

Experiences and strategies in Linking Real-World Data from different sources

Nov-2018

Agenda

- Background
- Example 1: linkage of clinical study data with outpatient pharmacy data
- Example 2: linkage of GP data with hospitalization data
- Probabilistic linkage: Pharmo database network

Data linkage

- Exact Deterministic
- Iterative Deterministic
- Probabilistic

Steps in the linkage process

- Choose identifiers
- Select and upgrade the patient data
- Blocking / Exact match
- Actual linkage process
- Validate results

Privacy

- Trusted third party (TTP): STIZON
- Proportional
- Pseudonimization



Example 1: Linking a clinical study to outpatient pharmacy data

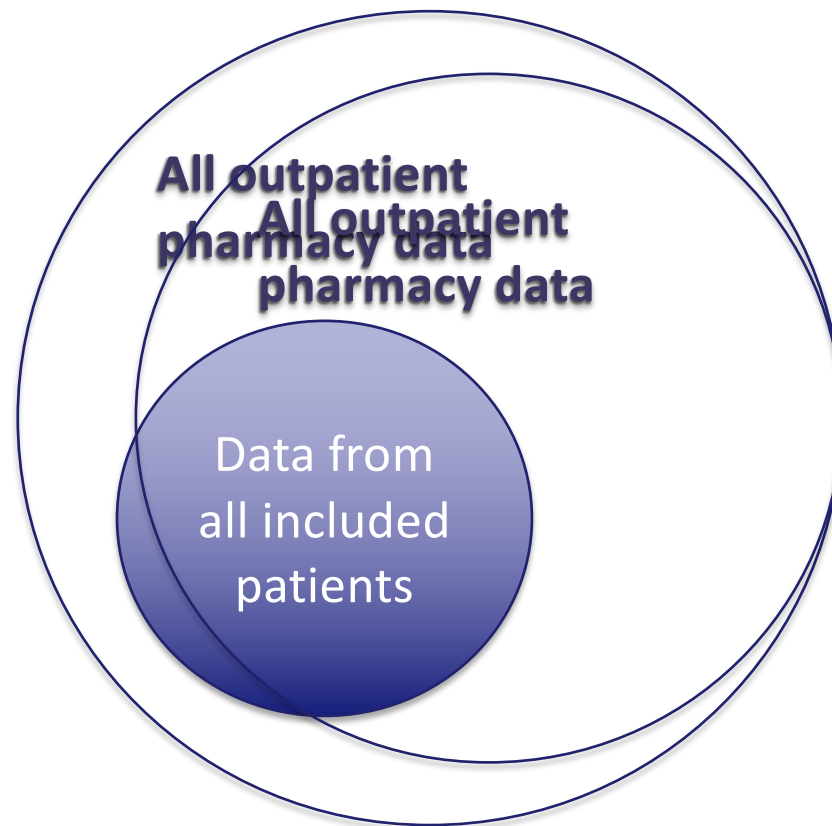
- Context
 - Philips research, IQ Health Care and Radboud UMC
 - Adherence control engine
 - Patients starting on oral antidiabetic treatment, cardiovascular treatment
 - Recruitment via pharmacies
 - Medication adherence scores to be linked to other study results

Example 1: Linking a clinical study to outpatient pharmacy data

- Iterative deterministic linkage
- Identifiers
 - Social security number, gender, birth date
 - GP
 - Study inclusion vs start of new treatment
 - Indicated study treatment versus started treatment

Example 1: Linking a clinical study to outpatient pharmacy data

- Data selection

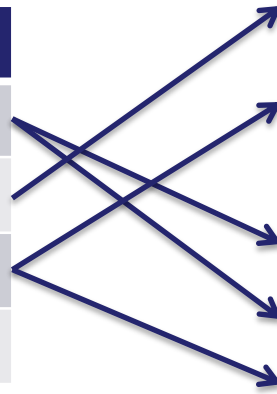


Example 1: Linking a clinical study to outpatient pharmacy data

- Blocking
 - social security number
 - gender and birth date

Patient	SocNo	Gender	Birth date
101	328312	M	31/01/1947
102	123467	F	13/05/1950
103	139091	M	17/08/1972
104	232121	F	29/11/1964

Patient	SocNo	Gender	Birth date
365		M	13/05/1950
791		F	13/05/1950
011	139091	M	17/09/1972
231		F	29/11/1965
871	328312	M	31/01/1947
332		M	31/01/1947
126		M	17/08/1972



Example 1: Linking a clinical study to outpatient pharmacy data

- Blocking

```
PROC SQL;  
    CREATE TABLE MyMatch as  
        SELECT a.*, b.* FROM MyPharmacyPatients as a  
            LEFT OUTER JOIN MyStudypatients as b  
                ON (a.birthdate=b.birthdate and  
                    a.gender = b.gender)  
                OR (a.SocNo=b.SocNo and a.SocNo <> .)  
        ORDER BY a.PatIdPh, b.PatIdStud;  
QUIT;
```

Example 1: Linking a clinical study to outpatient pharmacy data

- Define match compatibility for each possible match:
 - Social security number overrules everything
 - Difference between inclusion date and start relevant treatment
 - 0-15 days, 15-30 days, 30-60 days
 - Corresponding GP?
 - Started on correct treatment?

Example 1: Linking a clinical study to outpatient pharmacy data

- Iterative linking
 - Take match with highest score

```
PROC SORT DATA=mymatch3;  
    BY PatIdStud DESCENDING matchOk;  
RUN;
```

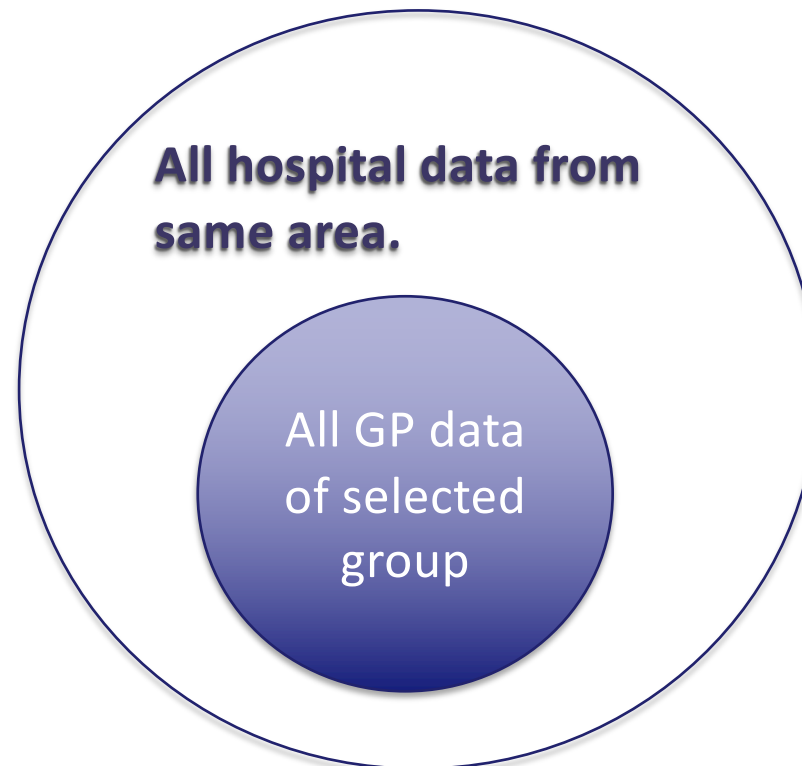
```
DATA MyMatch4;  
    SET MyMatch3;  
    BY PatIdStud DESCENDING matchOk;  
    IF FIRST.patIdStud THEN OUTPUT;  
    KEEP patIdStud patIdPh;  
RUN;
```

Example 1: Linking a clinical study to outpatient pharmacy data

- Results: match 89% of study patients
- Issues during the matching process
 - Study patients without linked data
 - Time window: when is inclusion date corresponding to medication start date
 - Medication differences
 - Patients switching pharmacies

Example 2: Exact deterministic linkage

- Goal is to improve registration of CVR patients
- Hospital database linked to GP database
- Identifiers
 - Gender, birth date,
 - Postal code,
 - GP
- Data selection



Example 2: Exact deterministic linkage

- Match on all variables excluding GP
- Selected right GP matches based on initial matching results
- Match again on all variables including GP related to the practice

Example 2: Exact deterministic linkage

- Examination of results
 - 184159 (39%) of patients matched with hospital database.
 - Relatively more children and elderly patients matched
 - Relatively more women matched (53%)
 - Two practices not matched

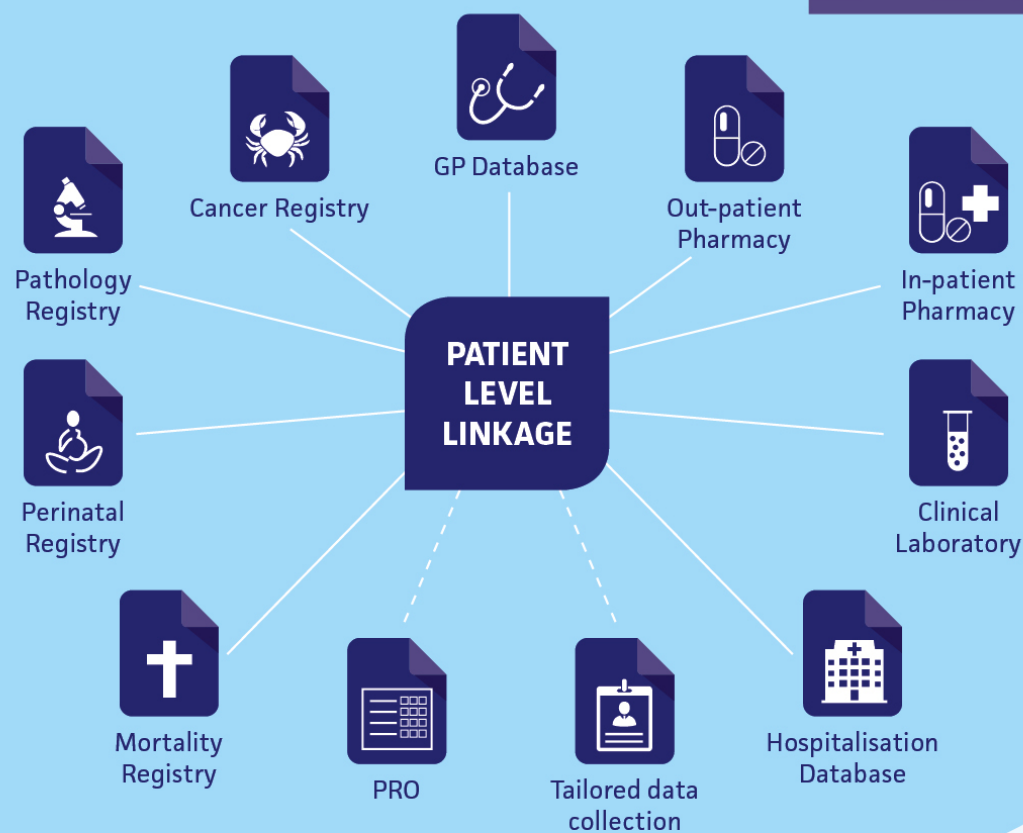
Probabilistic matching

- Outpatient pharmacy data vs hospitalization
- No social security number available
 - Need for alternative logical related identifiers
- Based on bayesian statistics
- Recurrent matching process
- Ron M.C. Herings: A Record linkage system for postmarketing surveillance of prescription drugs in The Netherlands

PHARMO Database Network (NL)

PHARMO Database Network

Rapid access to primary and secondary healthcare data



Application

- Post authorisation safety studies (PASS)
- Real World effectiveness studies
- Market access studies
- Clinical practice
 - Improve registration
 - Identifying special interest groups

Selection of matching method

- Number of possible identifiers
- Quality of the data
- Single vs recurrent matching
- Purpose of the results

Conclusion

- There is not a one fits all solution
- Method depends on the purpose of the linkage
- Keep reviewing and checking individual matching results to understand the strength and the possible biases

WWW.PHARMO.COM

Contact details

T +31 (0)30 744 08 00

E pharmo@pharmo.com

