Use Cases – Advanced Dataset Validation

Case 1 – Undead Prescriptions

1. Data user wants to know what proportion of Prescribing records were collected after the patient died or before he was born

Case 2 – Standardize Directions in Prescribing

1. Data analyst wants to know what proportion of prescribing directions do not follow syntax rule (e.g. 1/d = take 1 a day)
2. Data analyst wants to write a rule that transforms common mistakes or mis-ordered directions into correct standardized directions format.
3. Data user wants to know the daily dosage rate of any given drug based on the drug (res seq no) and the directions given to the patient. (requires machine parsing of directions)

Case 3 – Multiple Mortality

1. Data user wants to know how many people have multiple dates of death.
2. Data user wants to know which date of death is ‘correct’

Case 4 – Multiple Diagnosis (of Diabetes)

1. Data user wants to know earliest date a person has been diagnosed with diabetes
2. Data user wants to know how many prescriptions for diabetes related drugs exist prior to this date (for each patient with diabetes).

Case 5 – GP Practice

1. Data analyst wants to know how many GP Practices are not in our lookup
2. Data analyst wants to know how many GP Codes in CHI are not in our GP lookup
3. Data user wants to know how many GPs are not part of a practice
4. Data user wants to know how many people do not have a known GP Practice

Case 6 – Time Series CHI

1. Data user wants to validate the record of stay in healthboards against prescribing
2. What proportion of records have a date\_out / date\_in that does not correspond to a change in prescriptions collected per year/month

Case 7 – Res Seq No not found

1. Data analyst wants to know what proportion of ResSeqNos are not in our prescribing item