

# Validation in epidemiological studies



Atsumi, living with epilepsy

# Validation or verification?

- We all know how to validate ! (?)
- From clinical trials?
- From Epidemiological studies?
- Via double programming?
- What exactly is validation?
- What is verification?
- What is feasible in epidemiology?
- What is feasible on claims data?
- Lets have a look on some “advices”

# (Some) Literature about pharmacoepidemiology:

- ISPE: Guidelines for good pharmacoepidemiology practices (GPP)
- ISPE: Guidelines for Good Database Selection and use in Pharmacoepidemiology Research
- FDA: Good Pharmacovigilance Practices and Pharmacoepidemiologic Assessment
- FDA: Best Practices for Conducting and Reporting Pharmacoepidemiologic Safety Studies Using Electronic Healthcare Data Sets
- ISPOR: Using Real-World Data for Coverage and Payment Decisions
- EMA: Guideline on good pharmacovigilance practices (GVP)
- EMA: Guide on Methodological Standards in Pharmacoepidemiology (Revision 1)

# Good Pharmacovigilance Practices and Pharmacoepidemiologic Assessment FDA March 2005

- For most pharmacoepidemiologic studies, FDA recommends that sponsors validate diagnostic findings through a detailed review of at least a ***sample of medical records***.
- Double programming?
- Feasible?
- Sample size?

# Best Practices for Conducting and Reporting Pharmacoepidemiologic Safety Studies Using Electronic Healthcare Data Sets

- ⊗ Because electronic administrative claims data are ***not collected for investigative purposes***, but rather for patient care or reimbursement purposes, it is vitally important to ensure that medical outcomes of interest are validated (Lanes). Validation of administrative claims data is the process through which primary medical data (generally medical charts) are abstracted and reviewed to determine whether the patient actually experienced the event coded (or suggested by the algorithm if applicable) in the electronic data.
- ⊗ Double programming?
- ⊗ Feasible?

## Best Practices for Conducting and Reporting Pharmacoepidemiologic Safety Studies Using Electronic Healthcare Data Sets – EMR data

- ⊗ There is still a scientific need, however, to develop and employ strategies for ensuring that the electronic data accurately reflect patient experience.
- ⊗ As implementation of EMRs becomes more widespread, investigators will be challenged to ***develop innovative strategies*** to confirm electronic exposure and outcome data, and FDA encourages such efforts as they are critical to ensure the validity of studies relying upon these data.
- ⊗ Double programming?
- ⊗ Feasible?

# ISPE: Guidelines for good pharmacoepidemiology practices (GPP)

- Use ***validated instruments*** and measures whenever such exist, and describe the validation method.
- Reasonable effort should be made to document and validate ***interim steps*** in the analysis.
- Double programming?
- Feasible?
- What is a validated instrument?
- What are the right interim steps?

# ISPE: Guidelines for Good Database Selection and use in Pharmacoepidemiology Research

- ⊗ Data quality sufficient to complete and interpret analysis
- ⊗ completeness and accuracy of key study variables
- ⊗ Quality control of study population to validate study-specific variables
- ⊗ Programme review/testing at each stage of extraction/analysis.
- ⊗ create an artificially engineered (?) sample of complex patient data that include every(?) permutation of the extraction.
- ⊗ Simple consistency in data capture over time
- ⊗ Interdependence of variables within a case may be examined.
  
- ⊗ Double programming?
  
- ⊗ Feasible?

# IEEE: Definition of validation/verification

- ⊗ Verification - ensure product fulfills requirements during development cycle
- ⊗ Double programming?
- ⊗ Feasible?
- ⊗ Validation - evaluate end product to ensure compliance with (software) requirements
- ⊗ Double programming?
- ⊗ Feasible?

# Proposal claims data:

- Derive a standard access to any database in use, e.G. a validated approach for the first data selection and merge.
- Validate the standard access.
  - You may use an artificial sample with all cases if this is feasible
  - You may use a sample with the most important cases only.
  - You may also use a sample of the database to develop.  
Depending on the sample size you will cover all issues of interests.
  - Amend your standard over the time.
- Document your standard well. It's a good way to improve your programming documentation.
- Benefit:
- Different databases can be handled similar afterwards.
- Faster development.
- Avoid errors.

# Proposal claims data:

- ⊗ Build (validated) standard tools for common tasks.
  - Analysis datasets:
    - For final selections.
    - For matching with control groups.
    - For deriving common covariates.
    - To define episodes
    - .....
  - Outputs:
    - Deliver standardised tables figures listings
    - Documentation of selection
    - .....
- ⊗ Benefit: Your results will be more comparable between studies

# Proposal claims data :

Don't end up like this guys from the stone age.

Obviously they don't have a standard direction for sun light to compare their time measurement.



# Proposal claims data :

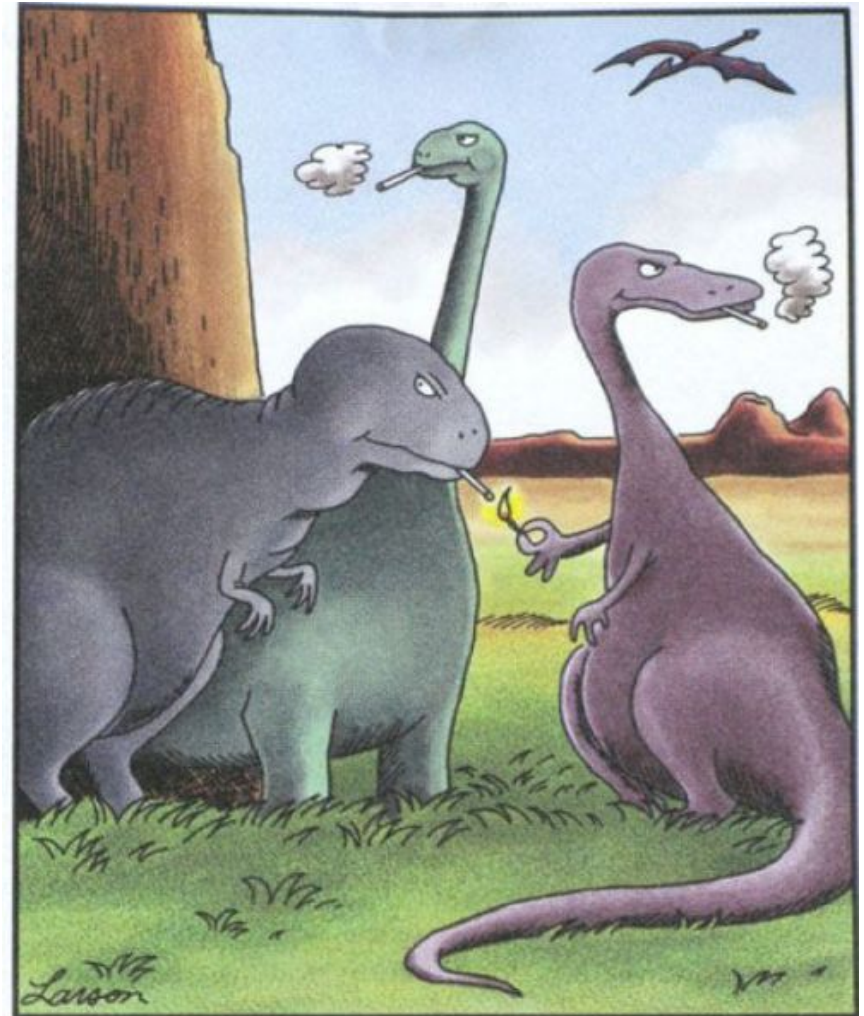
Challenge your standard instruments.

Compare your standard instruments with others

Don't believe your standards are perfect because you applied them with success.

Challenge any new theory before implementation.

Implement only what you really understand.



The real reason dinosaurs became extinct

# Proposal claims data: „Standard tools, data access“



Review  
database  
documen-  
tations



Consider a  
common  
structure  
delivered  
by first  
data  
access.



Implement  
joins  
between  
datasets  
(Little  
sample is  
enough)



Review  
result for  
some  
records to  
review  
endpoint  
definitons

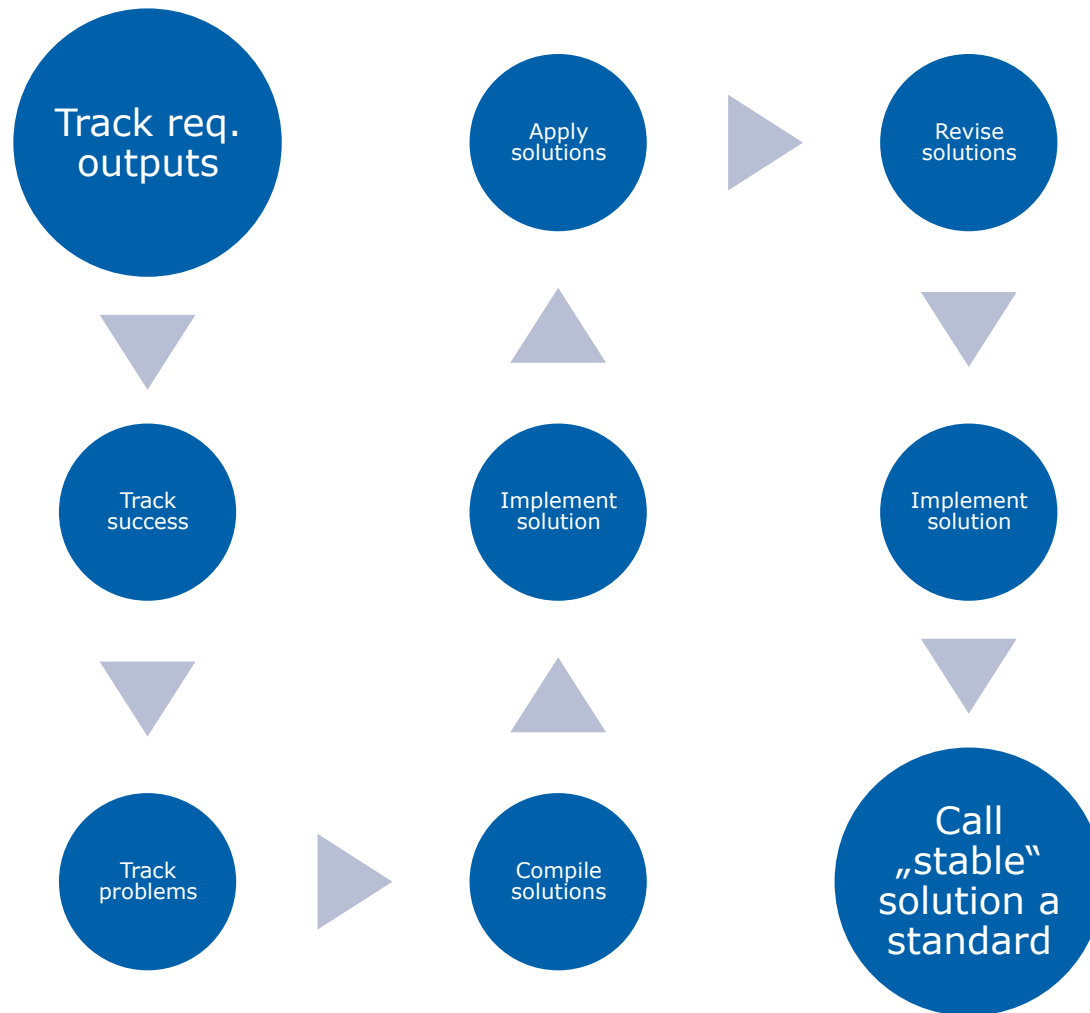
For example, to address any issue that may occur in any 100`th record with a propability of 0.95 you need a sample size of 298 records.

# Proposal claims data: „Standard tools, derivations“

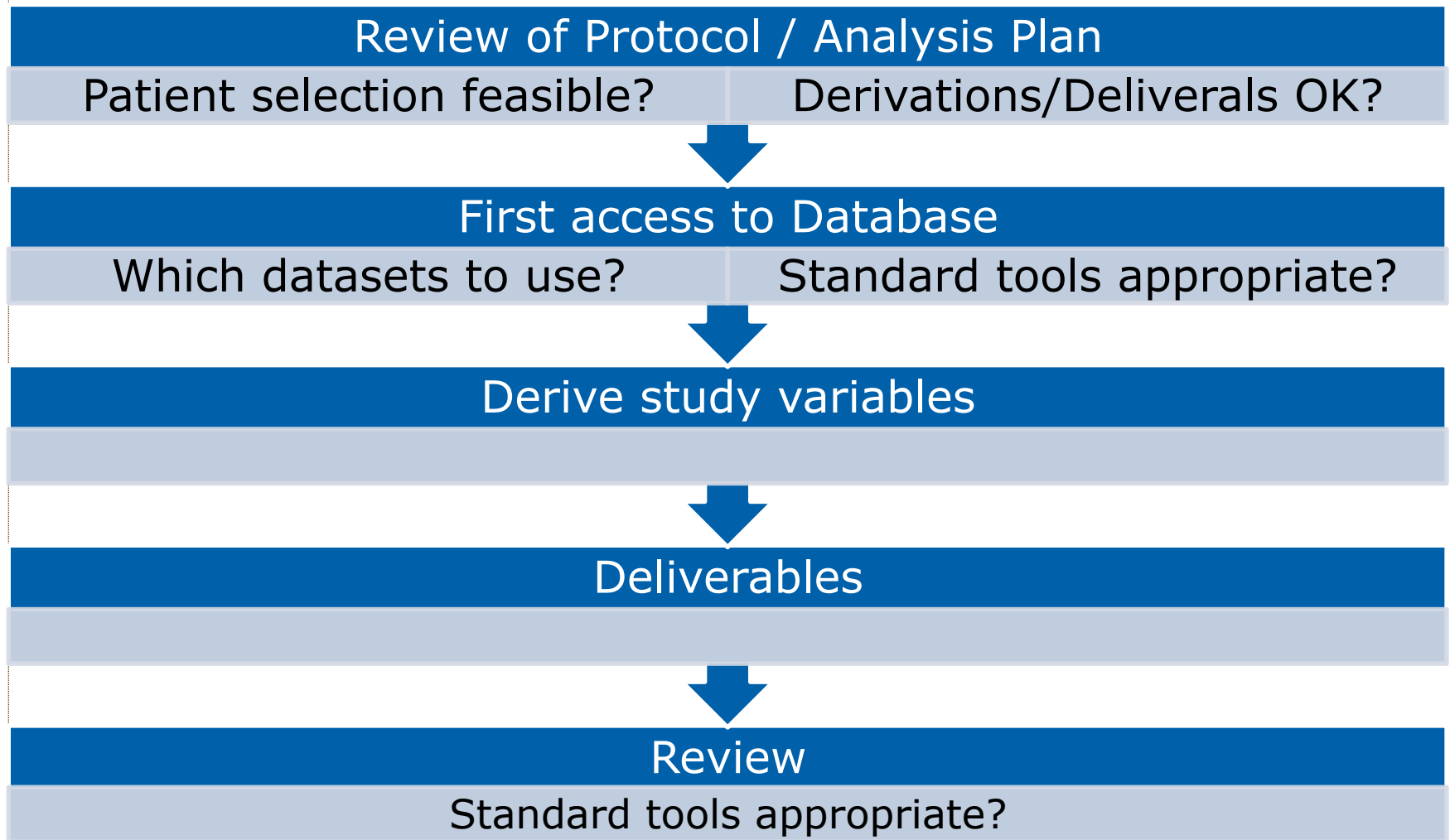


To ensure that you address any issue that occur in one of 1000 observations with a probability of 0.9 you need 2302 observations. May be not small data. And also no guarantee that you get all things right. But if you want to design a dataset with all possible combinations for a complex program this might become even bigger.

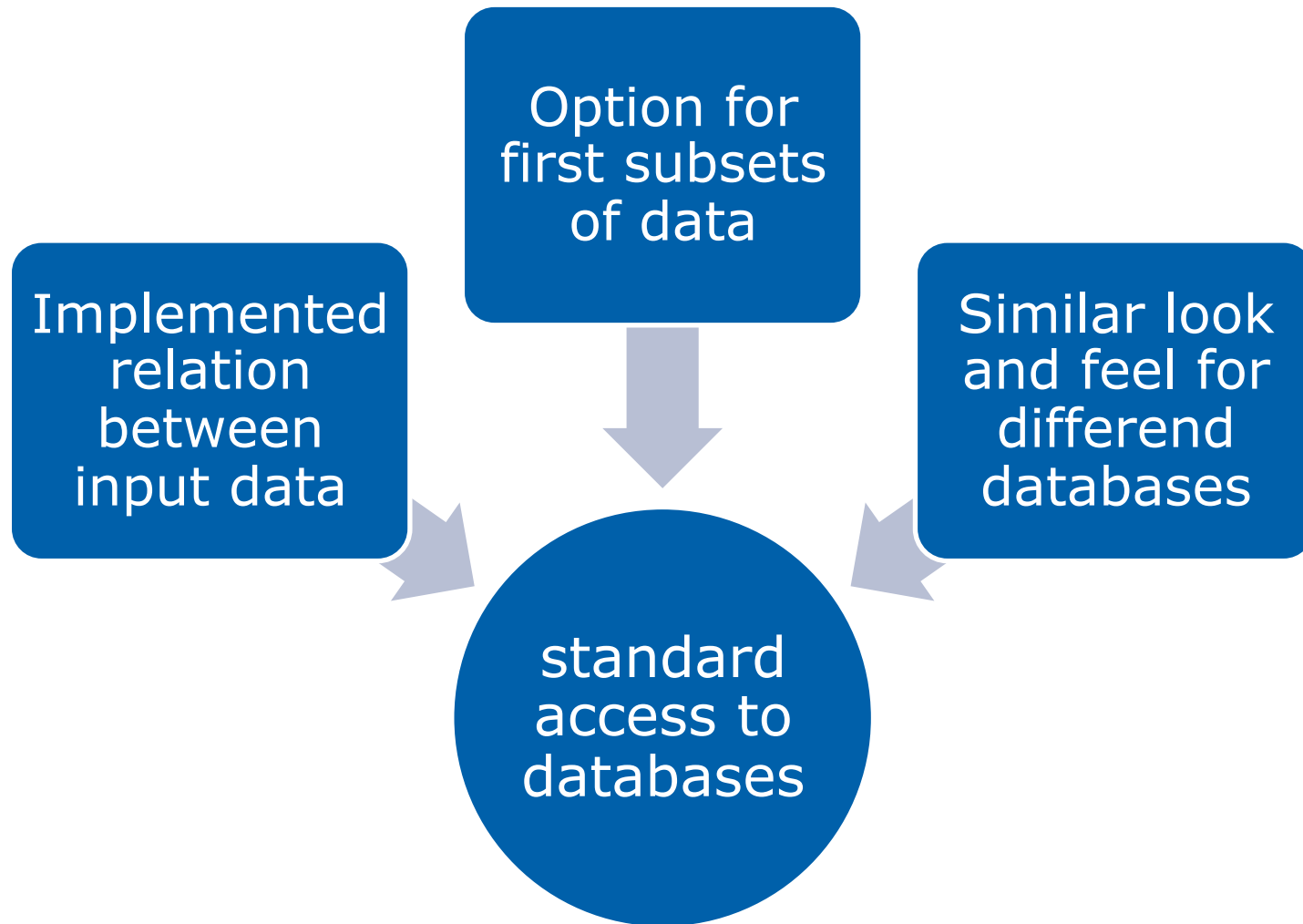
# Proposal claims data: „Standard tools, output“



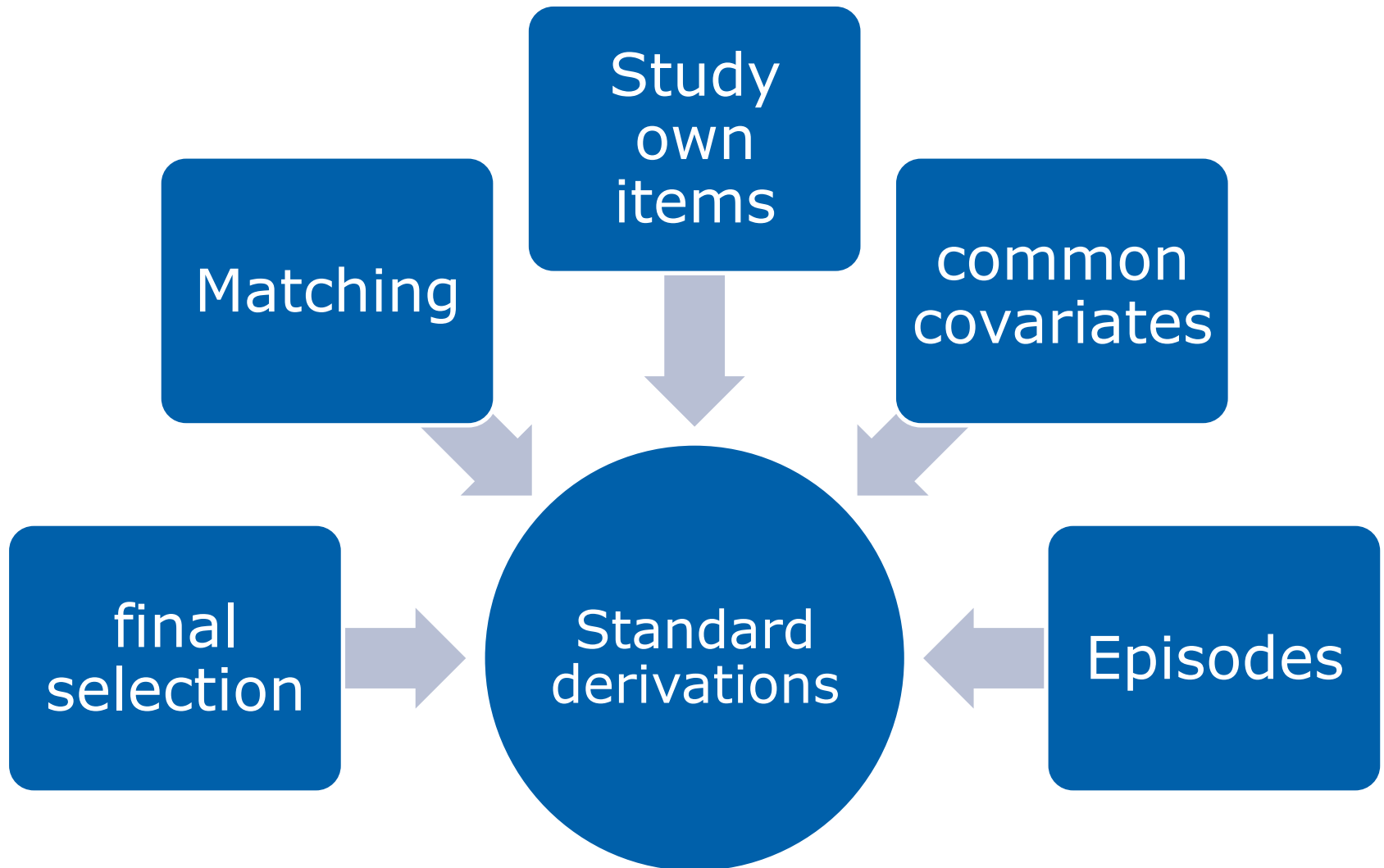
# Proposal claims data : Study Roadmap 0



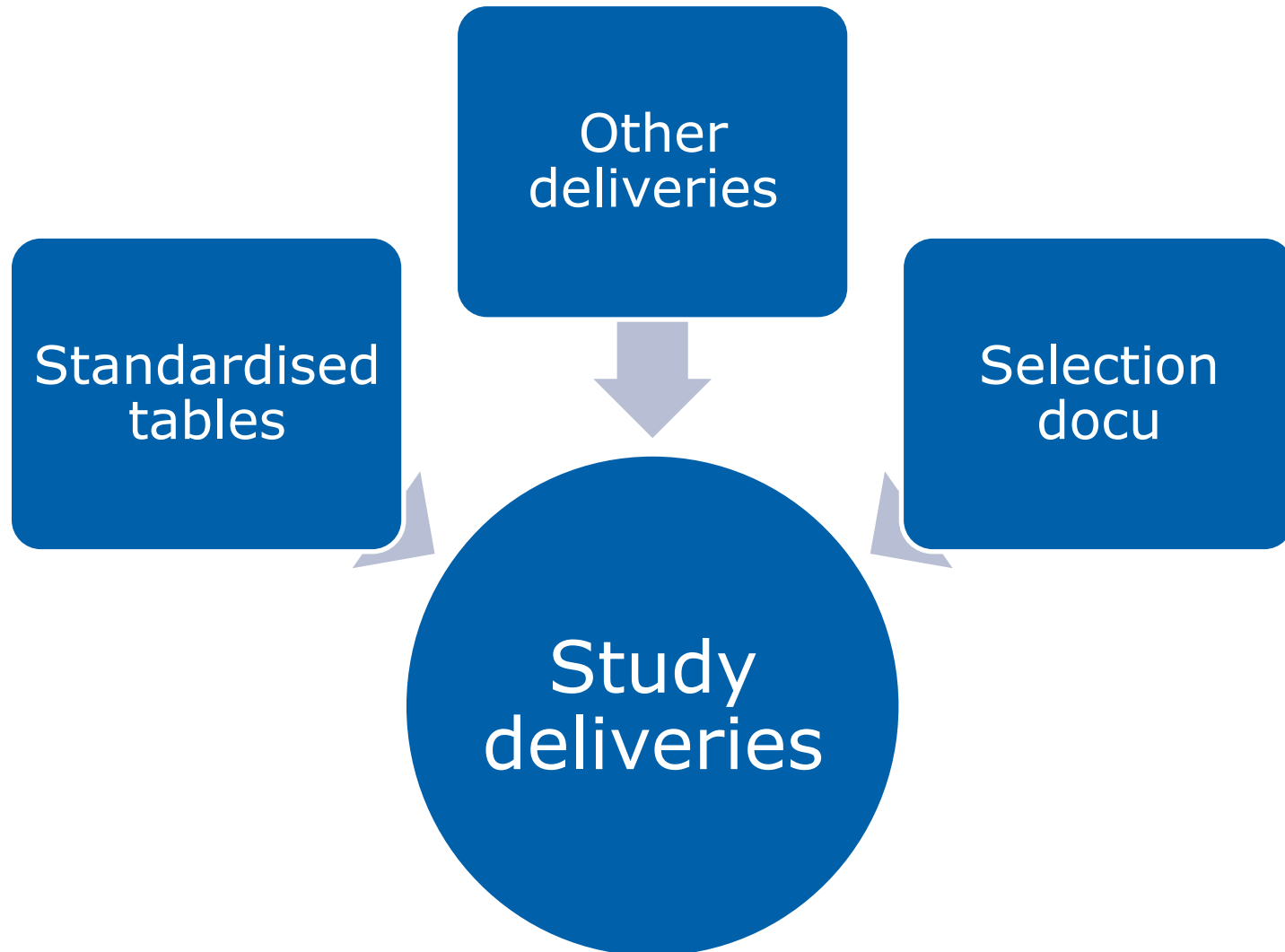
# Proposal claims data: Study Roadmap 1



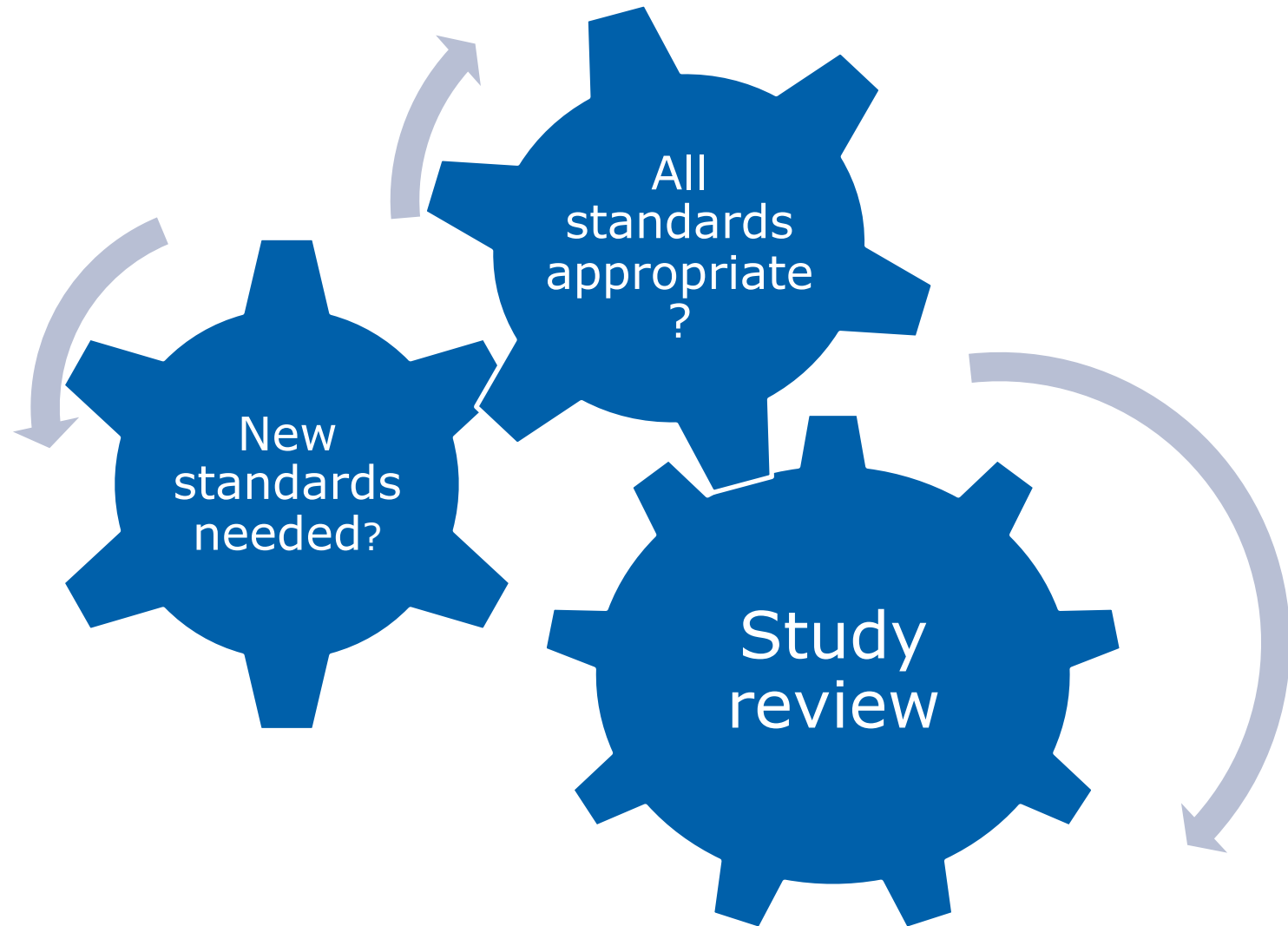
## Proposal claims data: Study Roadmap 2



# Proposal claims data: Study Roadmap 3



# Proposal claims data: Study Roadmap 4



# Questions?



# Disclaimer

23

This presentation is meant for a general audience, and is not intended for healthcare professionals, patients or patients associations.

This presentation includes “forward-looking statements” relating to UCB group of companies (“UCB”) that are subject to known and unknown risks and uncertainties, many of which are outside of UCB’s control and are difficult to predict, that may cause actual results to differ materially from any future results expressed or implied from the forward-looking statements. In this presentation, the words “anticipates,” “believes,” “estimates,” “seeks,” “expects,” “plans,” “intends” and similar expressions, as they relate to UCB, are intended to identify forward-looking statements. Important factors that could cause actual results to differ materially from such expectations include, without limitation: the inability to obtain necessary regulatory approvals or to obtain them on acceptable terms; the economic environment of the industries in which UCB operates; costs associated with research and development; changes in the prospects for products in the pipeline or under development by UCB; dependence on the existing management of UCB; changes or uncertainties in tax laws or the administration of such laws; changes or uncertainties in the laws or regulations applicable to the markets in which UCB operates. All written and oral forward-looking statements attributable to UCB or persons acting on its behalf are expressly qualified in their entirety by the cautionary statements above. UCB does not intend, or undertake any obligation, to update these forward-looking statements.

■

“If the whole materia medica, as now used, could be sunk to the bottom of the sea, it would be all the better for mankind, and all the worse for the fishes.”

Oliver Wendell Holmes

Medical Essays, “Comments and Counter”

Currents in Medical Science