Linking Patients in PDMP Data

Kentucky All Schedule Prescription Electronic Reporting (KASPER)

PDMP Training & Technical Assistance Center Webinar
October 15, 2014

Jean Hall
Lindsey Pierson
Office of Administrative and Technology Services
Kentucky Cabinet for Health and Family Services
Standardizing Prescription Data

• Create standardized secondary data fields for all fields addressed in the process. These fields are used for clustering.

• Pt ID – This field is the patient SSN or Driver’s license
  – All upper case
  – Remove all non-alphanumeric characters
  – Compare to known invalid SSNs
  – Additional SSN and DL logic applied. Pt ID will be set to null under certain circumstances.
Standardizing Prescription Data

• Last Name
  – All uppercase
  – Set to null if last name is NOLASTNAME, DUMMY or TEST
  – Replace all non-alphanumeric characters with a blank space
  – Remove spaces at beginning and end
  – Only characters up to the first space are considered to be the last name
Standardizing Prescription Data

• First Name
  – All uppercase
  – Set to null if first name is NOMIDDLENAMESPACE, DUMMY, TEST or NOFIRSTNAME
  – Replace all non-alphanumeric characters with a blank space
  – All characters prior to the first space are set to First Name and all characters after the first space are set to Middle Name.
Standardizing Prescription Data

• Address
  – We use a vendor proprietary process to verify patient addresses against the United States Postal Service address database.
  – If a match is found, the secondary fields are updated with the address data found in the USPS database (incorrect data is updated with correct data)
  – If no match is found, the secondary fields are updated with the original data and a flag indicating:
    • No match
    • City was blank or could not be found
    • Partial street data or could not be found
    • Multiple matches were found
Standardizing Prescription Data

• Date of Birth
  – Date of Birth is set to null if it is:
    • Greater than Date Filled
    • Greater than today’s date
    • Is older than 01/01/1600

  – Date of Birth is older than 1/1/1900, we set the century to 19.

• Geocoding
  – The geocoding process that we use is a vendor proprietary process that returns a longitude and latitude based on the patient zip code
As new records are inserted into our prescription table and after standardization is complete, patient data is clustered and a cluster ID is assigned to records that are determined to be for the same patient.

Some data matching is based on an exact match and others on a weighted match. In the illustration that follows,

- * Indicates an exact match
- (85) is not percentage, it is a weighted algorithm proprietary to the vendor. For the name fields this weighted logic incorporates common nicknames, gender, misspellings, regional spelling and other proprietary statistical matching. Addresses that are weighted are standardized using USPS integrated logic.
Clustering Prescription Data

Initial Rules Set

- #1 SSN* & First Name (85)
- #2 SSN* & DOB*
- #3 Last Name (85), First Name (85), & DOB*
- #4 Last Name (85), First Name (85), Street Address (85) & Zip Code*
- #5 DOB*, Street Address (85) & 5 digit Zip Code*

Current Rules Set

- #1 SSN*, Last Name (85) & First Name (85)
- #2 Last Name*, First Name (85), Address (85) & 5 digit Zip Code*
- #3 Last Name*, First Name (85), & DOB*
- #4 SSN* & DOB*

93% Auto-Match

97% Auto-Match
Clustering Prescription Data

Changed logic to reduce blended patients:

• #1 - Poor SSN Quality combined with weighted First Name resulted in higher rates of blended patients, added Last Name.

• #2/#3 - Tightened logic to use exact Last Name to prevent blended patients with similar Last Names.

• #5 – Incorporated this rule into our request process to help identify blended patients.
Report Request fields are standardized based on the same criteria as prescription data:

- Patient ID (SSN)
- Last Name
- First Name
- Street Address
- Zip Code
- Date of Birth
KASPER standardized report requests are compared to the standardized and clustered prescription data based on the following rules:

- SSN, Last Name (85) & First Name (85)
- Last Name*, First Name (85), Address (85) & 5 Digit Zip Code
- Last Name*, First Name (85) & DOB
- SSN & DOB
- DOB, Address (85) & 5 Digit Zip Code
Thank You!

Jean Hall
502-564-0105 ext. 2499
Jean.Hall@ky.gov

Lindsey Pierson
502-564-0105 ext. 2646
Lindsey.Pierson@ky.gov
PATIENT LINKING

Chad Garner
Director of OARRS
Ohio State Board of Pharmacy
WHY DO WE NEED PATIENT LINKING?

- Patients use multiple pharmacies
- Patients move, therefore information changes
- Patients get married (name changes)
- Parents lack creativity when naming children
- No standardization for entering addresses
- Data entry errors (spelling, inverted digits, typos)
- Drug seekers use aliases
- **The Birthday Problem**
  - With just 23 people selected at random, there is a 50% chance that 2 of them share the same birthday (not the same year)
  - With 70 people, the chance is 99.9%
- **Popular Surnames**
  - 1% of US Population named Smith
- **Given Names Cycle by Year in Popularity**
  - 2.2% of US Population born in 1961
  - 1.6% of people born in 1961 named Michael
- **Ohio Population 11.5 million**
- **83% chance of 2 Michael Smiths born on same day in 1961**
- **Verified: 2 patients named Kim Smith with same DOB living 10 miles apart**
WHAT INFORMATION IS AVAILABLE?

Incoming Rx Record
- First Name (required)
- Last Name (required)
- Date of Birth (required)
- Street Address (required)
- City (required)
- State (required)
- Zip Code (required)
- Phone Number (required)
- Gender (required)
- Previous 2 years Rx History

Incoming Rx History Request
- First Name (required)
- Last Name (required)
- Date of Birth (required)
- Street Address (optional)
- City (optional)
- State (required – OH default)
- Zip Code (required)
- Phone Number (optional)
- Gender (required- unknown default)
Key – Dataflux uses soundex, address standardization, nickname databases etc. as appropriate for a type of data to generate a key. Example: John = $$$$$$$CB$$$$$$, Jonathan = $$$$$$$CB$$$$$$

Zip-# - leftmost # digits of the zip code. Example: Zip-2 of 43215 is 43

Computed Gender – Gender assigned by individual’s first name. If 95% of patients with given name in database are female/male, computed gender is female/male. If less than 95% are female/male, computed gender is neutral.

Group – set of patient records that all belong to one real-world person

Manually Grouped Flag – Boolean value – false by default. Set to true when a group is processed manually
STEP 1 – GENERATE KEYS

- First Name Key
- Last Name Key
- Address Key
- Phone Number Key
STEP 2 – PROCESS RULES

**Rule 1**
- Last Name Key
- First Name Key
- Zip-2
- DOB
- Address Key

**Rule 2**
- Last Name Key
- First Name Key
- Zip-2
- DOB
- Address Key

**Rule 3**
- Zip-5
- Address Key
- DOB

**Rule 4**
- Phone Number Key
- First Name Key
- DOB

**Rule 5**
- Last Name Key
- First Name Key
- Common Prescriber
- DOB
Is difference in ages of patients more than 3 years?
- No, no exception
- Yes? Continue

Is the difference in ages of patients more than 16 years AND the month and day match AND the difference in years evenly divisible by 10?
- Yes, no exception (it’s a typo)
- No, Continue

Is the difference in ages of patients more than 16 years, but month or day do NOT match or difference in years not evenly divisible by 10?
- Yes: This is a father and son with the same name. Split the patient records accordingly, if necessary, determine if incoming record is father or son, remove the others from the buckets. No exception.
- No: Mark as exception
STEP 4: EXCEPTIONS – TWINS

- Is the request for someone 18 or younger?
  - Yes – remove any records in Rule 3 or Rule 4 buckets, no exception
  - No – No exception
STEP 5: EXCEPTIONS – COMPUTED GENDER

- Do we have multiple computed genders?
  - Yes: Is the gender on the incoming record “unknown”?
    - Yes – keep all records, no exception
    - No – Keep records where computed gender matches incoming record, no exception
  - No: No exception
Multiple Groups, all are marked as manually reviewed
  - If an exact match is found, add request to that group, processing complete. Otherwise, exception.

Multiple Groups, at least 1 is not marked as manually reviewed
  - Exception

Single Group, no exception
  - Add request to group, processing complete

No groups
  - Send “No patients found” report.

Still have Exception?
  - Add to manual process queue
**STEP 6: INCOMING RX RECORD**

- **Multiple Groups, no exceptions, all manually grouped flags = false**
  - Merge groups into single group, add new patient to this group, processing complete

- **Multiple Groups, no exceptions, all manually grouped flags = true**
  - Add it to a group – it will cause an exception if the patient is requested.

- **Multiple Groups, no exceptions, some manually grouped flags = true**
  - Exception

- **Single group, no exceptions**
  - Add patient to group, processing complete

- **No groups**
  - Clear exceptions, Create a new group, add patient to it, complete

- **Still have an exception?**
  - Create a new group and add patient to it – it will cause an exception if patient is requested.
2006
- 4 rules (no rule 5)
- No exceptions on incoming Rx records
- No manually grouped flag
- 1 exception on requests: DOB range > 3 years
- All exceptions were manually processed – even if the same patient was requested multiple times
- 95% auto-completion rate

2009
- Added rule 5
- Still 95% auto-completion rate, but we started finding patients who moved across the state
HISTORY OF CHANGES

- **2011**
  - 1\textsuperscript{st} mandatory usage bill passed. Manually processing 5\% of requests was already taking 1 FTE.
  - Added ability to detect father/son
  - 97\% auto-completion rate

- **2012**
  - Added Manually Grouped Flag
  - 98\% auto-completion rate initially; gradually grew to current 99.8\% auto-completion rate
  - Added Computed Gender in order to deal with husband/wife appearing on same report
  - Remaining manually processed requests are much more difficult. Often require looking at Rx histories and/or calling a pharmacy or physician treating the patient

- **Future?** May change Zip-2 to a radius search