Building a Disease Registry for Research & Analytics: A Breast Cancer Disparities Case Study

Edward M. Rafalski, PhD, MPH, FACHE
Data Warehousing, Mining & BIG Data

• Important to clarify the differences
  • A data warehouse consists of a set of programs that are used to extract data from the transactional system
  • Mining is a term used to describe the analysis of warehoused data in order to generate new insights
  • BIG Data is the term for a collection of data sets so large and complex that it becomes difficult to process using on-hand database management tools or traditional data processing applications. The challenges include capture, curation, storage, search, sharing, transfer, analysis, and visualization. The trend to larger data sets is due to the additional information derivable from analysis of a single large set of related data, as compared to separate smaller sets with the same total amount of data, allowing correlations to be found to spot business trends, determine quality of research, prevent diseases, etc.
Why Build a Disease Registry?

- Most data platforms in healthcare are not integrated
- Much of what is integrated is administrative data designed to drop a clean bill
- Much of what is captured in administrative data is driven by Finance’s desire to bill and maximize revenue, as opposed to capture behavioral, clinical or outcome data
- Administrative data is *dirty*
  - Methodist Healthcare, until relatively recently (2009), was not collecting race/ethnicity data correctly using a REL standard
  - In 1997, OMB issued a [Federal Register](https://www.federalregister.gov/) Notice regarding revisions to the standards for the classification of federal data on race and ethnicity. OMB developed race and ethnic standards in order to provide "consistent data on race and ethnicity throughout the Federal Government. The development of the data standards stem in large measure from new responsibilities to enforce civil rights laws." Among the changes, OMB issued the instruction to "mark one or more races" after noting evidence of increasing numbers of interracial children and wanting to capture the diversity in a measurable way and having received requests by people who wanted to be able to acknowledge their or their children's full ancestry rather than identifying with only one group. Prior to this decision, the Census and other government data collections asked people to report only one race.
Equity Quality Improvement Collaborative:
Memphis, Tenn.

Methodist North Hospital:
AF4Q ALLIANCE LEAD: Healthy Memphis Common Table
LOCATION: Memphis, TN
# OF BEDS: 250 licensed beds

Assuring Health Equity in Memphis: It Takes a City

When announcing our Equity Quality Improvement Collaborative selection, this important work was positioned as a community initiative to ensure health care equity and not simply a Methodist North project. With evidence-based cardiac care an ongoing focus, segmentation confirmed that measures of ideal cardiac care were indeed equitable for all patients regardless of race or ethnicity. But analysis of early data showed that disparities did exist, and in the most important of areas – quality of life and longevity.

Black heart attack patients, on average, are nine years younger than White patients. The age difference for congestive heart failure was even more disparate, with Black patients on average 15 years younger than their White counterparts. More significant was the realization that fewer Black patients were seeking timely help. In review of patients who experienced collapse outside of the hospital, we found that a disproportionate number of sudden death patients were Black. Moreover, the White patients were found to have an additional 8.5 years of life.
Racial Disparities in Breast Cancer Mortality in the 25 Largest US Cities

Memphis
July 31, 2012

Steve Whitman, PhD
steve.whitman@sinai.org
Director, Sinai Urban Health Institute
www.suhichicago.org
All Cause Mortality

Log Scale (relative difference)

<table>
<thead>
<tr>
<th>Year</th>
<th>NHB CHI</th>
<th>NHW CHI</th>
</tr>
</thead>
<tbody>
<tr>
<td>1990</td>
<td>36%</td>
<td>42%</td>
</tr>
<tr>
<td>2005</td>
<td>36%</td>
<td>42%</td>
</tr>
</tbody>
</table>

*p<0.05
Female Breast Cancer Mortality

Log Scale (relative difference)

- NHB CHI
- NHW CHI

*p<0.05

Rate

100

20%

10

1990 2005

Year

99%*

*p<0.05
Methodology: NHB/NHW Mortality Rate Ratios

- 24 most populous cities in US (no data for Indianapolis available)
- National Death Certificate files from CDC
- Computed NHB/NHW breast cancer mortality rate ratios to measure the NHB:NHW disparities
  - Many other disparities were…not measured
- All mortality rates are age-adjusted and expressed per 100,000 women
- Used 95% confidence intervals
Methodology: Ecologic Variables

Used 7 city-level (ecological) risk factors seeking correlates of the disparity:

1. Population of City
2. % NHW
3. % NHB
4. Median Household Income
5. % Below Poverty
6. Gini Index: measure of income inequality
7. Index of Dissimilarity: measure of racial segregation
## 5 Cities with Highest Rate Ratios

<table>
<thead>
<tr>
<th>City</th>
<th>NHB Rate</th>
<th>NHW Rate</th>
<th>RR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Memphis</td>
<td>44.6</td>
<td>21.3</td>
<td>2.09</td>
</tr>
<tr>
<td>Denver</td>
<td>30.8</td>
<td>17.7</td>
<td>1.74</td>
</tr>
<tr>
<td>LA</td>
<td>46.5</td>
<td>27.4</td>
<td>1.70</td>
</tr>
<tr>
<td>Houston</td>
<td>47.3</td>
<td>28.7</td>
<td>1.65</td>
</tr>
<tr>
<td>Chicago</td>
<td>37.8</td>
<td>23.4</td>
<td>1.62</td>
</tr>
</tbody>
</table>
AVON Planning Grant

• Build a breast cancer registry
• Attempt to answer three initial questions:
  1. Non-Hispanic black women and the uninsured wait longer than non-Hispanic white women to get screened after noticing a lump or other symptoms
  2. Under-served women manifest breast cancer (or more aggressive cancer growth/advanced stage cancer) at younger ages, yet screening guidelines do not promote early detection among women under 40
  3. Non-Hispanic black women need navigation via a trusted liaison or navigator to access available treatment and ensure that patients aren’t lost to follow-up.
• We are positioned to garner metrics across the continuum of care, including treatment, to identify barriers affecting mortality at all touch points. In addition, we propose to measure the impact of a trust element to improve process delays.
Registry Structure

- 165 Variables identified (in initial structure)
- Five (5) years of historical encounters
- All woman >=19
- Many more (datasets & variables) to come...
  - Tennessee Cancer Registry (TCR)
  - Methodist Primary Care Physicians (PCG)
  - Other interested stakeholders
Next Steps

• Envisioning a ‘prospective/concurrent database capture design’ using SlimPrim
  – Navigators using iPad devices in the field to capture behavioral variables

• Additional Grant Funding
  – AVON
  – Komen
  – CMS Innovation
  – PCORI
  – Others TBD
QUESTIONS?

Ed.Rafalski@mlh.org