Thoracic Oncology Padova
March 29-30, 2019

Challenges and opportunities of Real World Datasets

Paolo Bruzzi - Genova

Contents (15')

- Glossary
  - Datasets vs Registries vs Studies vs Aims
  - Trials vs Observational Studies
  - Efficacy vs Effectiveness
  - Internal Validity vs External Validity
  - Scope of observational studies
- *Challenges and opportunities of Real World Datasets?*

Contents (15’)

- Glossary?

“There is great chaos under heaven …”

Mao Zedong, 1966, letter to his wife

What are “real life” data?

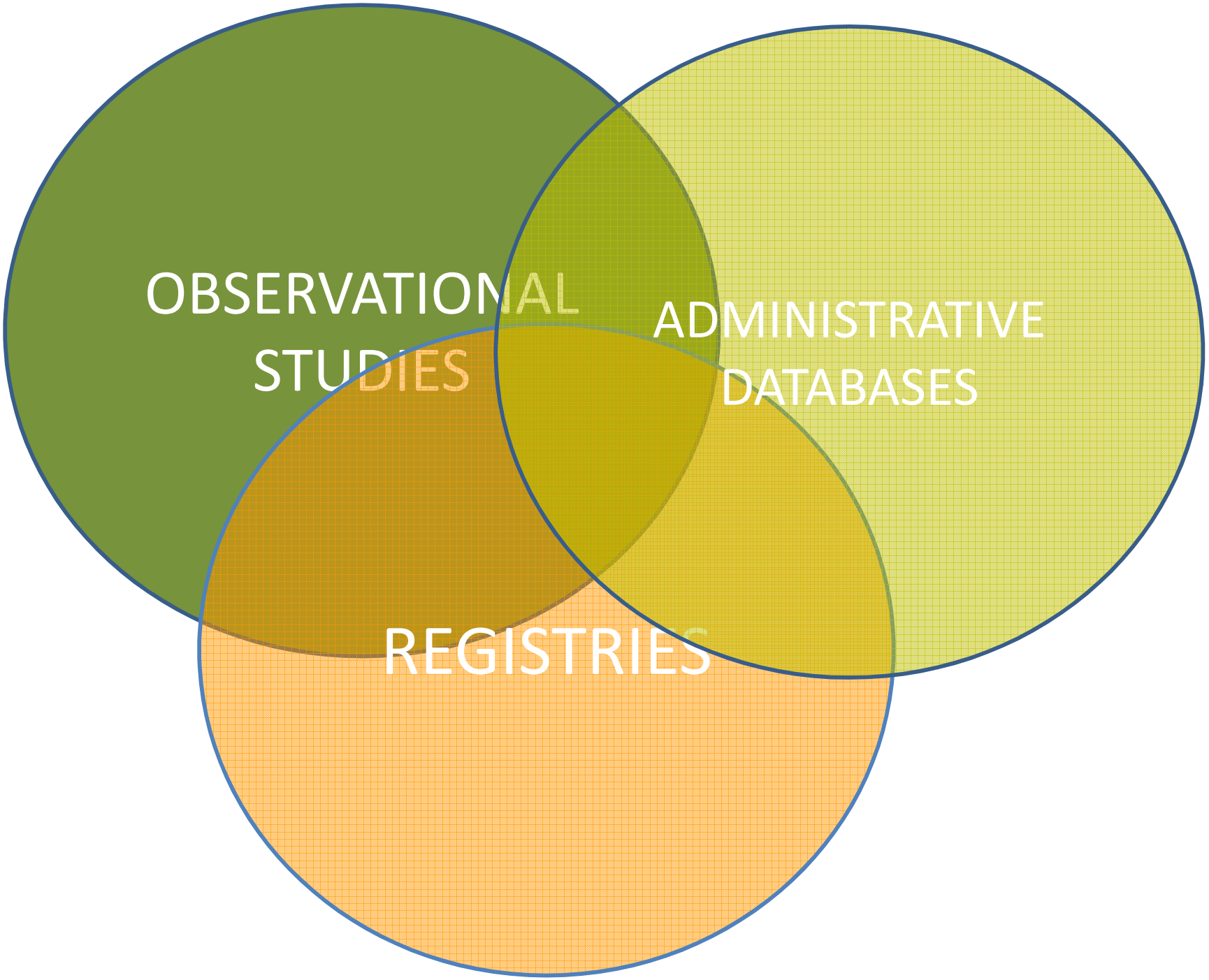
“Data used for decision-making that are not collected in conventional randomized clinical trials”

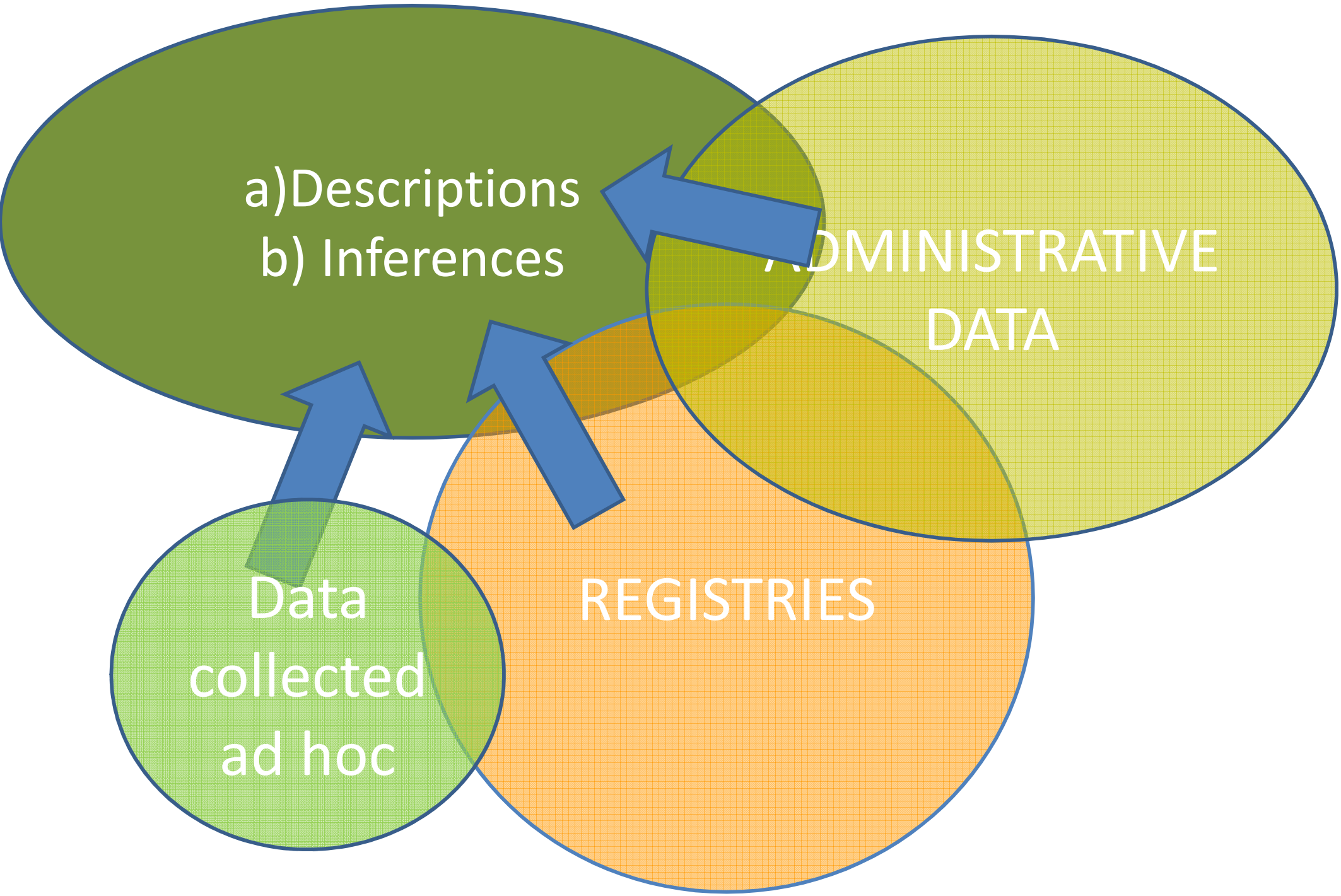
International Society of Pharmacoeconomics and Outcomes Research

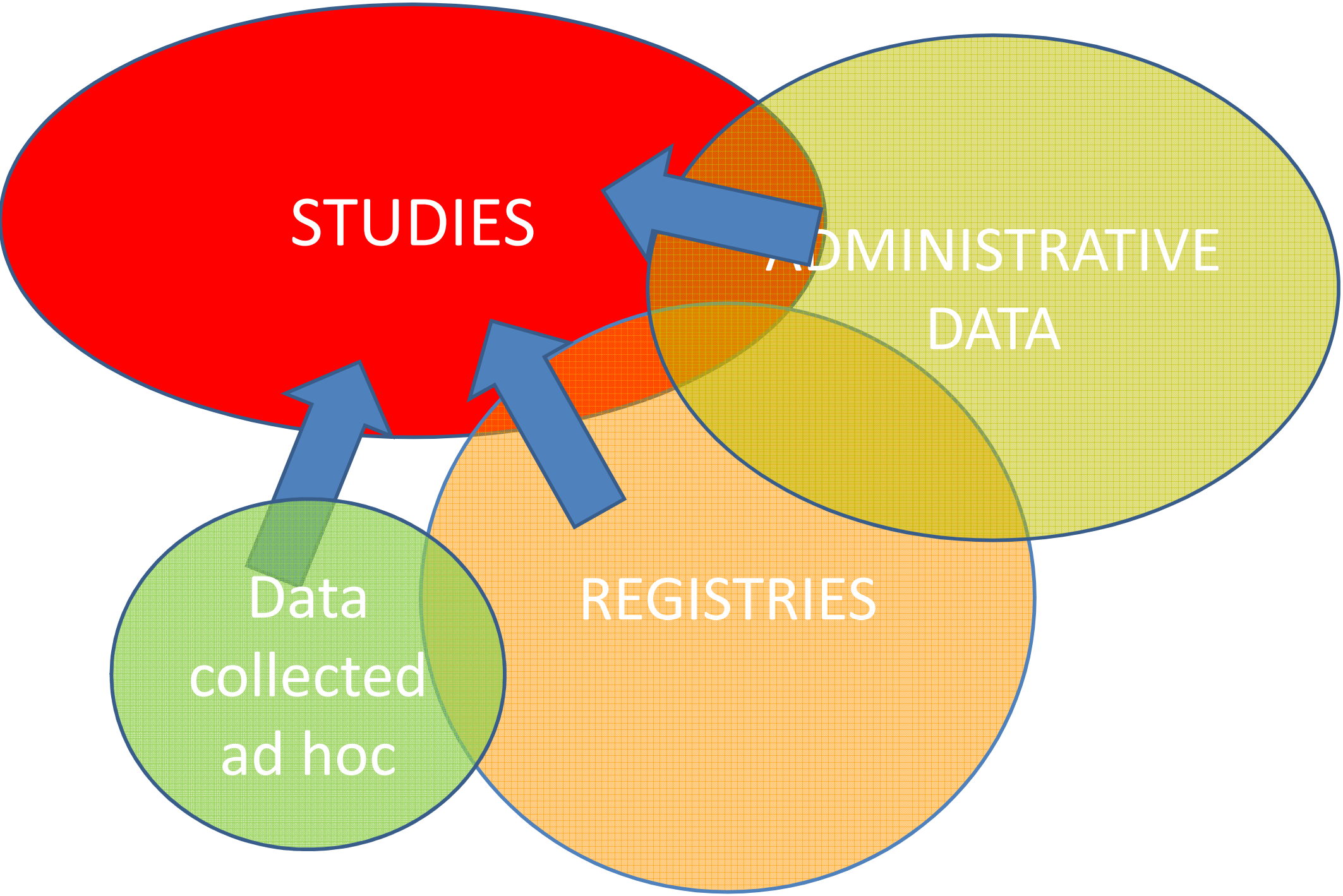
Sources of Real-World Data

- **Pragmatic clinical trials**
- Prospective observational studies and patient registries
- Administrative claims data
- Patient surveys
- Electronic health records/medical chart reviews
- Government- or third-party-sponsored systematic surveys that assess public health, resource consumption, practice patterns, and clinical trends

Garrison LP, et al. *Value Health*. 2007;10(5):326-335.


OBSERVATIONAL STUDIES
ADMINISTRATIVE DATABASES
REGISTRIES


a)Descriptions
b) Inferences
ADMINISTRATIVE DATA
Data collected ad hoc
REGISTRIES


STUDIES
ADMINISTRATIVE DATA
Data collected ad hoc
REGISTRIES

Contents (15’)

- Glossary
  - Datasets vs Registries vs Studies vs Aims

Definitions

- Dataset -> A collection of data
  - No aims
  - Specific/generic aims ≈ Registry
- Registry -> A planned & structured data collection

- Study -> An endeavor aimed at answering questions
- Aim -> Aim

Examples

- Dataset ->
  - No aims
  - Specific/generic aims
- Registry ->
- Study ->
- Aim -> Aim

- Any set of records (computer, paper, etc.)
  - Clinical database, big data
  - Birth Records, Hospital DF's
- Cancer Registry
- Trial, Observational Study
- To evaluate..
  - the incidence of...
  - The efficacy of...

Main Differences

Dataset

A collection of data from one or multiple sources

Registry

- **Prior definition of**
  - **Target Population**
  - **Data to be collected**
  - **Sources**
    - Existing databases
    - Ad hoc data collection
  - Format of data (structure of DB)
- Active search of missing data
- **Multiple aims (No specific aim)**

Main Differences

Registry

A planned & structured collection of data, usually with multiple but generic aims

Study

- Prior definition of
  - **STUDY AIM(S) = Study question**
  - **Eligible Population**
  - Data to be collected
  - Sources
    - Existing databases
    - Ad hoc data collection
  - Format of data (structure of DB)
- Active search of missing data

Main Differences

Registry

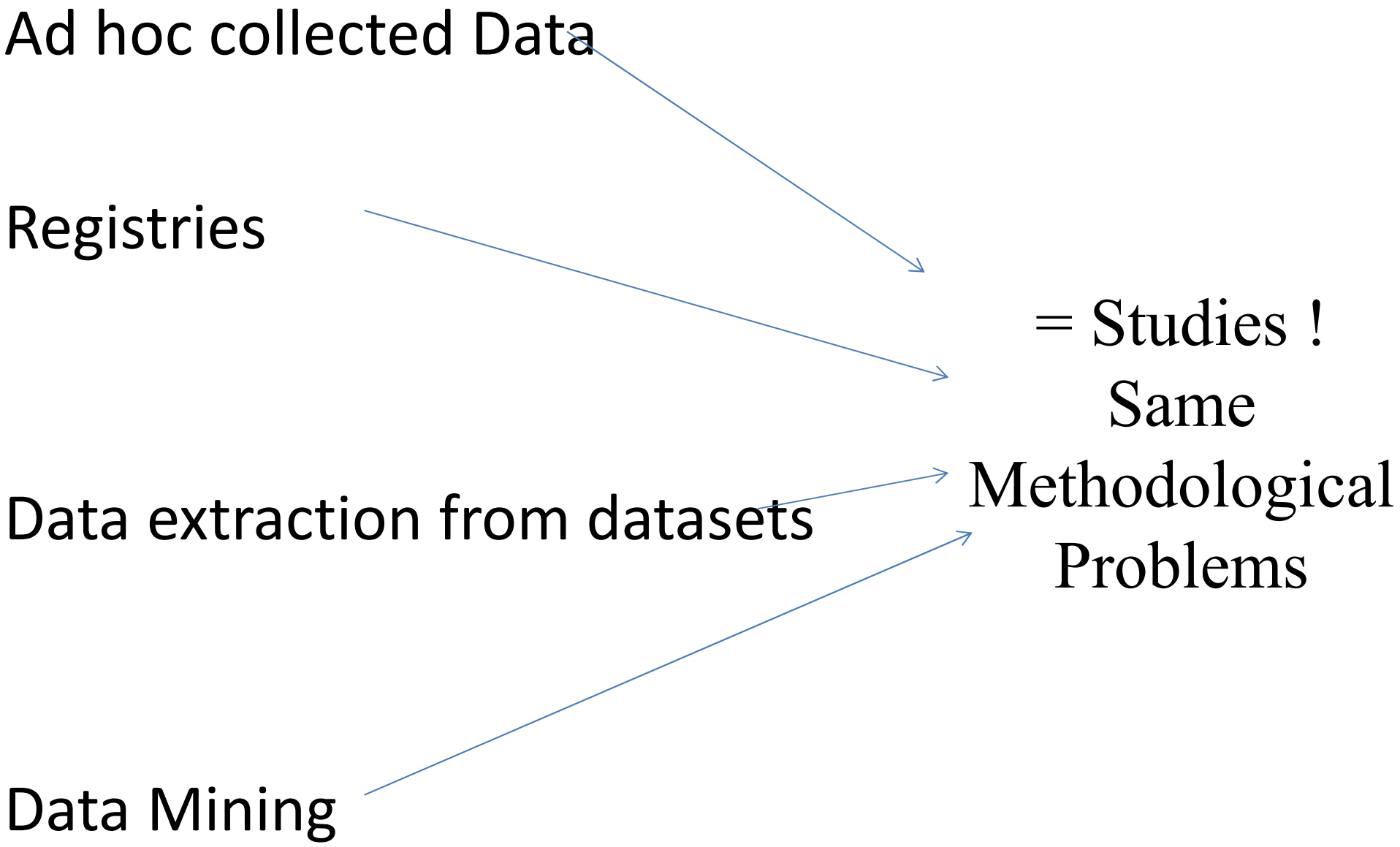
A structured collection of data, usually with multiple but generic aims

Study

- **Study design**
- **Statistical Plan**
- **Timeline**
- **Intervention (if any)**

Exam in any basic course in statistics or research methodology

- Question:
  - Why careful design and skilled statistical planning are required before doing a study?
- Answer:
  - To prevent or assess the classical errors due to selection bias, assessment bias, etc.
  - To avoid the confusion between the generation of hypotheses and their demonstration
  - To accomodate the problems related to multiplicity


Ad hoc collected Data
Registries
Data extraction from datasets
Data Mining
= Studies !
Same
Methodological
Problems

1st Conclusion

- Datasets and Registries DO NOT provide knowledge, but information that may prove useful to create knowledge
- Building Knowledge is possible only by means of (adequately planned and conducted) STUDIES
- The Analysis of a dataset without a prior adequate study design is doing a study of poor quality

Observational Studies vs Trials

- **Observational Studies** = Studies in which
  - No change is introduced in the standard routine clinical practice
  - With the exception of tests/exams/questionnaires/etc aimed at gathering information needed for the study
  - **That is not used in the management of the patient**

- **Trials** = All clinical studies that cannot be classified as observational

STUDIES

Observational study

- Prior definition of
  - **STUDY AIM(S) = Study question**
  - **Eligible Population**
  - Data to be collected
  - Sources
    - Existing databases
    - Ad hoc data collection
  - Format of data (structure of DB)
- Active search of missing data

Experimental Study (= Trial)

- Prior definition of
  - **STUDY AIM(S) = Study question**
  - **Eligible Population**
  - Data to be collected
  - Sources
    - Existing databases
    - Ad hoc data collection
  - Format of data (structure of DB)
- Active search of missing data

STUDIES

Observational study

- **Study design**
- **Statistical Plan**
- **Timeline**

Experimental Study (= Trial)

- **Study design**
- **Statistical Plan**
- **Timeline**
- **Intervention**

TRIALS

Uncontrolled Trial

- **Study Aim(s)**
- **Selection of treated & Controls**
- **Unbiased assessment of Outcome**
- **Statistical Plan**

Randomised Trial

- **Study Aims**
- **Random assignment of treatments**
- **Unbiased assessment of Outcome**
- **Statistical Plan**

Note: The GRADE method

According to the Grade method (*Grading of Recommendations Assessment, Development and Evaluation)* **all clinical studies that are not randomised are referred to as observational studies**

Observational studies = Uncontrolled trials?

Observational Trials???

Observational Trial? -> Oximoron!

Observational study

- Therapies already in use
- Unselected centers
- Unselected patients
- Routine assessments

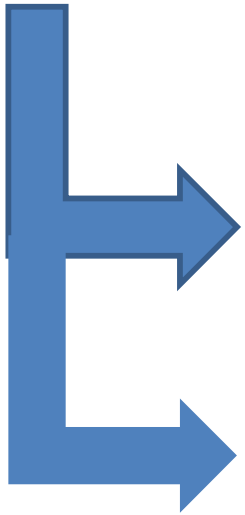
Uncontrolled trial

- New therapies
- Specialised centers
- Selected patients
- Special assessments

Dominant Thinking

The Randomised Clinical Trial is the only source of reliable evidence on the efficacy of medical interventions

- Uncontrolled trials, observational studies



Bias in the selection of treated (and controls, if any)

Bias in the assessment of outcome

Contents

- Glossary
  - *Datasets vs Registries vs Studies vs Aims*
  - *Trials vs Observational Studies*
  - Efficacy vs Effectiveness

Efficacy vs Effectiveness

- **Efficacy** is the extent to which an intervention does more good than harm
  under ideal circumstances

- **Effectiveness** assesses whether an intervention does more good than harm when provided under usual circumstances of healthcare practice

Cochrane AL. Effectiveness and efficiency: random reflection on health services. London: Nuffield Provincial Hospitals Trust; 1972.

Argument

- Do (Randomised) Clinical Trials provide evidence that is sufficient to evaluate the true impact of a new therapy in routine clinical practice?
  - “Fit” patients
  - Young patients
  - Specialised Centers
  - Special procedures
  - “Carefulness”

Argument

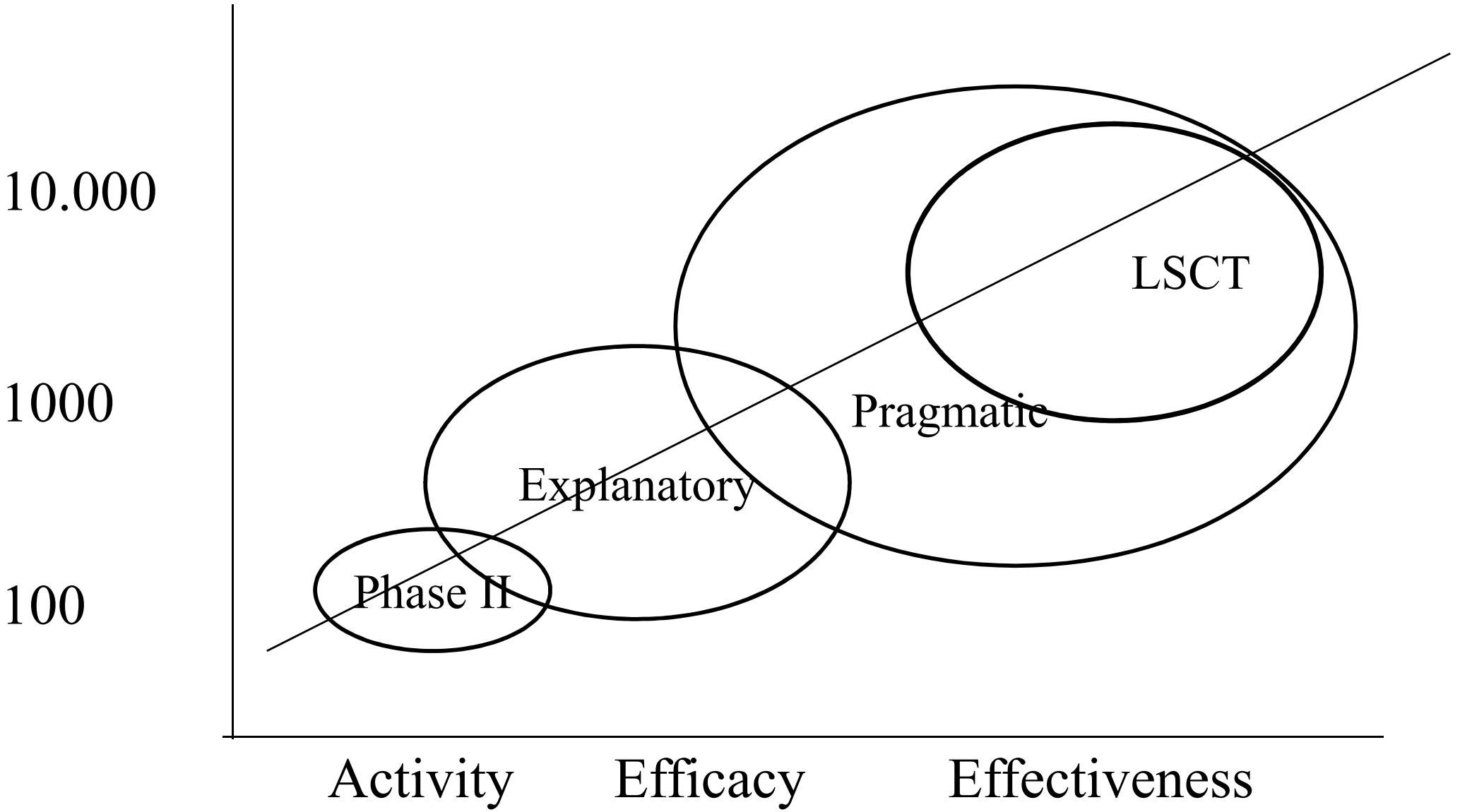
- Do (Randomised) Clinical Trials provide evidence that is sufficient to evaluate the true impact of a new therapy in routine clinical practice?

NO!

Solution

  - Explanatory trials -> Efficacy
  - Pragmatic trials -> Effectiveness

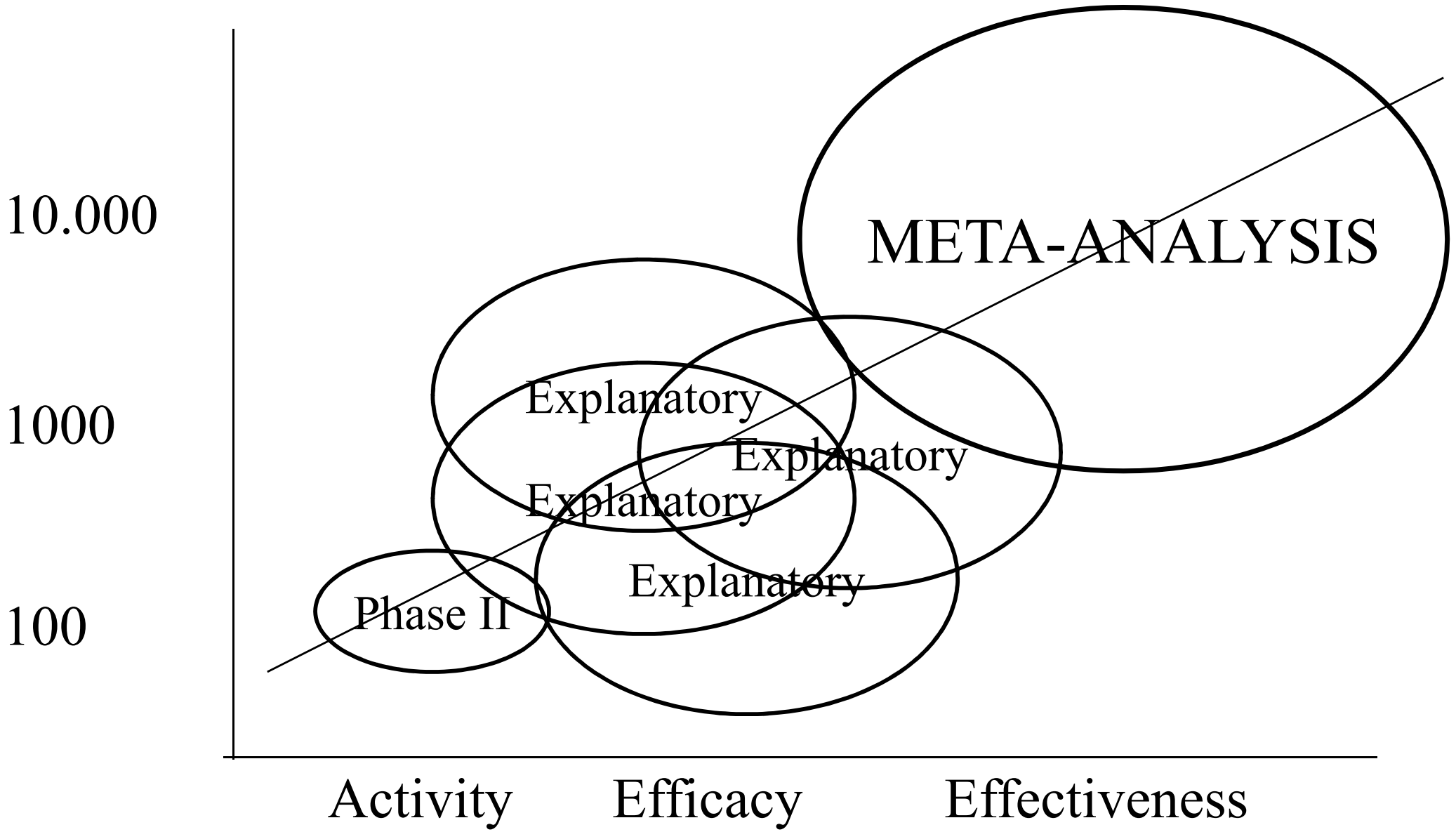
  (Unselected patients and centers, routine care +/- experimental at random) -> Large &Simple Clinical trials


10.000
1000
100
LSCT
Pragmatic
Explanatory
Phase II
Activity
Efficacy
Effectiveness

Problem

IT IS NOT ACCEPTABLE, FROM AN UNETHICAL VIEWPOINT, TO LAUNCH A PRAGMATIC RANDOMISED TRIAL WHEN ONE TREATMENT HAS BEEN SHOWN IN AN EXPLANATORY RANDOMISED TRIAL TO BE BETTER THAN THE OTHER

Solution: Meta-analysis!?


10.000
1000
100
META-ANALYSIS
Explanatory
Explanatory
Explanatory
Explanatory
Phase II
Activity
Efficacy
Effectiveness

The EBCTCG involves almost all trialists worldwide who have done relevant randomised trials of the treatment of women with breast cancer, …

Several hundred research groups have shared individual patient data on more than **450,000 women in 400 randomised trials** for meta-analyses

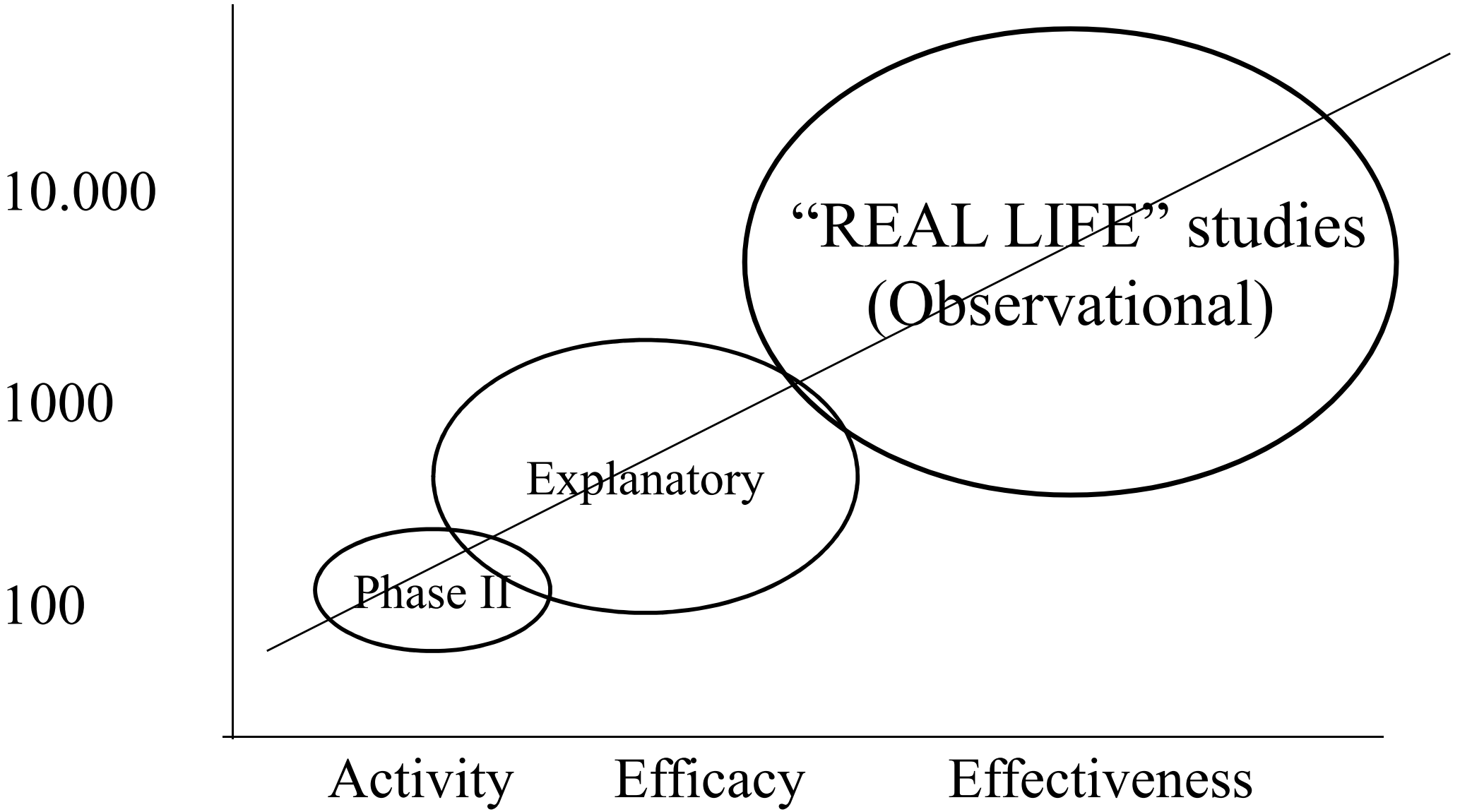
Problem

Modern Oncology

1 drug -> 1 condition -> 1 trial-> Registration

What do we meta-analyse?

How do we treat patients excluded from trials?


10.000
1000
100
"REAL LIFE" studies
(Observational)
Explanatory
Phase II
Activity
Efficacy
Effectiveness

Contents

- Glossary
  - *Datasets vs Registries vs Studies vs Aims*
  - *Trials vs Observational Studies*
  - *Efficacy vs Effectiveness*
  - **Internal Validity vs External Validity**

Interpretation the result of a study

1. Internal Validity

Is it true?

2. External Validity

So what?

Internal Validity

- Research protocol
- Identification of the primary aim(s)
- **Randomization Procedures**
- **Endpoint – Choice/Assessment/Blinding**
- **Statistical Plan**
- **Lost to follow-up/not evaluated**
- **Analysis ‘intention to treat’**

Interpretation

- *Internal Validity:*

- **<u>External Validity</u>**:

Possibility to
- Extrapolate
- Generalise
- Apply

the results of the study

Determinants of External Validity

- **INTERNAL VALIDITY!!!**
- *Study Design (Contrast)*
- *Selection Criteria/Patients Characteristics*
- *Participating Centers*
- *Treatment Protocol –Follow-up Protocol*
- *Endpoint*
- *Compliance – Contamination*
- *Precision of the estimates*
- *Analysis 'intention to treat'*

Observational Studies

Due to any (or most) of the following drawbacks

- Lack of
  - – active selection of cases receiving the new therapy
  - – concurrent, comparable controls receving standard treatment **(Randomization)**
  - – adequate follow-up (ITT) & unbiased endpoint assessment (Masking)

The Internal validity of observational studies is ALWAYS QUESTIONABLE

Observational Studies

METHODOLOGIC PROBLEMS

Depend entirely on the study aim!

Observational Studies: Possible aims

To gather (further) information on

- Efficacy
- Toxicity/Adverse effects
- Modes of use
- Patterns of Care

Observational Studies: Possible aims

To gather (further) information on

- Efficacy in specific care settings or patients subgroups

Question

- To estimate the efficacy of a new drug we require a randomized trial with strict methodological & statistical standards (masking, Intention to treat, Planned statistical analyses, etc)
- Is it conceivable that we modify these estimates in specific care settings or patients subgroups based on the results of (observational) studies of questionable validity?

2nd Conclusion

Observational studies should not be used to modify the estimates of the efficacy of treatments obtained in randomised clinical trials, unless dramatic effects (lack of) are seen in subgroups previously not included in efficacy trials

Observational Studies: Possible aims

To gather (further) information on

- ~~*Efficacy*~~

- Toxicity/Adverse effects
  - Rare events
  - Rare Subgroups
  - Patients not included in trials
- Pharmacoepidemiology - Data Mining (Big Data)?

Observational Studies: Possible aims

To gather (further) information on

- ~~*Efficacy*~~
- *Toxicity/Adverse effects*

- Modes of use
- Patterns of Care

} Quality of care
Critical Research area
Few Studies usually of poor quality

Observational Studies of quality of care (appropriateness)

Methodological Problems

1) Selection Bias
2) Assessment
3) **CONFOUNDING**

Association does not mean causation!

(e.g. lower % of histological confirmation in geriatric oncologies)

Contents

- Glossary
  - *Datasets vs Registries vs Studies vs Aims*
  - *Trials vs Observational Studies*
  - *Efficacy vs Effectiveness*
  - *Internal Validity vs External Validity*
  - **Scope of observational studies**

Summary
Scope of observational *real life* **studies**

Descriptive Epidemiology
- Epidemiological Determinants
- Outcomes

} Needs

- Efficiency/Costs
- Appropriateness
- (*Effects of care??Efficacy??)*

} Quality of Care

(**Historical controls for uncontrolled trials?**)

Contents (15')

- Glossary
  - *Datasets vs Registries vs Studies vs Aims*
  - *Trials vs Observational Studies*
  - *Efficacy vs Effectiveness*
  - *Internal Validity vs External Validity*
  - *Scope of observational studies*
- *Challenges and opportunities of Real World Datasets?*

Real world data?

- Real world datasets will not *(are not likely to)* prove useful sources of information on the efficacy of treatments
- They may prove very important source of information for observational studies, including those aimed at the assessment of the health impact of new technologies
- However, so far, **most Real Life studies have been poorly planned, conducted and reported, while their methodology is more challenging than that of Randomised Clinical Trials**

Contents (15’)

- Glossary?

“There is great chaos under heaven ….

….and the situation is excellent”

Mao Zedong, 1966, letter to his wife