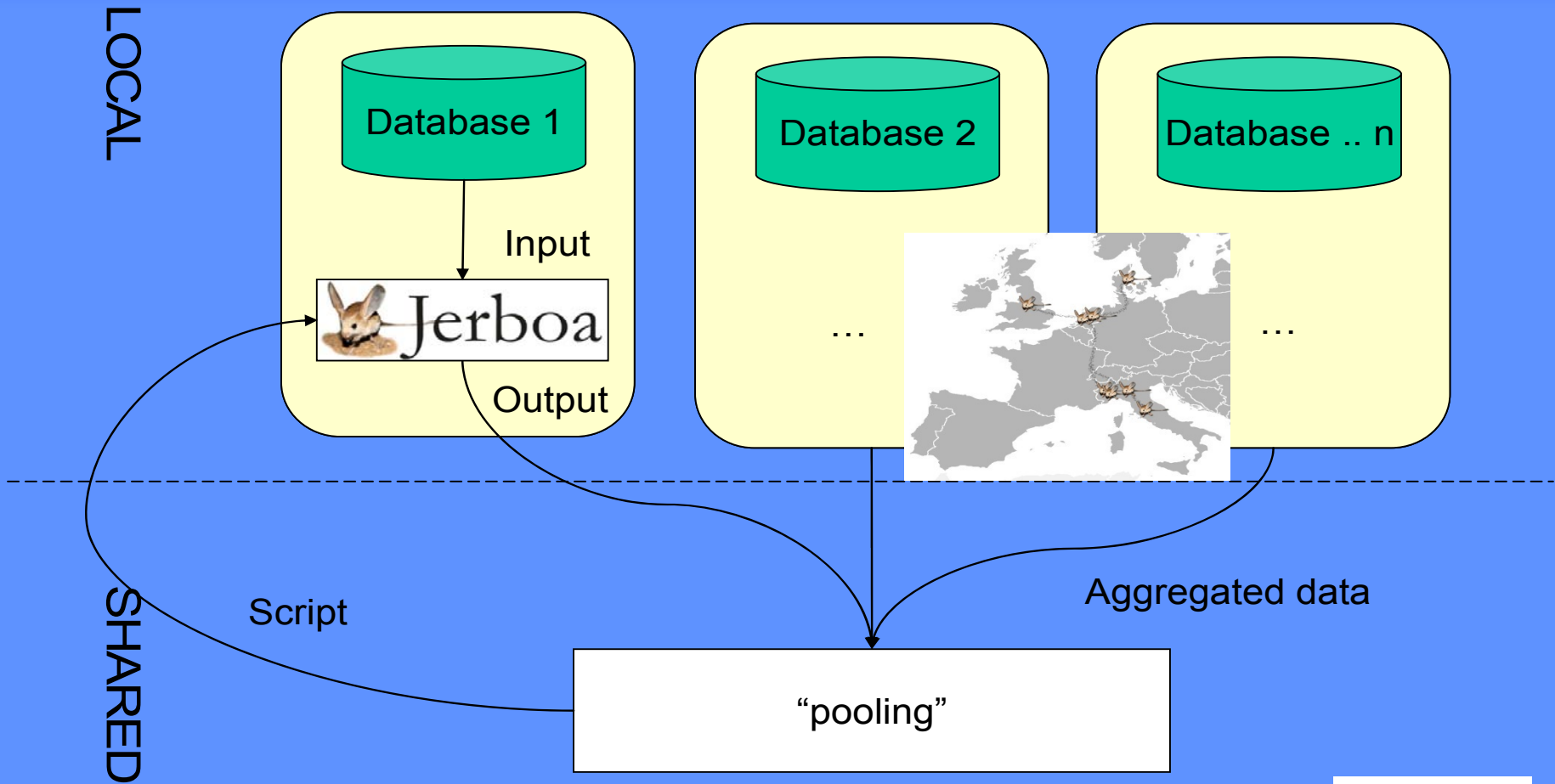
**517**

**4274** V2

**4544** V3

# Distributed Data Network

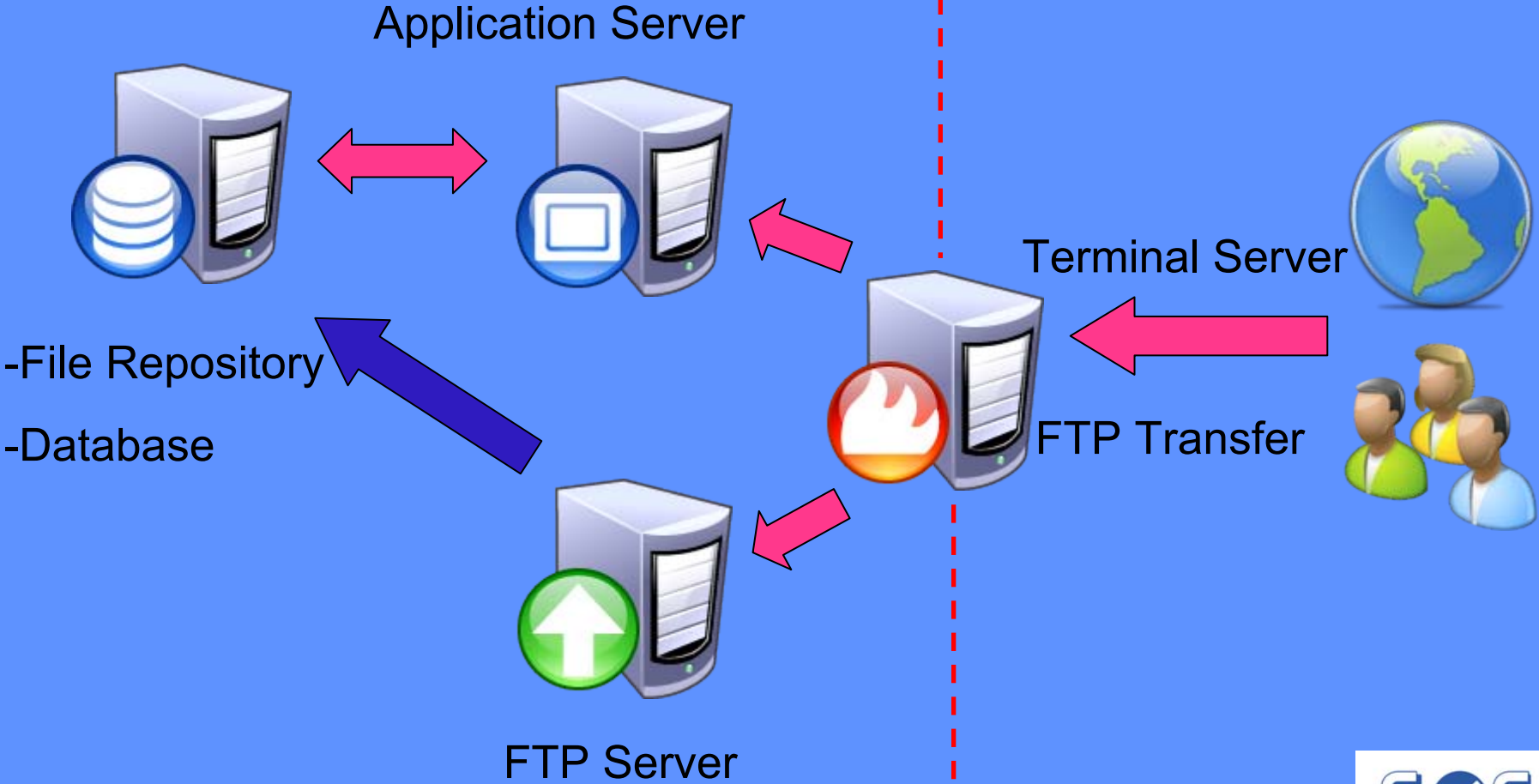
Jerboa



# Remote Research Environment

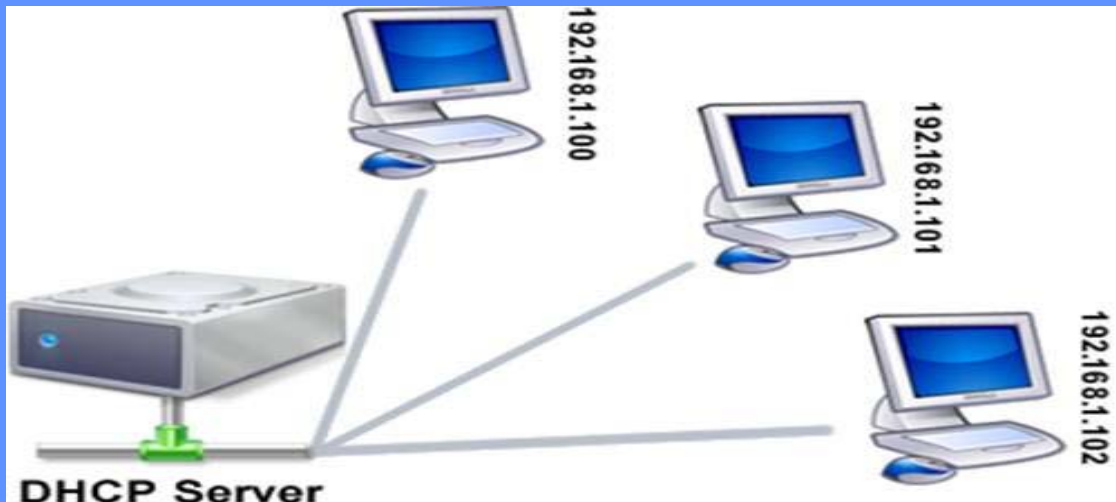
University of Milan-Bicocca

SOS Partners



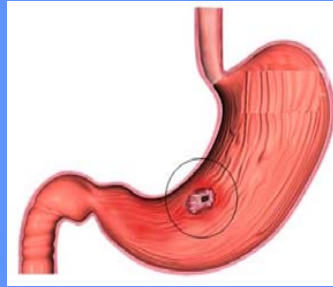
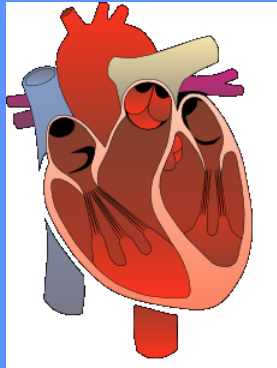
# Remote Data Access: Security & privacy issues

- Each partner has personal account
- Each partner has access only to specific folder
- List of specific authorized IP addresses of the SOS partners
- Saved credentials are not allowed
- Personal Token authentication required.



# Key output: decision model

- Many different NSAIDs
- Different risks



- Used across various populations

Which NSAID for given patient?



# Key methodological question at the end of project

- What do the collaborative database studies add to IPD, regular meta-analysis?



How to go ahead?

Two proposals for discussion

Nawab Qizilbash

## Areas for Development by ENCePP?

1. ENCePP Methodological Guidelines  
document for meta-analysis of controlled  
epidemiological studies
- 2.A 'structure' for data integration

# ENCePP Methodological Guidelines document

- Is such a document needed?
- What should be the purpose?
  - Protocol development guideline
  - Reporting guideline for safety
- What should be included?
- Timelines?

# Is there a need for a 'structure'?

Create an 'ENCePP Systematic Review Group/Secretariat' via 'ENCePP Data Sources Working Group'?

## WHAT MIGHT BE THE TASKS

- Receive issues for pooling (PRAC?)
- Evaluate suitability for pooling
- Assess scope of the SR
- Identify available potential data sources: trials, epidemiological studies, spontaneous AE database analyses and non-human data
- Seek collaboration for data
- Identify people for the Steering, Secretariat and Writing groups for each issue
- Protocol development
- Perform the overview / meta-analysis
- Reporting (to whom?)
- Develop methodology
- [Use existing resources where possible, e.g. UK NICE model?]

# Some Elements for Data Integration

- Trials and controlled epidemiological studies should be meta-analysed separately and together
- SR should include disproportionality analyses and non-clinical data

Thank you

&

Discussion