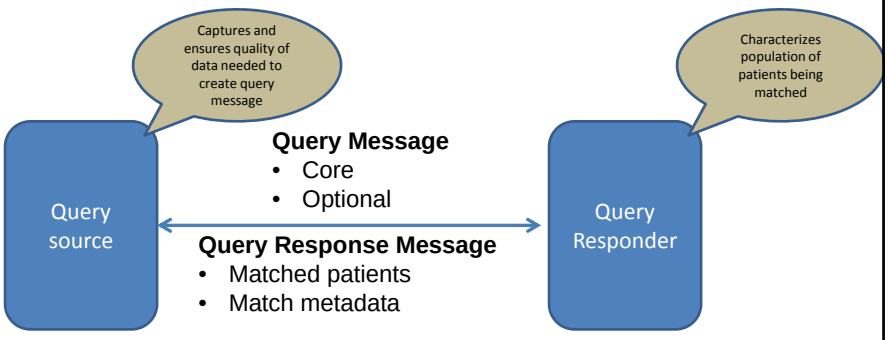
Draft Recommendations

Patient Matching Power Team

July 1, 2011

Principles

- Specificity more critical than sensitivity (need policy committee decision about the rate of false positives that can be tolerated) – we focused on direct patient care
- Make it easy to include additional patient attributes which may prove useful in matching in the future
- Align efforts to improve data with the importance of the data for matching (optimize value) –> getting the data right
- We will not disclose PHI inappropriately



Matching Fields

- Core
  - Name*
  - Date of birth
  - Administrative Gender
  - Zip code
  - Social Security Number -- when it is possible that the last patient contact with the data source was more than one year ago†
- Examples of optional attributes
  - Cell phone (leading indicator)
  - Future cyberidentities
  - Street address, city and state
  - Maiden name
  - Mother's maiden name
  - Other

*Begin capturing and including the full middle name rather than just the initial
†The rational is that US residents move an average of every 5 years

Data Quality

- Registration processes should provide a consistent method to identify missing/unavailable data, approximate values or questionable values at the time of data entry
- Methods to allow the patient to check the entry such as sharing the entry screen, a printed summary or on-line access may identify data quality issues
- Apply basic "edits" to attributes
  - valid dates, no more than six 9s or six 0s in a row is SS# etc
- Apply advanced "edits"
  - Validate that zip code and street address are consistent
  - Validate other attributes such as SS# or DL3

Data Formats / Content

- Follow CDA R2 header format for representing patient attributes
- Adopt the NwHIN patient query implementation guide or the IHE PDQ implementation guide (under discussion)

Match Quality Reporting

- List of matched patients
  - Might return no patients
  - Might return several patients that match*
  - Match confidence level (still under discussion)
- Data about the matching algorithm and process

* Limit what data about the individual is returned? eg SS# last 4 but allow more to be used behind the scenes

Discussion candidates for data about the algorithm and process

- What other field values could you provide that would improve your match?
- What algorithm was used – transparency, could be distributed
- Common identifier flag? – ie if one of the matching parameters is "common" in the matched population, inform the requester
- Population characteristics
- Population size?

Draft letter

1. Brief description of the Use Case/Scenario we are working of (direct patient care)
2. Assumptions
3. Principles
4. Recommendations (four categories?)
   - -- Data Elements
   - -- Query Formats, Content
   - -- Data Matching, definition of confidence levels
   - -- Data Quality
5. Further Policy Needs (from Policy Committee?)
   - -- Definition of a match, a 'perfect' match
   - -- Degree to which False Positives are acceptable