EUROPEAN MEDICINES AGENCY
SCIENCE MEDICINES HEALTH

A Regulatory Approach to Validation of the CDM

Structuring a CDM to improve validity of analyses.

Common data model Workshop in Europe - 11-12 December 2017

Requirements for regulatory evidence

Provide agreed, transparent decision process

- Industry
  - Need clear study design criteria
- Regulators
  - Need clear assessment criteria
- Public
  - Need to understand and trust process

Scientifically sound

Clinically convincing

Evidence generation processes verifiable

WHY HAVE PRESPECIFIED, PROSPECTIVE INTERVENTIONAL STUDIES AND, IN PARTICULAR, RCTS BEEN PREFERRED TO RETROSPECTIVE OBSERVATIONAL STUDIES IN REGULATORY DECISION MAKING?

Variable measurement

Interventional study

Pre-specified uniform measurement criteria

Pre-specified timing relative to treatment allocation

Pre-specified interpretation of measurements

Dedicated expert review temporally close to measurement if required

Observational data

Variable must be inferred from a range of unsystematically recorded observations

Timing is not controlled

Variation across doctors in preferred codes and propensity to record.

Allocation of treatment

Interventional study

Study group selected – good?

Use of placebo an option

Randomisation usual to ensure allocation independent of patient characteristics

Use of treatment can be checked

Patients and clinical staff are aware that they are involved in a study

Observational data

Treatment given selectively according to perceived patient need

Not always clear that the prescribed treatment has been taken or for how long

Data validation

Interventional study

Standardised forms and trial management procedures check timing and completeness of data

Major errors and omissions are queried

Monitoring mandated

External inspections can be implemented

Observational data

Some data collection systems will facilitate checks against medical notes

Some statistical checks may be run but remedial measures are usually crude – EG exclusion of whole practices

Because the studies are not specified at the time of data collection, concurrent validation cannot be focused.

Specification of analytical procedures

Interventional study

Success criteria can be specified prior to data collection

- In an RCT these can be quite simple

Analysis specified prior to data collection

Time points/numbers of patients required pre-specified

Control of multiplicity

Observational data

Best practice requires a formal protocol, stating success criteria, which should be prepared without prior looks at the outcomes*treatment interaction.

Analysis should be described in detail but lack of control over data collection adds complexity.

Decisions based on results always require post-hoc assessment of credible bias.

Impact in terms of regulatory criteria

Interventional study (RCT)

Levels of evidence can be specified in terms of effect sizes and measures of statistical significance

We believe that current regulatory and company monitoring processes make deliberate malpractice difficult.

Observational data

In addition to statistical measures we must evaluate the extent to which we trust the results.

Even without deliberate malpractice many aspects of data quality and study design need to be assessed

With current research environment, malpractice is not difficult

The current state of science is arguably very poor. For medical observational studies over 80% of initial claims failed to replicate, Ioannidis, JAMA, 2005, Young and Karr, Significance, 2011.

Scientific fraud is common in retracted science papers, Fang et al., PNAS, 2012. So the evidence is that science claims usually fail to replicate and that fraud is being committed.

Promoting transparency is a key to solving problems of validity and integrity.

Stan Young

HOW COULD A VALIDATED COMMON DATA MODEL HELP TO CHANGE THE BALANCE ?

Possible CDM “package”

Pre-specification of a selection of data sources

Level of coding detail

- High detail, possibly hierarchical systems to accommodate a wide variety of studies
- Only major details to allow broad-brush epidemiology

Concurrent validation of data

- Credible differences between databases
- Checks against national statistics

Possible CDM “package”

Automated recording of analyses

- What was done
- Versions of data-bases

Gatekeeper role?

- Some data ONLY accessible via CDM

Variable measurement

Many useful concepts pre-defined

- An important reduction in multiplicity concerns

With pre-specification of selected databases, a potential to standardise some recording practices across databases

Allocation of treatment

Unchanged

Data validation

Validation checks can be pre-specified and periodically repeated

Pre-specification of databases would allow monitoring of coding procedures and completeness

Established validity – or lack of validity – can be taken into account in specifying analyses.

Specification of analytical procedures

If the system records analyses this adds a level of verification

- Can protect against inappropriate prior inspection of data
- Can allow exact replication and additional sensitivity analyses

Regulatory verification

If system also has a gatekeeper role some data can be reserved for checks

Multiple databases allow assessment of heterogeneity of treatment effects across the health systems represented

Conclusions

A carefully designed CDM has many attractions from a regulatory point of view

Cannot solve all the challenges of observational data analysis but can provide an environment that limits some of the potential sources of bias and facilitates verification

THE END

Any questions?

European Medicines Agency
30 Churchill Place • Canary Wharf • London E14 5EU • United Kingdom
Telephone +44 (0)20 3660 6000 **Facsimile** +44 (0)20 3660 5555
Send a question via our website www.ema.europa.eu/contact

Follow us on **@EMA_News**