What Have I Done?
Analysis Specification
Document for Retrospective Database studies
Disclaimer: The opinions expressed in this presentation and on the following slides are solely those of the presenter and not necessarily those of UCB. UCB does not guarantee the accuracy or reliability of the information provided herein.
Agenda

**Introduction:**
- Issue: What Have I done?
- What was done in study XYZ?
- Why and when is this an issue

**Retrospective Database studies – Process and Documentation**
- Stakeholders
- Documents
- Guidelines

**Analysis Specification Document**

**Evaluation and first experiences**
Issue: What has been done?

When is this an important question?

- Manuscript needs to be written based on a completed study
- Request to do additional analysis on study XYZ
- Request to repeat study XYZ with a different population
- Stakeholders ask for additional detail on a completed study

Why can this be a difficult question?

- Study has been completed 2 or more years ago
- Programmers are no longer available
- Detailed Documentation is spread over several documents or email chains
Retrospective database study

Observational study on retrospective data
Data is already collected
→ data is as is

Data was not collected for the purpose of the analysis
→ assumptions need to be made
→ variables of interest need to be derived

Real World Data
→ Data may be incomplete
→ Data may contain inconsistencies
Real World Data

- claims databases
- medical charts
- observational trials
- pharmacy data
- electronic health records
- registries
- patient surveys
- social media
Real World Data

- claims databases
- medical charts
- observational trials
- pharmacy data
- social media
- electronic health records
- registries
- patient surveys
Retrospective Database studies

Process and study team / stakeholders

- Requestor
- Study Owner, Epidemiologist, Health Economist
- Programmers

visionary/strategist

architect

protocol

builder

programming
Retrospective Database studies

Process flow and documentation

source data → program → derived data → outputs

protocol

DDT

TFL shells

Final report
Issue: What has been done?

How to answer this question?

- Study Protocol
- TFL shells

What am I gonna do?

- Data Definition Tables
- SAS Programs
- Tables and Figures
- Study Report

What was done and how was it done
Retrospective Database studies

Process flow and documentation

- source data
- program
- derived data
- outputs

- protocol
- TFL shells
- DDT

- Additional information that is necessary is spread over several documents
- Details are missing from the protocol
- Protocol may need to get updated after it's final

Final report
Analysis Specifications Document (ASD)

- Study Design
- Population
  - Detailed eligibility criteria
  - Detailed inclusion-/exclusion criteria
- Analysis periods and analysis timepoints
  - Study period, index date, baseline period, follow up period, ...
- Study variables
  - Detailed derivation rules (incl. order of application)
  - Examples in case of complex derivation rules
- Statistical Analysis
  - Detailed Description of procedures
  - Description of outcome of iterative or data-driven steps
  - Capture deviations from protocol after protocol is final
Analysis Specifications Document (ASD)

- **Study Design**

- **Population**
  - Detailed eligibility criteria
  - Detailed inclusion-/exclusion criteria

- **Analysis periods and analysis timepoints**
  - Study period, index date, baseline period, follow up period, ...

- **Study variables**
  - Detailed derivation rules (incl. order of application)
  - Examples in case of complex derivation rules

- **Statistical Analysis**
  - Detailed Description of procedures
  - Description of outcome of iterative or data-driven steps
  - **Capture deviations from protocol after protocol is final**
Analysis Specifications Document (ASD)

- **Study Design**
- **Population**
  - Detailed eligibility criteria
  - Detailed inclusion-exclusion criteria
- **Analysis periods**
  - Study period, index date, baseline period, follow up period, ...
- **Study variables**
  - Detailed derivation rules (incl. order of application)
  - Examples in case of complex derivation rules
- **Statistical Analysis**
  - Detailed Description of procedures
  - Description of outcome of iterative or data-driven steps
  - **Capture deviations from protocol after protocol is final**
Analysis Specifications Document (ASD)

Process Flow

source data

protocol

ASD + TFL shells

prescriptive
Analysis Specifications Document (ASD)

Process Flow

- source data
- program
- derived data

- protocol
- DDT
- ASD + TFL shells
- living document
Analysis Specifications Document (ASD)

Process Flow:
- Source data
- Program
- Derived data
- Outputs

Protocol

DDT

Final report

ASD + TFL shells

descriptive
Retrospective Database studies

Process and study team

- Requestor
- Study Owner, Epidemiologist, Health Economist
- Programmers
- visionary/strategist
- architect
- builder
- protocol
- programming
Retrospective Database studies

Process and stakeholders

- **Requestor**
  - Visionary/strategist

- **Study Owner, Epidemiologist, Health Economist**
  - Architect
  - Protocol
  - Engineer
  - Specifications
  - Builder
  - Programming

- **Observational Data Analysts**
First Experiences

Additional effort and time involved
Additional review necessary
Documents need to be checked against each other for consistency

Late additional requests can be handled much easier
Helpful for manuscripts and other publications
Collects all information that otherwise is spread in e-mail communication
Reference document for future studies
First Experiences

Study Owners
• Less documentation on their end (+)
• Less work with TFL shells (+)

Observational Data Analysts/Programmers
• Additional Work (-)
• Enhanced Role (+)
• More visibility (+)

Better mutual understanding (+)
Conclusions

ASD is helpful for ...
- complex studies
- studies that may be used as pilot studies
- complex feasibility assessments

ASD facilitates ...
- successor studies
- the identification of areas of standardization
- the collaboration between departments

Small, easy and one-off request may not need an ASD
A good protocol may be sufficient

Trade-off between additional effort for current study and possible blueprint for future studies
Questions?
Thanks!
Documentation

Protocol Checklists

Good Pharmacoepidemiological Practice (GPP)
Guideline

ENCEPP Checklist

STROBE: STrengthening the Reporting of OBservational studies in Epidemiology Checklist

ISPOR Checklist

Is that sufficient?
Is more specification too much for the protocol?
Guidelines

ISAC guidelines

Independent Scientific Advisory Committee for MHRA database research (ISAC)

Guidance Notes: Content of CPRD ISAC Research Protocols

p16.

If it is not possible at the time of the ISAC application to provide operational definitions of exposures and/or outcomes because these will be elucidated during the course of the study, an acceptable alternative is to describe the process by which these definitions will be reached.

Is that sufficient?
Is more specification too much for the protocol?
Analysis Specification Document

Examples

Based 'largely' on treatment line algorithm from EPD091, EPD116, and EPD127
Analysis Specification Document

Examples

Example 1

Example 2

- code for obesity
- code for glucose intolerance
- code for hypertension
- code for hypertriglyceridemia
- code for 277.7

diagnosis date for metabolic syndrome disorder
3.6.1.7 Notes on Health Care Resource Utilisation

For each date where a patient has a claim we look at the type of files contributing on the date.

Hospitalizations: If a date appears in any of the I (Inpatient), S (Inpatient Services) or F (Facility) files with non-missing CASEID then we consider a hospitalization to have occurred on this date. ED visits that result in a hospitalization are included as hospitalizations.

ED Visit: If a date has SVCSCAT ending in 20 and no hospitalizations on this date. (This is definition recommended by Marketscan)

Outpatient Visit: Any other dates from the O (Outpatient File) file or F (Facility) file with CASEID missing which are not hospitalizations or ED Visit.

O (Outpatient File) records or F (Facility File, with CASEID missing) records which have the variable STDPLAC=21 are removed prior to applying the above classifications, as they may represent inpatient services, because the claim was not incorporated into an inpatient admission (e.g., no room and board charge was found). Also see MarketScan User Guide.