EUROPEAN MEDICINES AGENCY
SCIENCE MEDICINES HEALTH

Report from the breakout session 5: Heterogeneity of results between data sources

Learnings Initiative webinar for Optimal Use of Big Data for Regulatory Purpose

30 November 2021

Presented by Daniel Prieto-Alhambra

An agency of the European Union

Is heterogeneity between results of database studies a **common feature in pharmacoepidemiology**? What are possible **explanations** for such heterogeneity? Can heterogeneity be a **source of knowledge**?

- **SUMMARY: the search for "wanted" heterogeneity**
- Four types of heterogeneity exist: measurement heterogeneity, information heterogeneity (different measurement/s availability, granularity, etc), and true heterogeneity. And a fourth one: "Interpretation heterogeneity" (e.g., statistical significance as a metric for trial emulation).
- The importance to come up with a set of metrics to measure database heterogeneity.
- It would also be interesting to have a set of metrics to measure consistent results**.**
- Heterogeneity is a source of knowledge and can be used to our advantage. We need different types of data to learn on this. We can use this to triangulate evidence, as confounding structure can be different.
- We want to see the complexity of different settings in RWE, and we should appreciate heterogeneity.

Classified as public by the European Medicines Agency

Can heterogeneity between data sources be **anticipated** when considering use of a range of databases, e.g. through set of **standard indicators (metadata) or other means**?

- **SUMMARY: Metadata could be necessary but not sufficient. Additional tools are needed.**
- The risk of heterogeneity being interpreted as lack of reliability/bias etc, especially in the past
- Metadata can be useful to anticipate heterogeneity but cannot solve the issue completely
- Potential value of phenotype libraries to identify important variables in different databases
- Value of cohort diagnostics/analytical diagnostics/summary scores (e.g. PS) before conducting a study.
- The importance of documenting DQA assessments and periodicity

Classified as public by the European Medicines Agency

What are possible remedies to **attenuate heterogeneity** at the stage of **study design**, for example through restriction of study population, exposure and outcomes to the minimum required?

- **SUMMARY: Restriction is too costly. Alternative tools exist and should be used more widely**
- Restriction, while appropriate for certain questions, comes at a very high cost in RWE as it sacrifices 'external validity'
- More useful solutions (in our opinion) could include:
  - the use of metadata to identify 'fit for purpose' databases for a specific RQ
  - the use of negative control outcomes to address systematic error, residual confounding, etc
- Study design choices are key to minimise "unwanted" heterogeneity. We discussed the value of visualization tools to prevent unwanted heterogeneity.

Classified as public by the European Medicines Agency

How can heterogeneity between databases be **analysed** and **interpreted** in the context of **regulatory evaluations**? Is there a place for **(meta-)analytical techniques**? Under which conditions?

- **SUMMARY: Meta-analyses carry challenges. It is essential to report database-specific results.**
- Visualization tools can be used to identify differences in study design and to interpret heterogeneity
- Important to report all estimates, not only a meta-analytical one by using forest plots
- Interpretation is very important as RWE should provide heterogeneous findings that are true/reliable
- The meta-analysis can discard useful data, by requesting them to be homogenous.
- Small cell count issue. The potential problem of analyses with 0 outcomes which probably means "safe" but we drop such databases for mathematical reasons when meta-analysing. Maental-Hanszel approach and some novel analytical/IT solutions e.g., synthetic data has been used before as possible solutions.
- We recognized the public health value of access to patient-level data for these specific studies.

Classified as public by the European Medicines Agency