Improving the Reporting of Race, Ethnicity, and Language in California Hospitals

David Zingmond, MD, PhD
The David Geffen School of Medicine at UCLA
November 12, 2013

Objectives and Approach

Overall Objective: To improve the reliability, validity, and completeness of self-reported Race, Ethnicity, and Language in data for patients seen in California hospitals (inpatient, ED, and ambulatory surgery)

1. Pre- and Post- needs assessments through structured surveys to hospital registrars (and others) in California hospitals.
2. Adaptation/development/implementation of training materials
3. Development of revised data auditing rules for evaluating data quality throughout the project and feeding back to hospitals
4. Post-collection data improvement (supplementation and imputation)

IOM Recommendations

Health Care organizations must have data on the race, ethnicity, and language of those they serve in order to identify disparities and to provide high quality care.

Detailed “granular ethnicity” and “language need” data, in addition to the OMB categories, can inform point of care services and resources and assist in improving overall quality and reducing disparities.

REL Data Standards

Measures should be self-reported by patients
• Parents report for children and
• Guardians report for legally incapacitated adults

1. Baseline Survey Results

• Out of 367 general acute care hospitals, 205 hospitals (56%) completed at least one survey
• Responding hospitals were similar to non-responding hospitals in terms of:
  – Size (# beds)
  – Ownership (private v. public)
  – Academic medical center (yes v. no)
  – Urban v. rural
Characteristics of Hospitals’ Patient Populations

- Almost half (49%) of hospitals estimated that 25% or more of their patients are minorities (non-White race/ethnicity)
- 26% of hospitals reported serving a “majority minority” patient population
- 1 in 5 hospitals reported that 25% or more of their patients do not speak English (require an interpreter)

Hospital Data Collection and Auditing Practices

<table>
<thead>
<tr>
<th></th>
<th>Data Collected</th>
<th>Standardized Form Used for Data Collection?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Place of Birth</td>
<td>59%</td>
<td>47%</td>
</tr>
<tr>
<td>Race/Ethnicity</td>
<td>98%</td>
<td>83%</td>
</tr>
<tr>
<td>Language</td>
<td>99%</td>
<td>80%</td>
</tr>
</tbody>
</table>

- 75% of hospitals surveyed reported auditing patient registration information

Strategies to Improve Data Collection

- Ask patients to fill out a standardized form: 63%
- Incorporate questions into existing hospital forms: 84%
- Ask registration/admissions clerks to ask questions at patient’s 1st visit: 90%
- Issue memo to encourage hospital personnel to be sensitive to importance of these data: 62%
- Conduct routine staff training on data collection: 86%
- Develop and enforce hospital-wide policies and procedures on data collection: 81%
- Make FAQs and answers available to admissions staff to address patient questions: 83%

Strategies to Improve Data Collection

- Ask patients to fill out a standardized form: 63%
- Incorporate questions into existing hospital forms: 84%
- Ask registration/admissions clerks to ask questions at patient’s 1st visit: 90%
- Issue memo to encourage hospital personnel to be sensitive to importance of these data: 62%
- Conduct routine staff training on data collection: 86%
- Develop and enforce hospital-wide policies and procedures on data collection: 81%
- Make FAQs and answers available to admissions staff to address patient questions: 83%

2. Improving Data Collection

- Develop materials easy to modify by end users (supplemental to existing resources)
- Target primarily at front-line staff
  - Staff training, script, FAQs, forms (on paper; interview)
- Materials formatted using a (1) face page, (2) instructions for use, and (3) actual forms

Creation and Dissemination of REL Materials

- Guiding principles
  - Ease of use
  - Ease of disseminating
  - Ease of customizing
Dissemination of materials

- Statewide webinars in Oct/Nov 2011 with 300+ total participants
- Creation of a wiki-site accessible to end-users
- Planned placement of materials on the OSHPD and HCUP websites
- **Insufficient funding** for extended directed training of hospitals and their staff

3. Development of revised data auditing rules

- Allow for evaluation of trends in reporting during the study period
- Allow for improvement of existing auditing rules currently in place
### Reporting Standards for OSHPD Data Sets

**INPATIENT**

- **Ethnicity**
  - 1 = Hispanic
  - 2 = Non-Hispanic
  - 3 = Unknown

- **Race**
  - 1 = White
  - 2 = Black
  - 3 = Native American/Eskimo/Aleutian
  - 4 = Asian/Pacific Islander
  - 5 = Other
  - 6 = Unknown

**Emergency Dept & Ambulatory Surgery**

- **Ethnicity**
  - E1 = Hispanic or Latino
  - E2 = Non-Hispanic or Non-Latino
  - 99 = Unknown

- **Race**
  - R1 = American Indian
  - R2 = Asian
  - R3 = Black or African American
  - R4 = Native Hawaiian or Pacific Islander
  - R5 = White
  - R6 = Other Race
  - R9 = Other Race
  - 99 = Unknown

In both cases, we allow a single choice for Ethnicity and a single choice for Race. This differs from the current OMB 1997 standard which provides for multiple selections for Race.

### Standard Data Audits

- **Data checks at the time of submission**
  - Hospitals should not have 100% or 0% in major categories
  - Missing/unknown does not exceed 10% of responses
  - Agreement across admits within the same hospital

- **Rules are context free and use two basic concepts**
  - All hospitals should have some variation in Race/Ethnicity
  - Minimize missing/invalid responses

- **Current auditing rules focus on completeness, but not accuracy**

### Missing Race/Ethnicity

- In CA, few hospitals are deficient based upon rates of missing/unknown RE
- 3.4% unknown race/ethnicity (mean across 349 hospitals reporting discharges in 2009)
  - 17 hospitals > 10% unknown
  - 1 hospital > 20% unknown

- **Completeness does not equal accuracy**

### Auditing: Standards for Comparison

- **Ideal**: comparison to validated patient-level, self-report data
- **Patient-level data**
  - Birth certificates (maternal self-report)
  - Cancer registry (hospital-based, augmented)
  - Cancer Epi Study database (very limited)
  - Death certificates (hospital-based)
  - Medicaid
- **Summary-level (contextual) data**
  - U.S. Census
  - American Community Survey (5 year average for language fluency)

### Assessing the accuracy of REL reporting

1. Identify self-report (gold standard) REL information for comparison to hospital-reported data
2. Compare REL information at the patient-level

### Inpatient data compared to Cancer Registry (& self-report)

<table>
<thead>
<tr>
<th></th>
<th>vs case epi study self-report</th>
<th>vs cancer registry</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Overall agreement</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Race</td>
<td>0.91</td>
<td>0.92</td>
</tr>
<tr>
<td>Ethnicity</td>
<td>0.91</td>
<td>0.92</td>
</tr>
<tr>
<td>Race/Ethnicity</td>
<td>0.90</td>
<td>0.95</td>
</tr>
<tr>
<td><strong>Sensitivity</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>0.94</td>
<td>0.93</td>
</tr>
<tr>
<td>non-white</td>
<td>0.82</td>
<td>0.91</td>
</tr>
<tr>
<td>Hispanic</td>
<td>0.66</td>
<td>0.66</td>
</tr>
<tr>
<td>NH White</td>
<td>0.95</td>
<td>0.95</td>
</tr>
<tr>
<td>NH non-white</td>
<td>0.79</td>
<td>0.82</td>
</tr>
</tbody>
</table>

N = 18,893 – patients in the cancer registry who also have self-report RE. Discharges from 2009
Agreement among newborn mothers: PDD versus birth records

<table>
<thead>
<tr>
<th>Overall agreement</th>
<th>Versus birth certificate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Race</td>
<td>0.70</td>
</tr>
<tr>
<td>Ethnicity</td>
<td>0.86</td>
</tr>
<tr>
<td>Race/Ethnicity</td>
<td>0.86</td>
</tr>
<tr>
<td>White</td>
<td>0.72</td>
</tr>
<tr>
<td>non-white</td>
<td>0.85</td>
</tr>
<tr>
<td>Hispanic</td>
<td>0.89</td>
</tr>
<tr>
<td>NH White</td>
<td>0.91</td>
</tr>
<tr>
<td>NH non-white</td>
<td>0.84</td>
</tr>
</tbody>
</table>

2009 Births
N = 513,456

No single answer (or source) for data audits

- Accuracy and agreement for RE depend upon the population to be studied
- “Majority” groups tend to be captured more accurately
  - Older non-Hispanic white cancer patients
  - Younger Hispanic women having babies

Mean-population comparison

Estimate disagreement between:
- Reported = Distribution of race categories as reported by the hospital
- Predicted = Population mean predicted distribution using zip-code level distribution for each patient

Comparison of mean population estimate with agreement for moms

\[ \rho = 0.5 \]

N = 513,456 at 254 Hospitals where births occurred in 2009
Agreement between PDD & birth versus PDD & Census
Audit measures by states and years

Rates of high ‘other’ RE reporting by state

Rates of high ‘unknown’ RE reporting by state

Average estimated disagreement in RE across hospitals by state

Rates of high* disagreement of RE reporting by state

Follow-up Hospital Survey

- 45% response rate
- 60% of original respondents
- Minimal acknowledgement and use of training materials
Remaining work

- Metric for assessment of language using language fluency (ACS 2007-2011)
- Indirect data improvements
  - Substitution using existing data
  - Imputation (statistical prediction)
- Trend analysis in California and other states through 2012

Conclusions / Lessons Learned

- Hospitals are not optimizing data collection of self-reported demographic measures
  - Incentives to improve reporting are absent
  - Materials developed are under-utilized
- New audit measures show promise and could be used in combination to paint a more comprehensive picture of accuracy
  - Most states have not yet attempted to improve data collection, as evidenced by trends