How Often Do Workplace Injuries Go Unreported?

HARVARD CATALYST BIOSTATISTICS PROGRAM SEMINAR
MARCH 21, 2012

Les Boden
Boston University
School of Public Health
Occupational Injury And Illness

- Estimates of 4-8 million workplace lost-time injuries annually, 6,000 deaths
- Unknown number of illnesses
- Injury cost $110 billion annually
- NIH estimates cancer costs: $226 billion
- Note: incident age much younger for occ. injuries: average 36 years old.
Motivation for this study

- Studies in individual work settings and of specific injuries show low reporting
- 1980s studies of fatal traumatic injuries
  - Multiple sources: death certificates, medical examiner reports, workers’ comp, etc.
  - Findings? (Stout and Bell 1991)
  - 40-70% of fatal traumatic work-related injuries went unreported
Reported Workplace Fatalities

- 1990: 2900
- 1992: 6271
Why Do We Care?
Better Surveillance Can

- Identify hazards
- Characterize burden of injuries and illnesses
- Guide prevention strategies
Participating States

- Minnesota
- New Mexico
- Oregon
- Washington

- West Virginia
- Wisconsin
- California: in process
## Data Sources

- **State Workers’ Compensation Data (WC)**
  - If medical or cash benefits paid
  - Reported to state WC agency

- **Bureau of Labor Statistics Annual Survey of Injuries and Illnesses (BLS)**
  - Stratified probability sample of employers
  - Provides state and national estimates of non-fatal injury incidence
What We Did

- Collect BLS and WC injury data
- Determine which injuries are reported to BLS only, to WC only, & to both
- Use capture-recapture to estimate the number of unreported injuries
Important Issues

- Accurate linking (no lost marks assumption)
- Harmonizing scope (closed population assumption)
- Source dependence
- Population heterogeneity
Linking cases across data sources

- Link datasets for specific time frame
- Deterministic linkage
  - cases match on predetermined number of variables
  - Soundex helps with alternate name spelling
- Probabilistic linkage
- Clerical review of indeterminate cases
### Case Linkage

<table>
<thead>
<tr>
<th>Common elements</th>
<th>WC</th>
<th>SOI I</th>
</tr>
</thead>
<tbody>
<tr>
<td>Worker first name</td>
<td>X</td>
<td>X (or F.I.)</td>
</tr>
<tr>
<td>Worker last name</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Employer name</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Employer identifier (FEIN)</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Employer address</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Date of injury</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Date of birth</td>
<td>X</td>
<td>X (or age)</td>
</tr>
<tr>
<td>Sex</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>ER_ID</td>
<td>DOB1</td>
<td>DOB2</td>
</tr>
<tr>
<td>-------</td>
<td>--------</td>
<td>--------</td>
</tr>
<tr>
<td>=</td>
<td>12/15/98</td>
<td>12/16/98</td>
</tr>
<tr>
<td>=~</td>
<td>09/28/70</td>
<td>06/22/71</td>
</tr>
<tr>
<td>=~</td>
<td>05/19/65</td>
<td>05/19/65</td>
</tr>
<tr>
<td>=~</td>
<td>12/15/64</td>
<td>02/15/64</td>
</tr>
<tr>
<td>=~</td>
<td>06/15/59</td>
<td>06/19/59</td>
</tr>
<tr>
<td>=</td>
<td>05/18/98</td>
<td>05/18/98</td>
</tr>
<tr>
<td>.</td>
<td>02/27/67</td>
<td>02/18/99</td>
</tr>
<tr>
<td>.</td>
<td>04/05/44</td>
<td>10/05/44</td>
</tr>
<tr>
<td>.</td>
<td>06/14/69</td>
<td>08/14/59</td>
</tr>
<tr>
<td>=~</td>
<td>04/07/66</td>
<td>08/10/00</td>
</tr>
<tr>
<td>=~</td>
<td>02/12/77</td>
<td>01/16/98</td>
</tr>
</tbody>
</table>
## Scope of Data Sources

<table>
<thead>
<tr>
<th>Data source</th>
<th>Injury or Illness report</th>
<th>Inclusion criteria</th>
<th>Excluded populations</th>
<th>Sampling</th>
</tr>
</thead>
<tbody>
<tr>
<td>BLS</td>
<td>Employer</td>
<td>&gt;=1 work days lost</td>
<td>Agriculture &lt; 10 employees</td>
<td>Stratified Sample</td>
</tr>
<tr>
<td>WC</td>
<td>Worker or employer</td>
<td>No min days, 0-3 days = &lt;4 (if medical costs)</td>
<td>Self-employed, other WC systems</td>
<td>Population</td>
</tr>
</tbody>
</table>


Harmonizing Scope

- Restrict BLS to 4+ days
- Restrict BLS to exclude
  - Longshore and harbor workers
  - Railroad workers
- Restrict WC to
  - Exclude agriculture < 10 employees
  - Include only establishments sampled by BLS
Source Dependence

- Capture-recapture generally assumes that the reporting sources are independent
- What about in this case?
Source Independence Unlikely

- No WC report means BLS report less likely
  - Doctor doesn’t diagnose (esp. disease)
  - Worker doesn’t report injury
  - Employer doesn’t think it’s work-related
  - Employer uses WC records to fill out BLS survey
- \(\Rightarrow\) Positive source dependence
- \(\Rightarrow\) Estimates of underreporting lower than actual number of missed cases
### Reporting Completeness – Source Independence (OR=1)

<table>
<thead>
<tr>
<th></th>
<th>WA</th>
<th>WV</th>
<th>OR</th>
<th>WI</th>
<th>NM</th>
<th>MN</th>
</tr>
</thead>
<tbody>
<tr>
<td>WC</td>
<td>93%</td>
<td>91%</td>
<td>77%</td>
<td>75%</td>
<td>67%</td>
<td>65%</td>
</tr>
<tr>
<td>BLS</td>
<td>55%</td>
<td>76%</td>
<td>56%</td>
<td>65%</td>
<td>51%</td>
<td>68%</td>
</tr>
<tr>
<td>Neither</td>
<td>4%</td>
<td>3%</td>
<td>12%</td>
<td>8%</td>
<td>16%</td>
<td>13%</td>
</tr>
</tbody>
</table>
### Underreporting Estimates – Impact of Source Dependence

<table>
<thead>
<tr>
<th>OR</th>
<th>WA</th>
<th>WV</th>
<th>OR</th>
<th>WI</th>
<th>NM</th>
<th>MN</th>
</tr>
</thead>
<tbody>
<tr>
<td>OR=1</td>
<td>4%</td>
<td>3%</td>
<td>12%</td>
<td>8%</td>
<td>16%</td>
<td>13%</td>
</tr>
<tr>
<td>OR=3</td>
<td>10%</td>
<td>7%</td>
<td>24%</td>
<td>19%</td>
<td>33%</td>
<td>26%</td>
</tr>
<tr>
<td>OR=5</td>
<td>14%</td>
<td>10%</td>
<td>32%</td>
<td>26%</td>
<td>43%</td>
<td>33%</td>
</tr>
</tbody>
</table>
Estimation Steps

- Initially: unadjusted estimates
- Then model heterogenous capture probabilities
  - Multinomial logit estimation
  - Corrects covariance-induced dependence
  - Informs about factors related to reporting
- Sensitivity analysis to examine potential effect of source dependence
Reasons for Underreporting

- From all data:
  - Injury occurred late in year
- From WC linked/unlinked data:
  - Injury reported after end of the year
  - Dispute over payment of WC benefits
  - Delay between date of injury and date injury reported to employer
Reasons for Underreporting II

- From employer survey:
  - Employer didn’t think injury was work-related
  - Employer was unclear about reporting requirements
  - When injury not initially reportable (no lost work) did not add if became reportable later
Difficult to Quantify Reasons

- Exclusions: “independent contractors”
- Worker afraid to report
- Worker doesn’t know about WC
- Employer suppresses reporting to avoid
  - WC costs
  - Disadvantage in bidding for contracts
- “Safety bingo”
Difficult to Quantify Reasons II

- WC law places barriers to benefits
  - “Objective evidence” of disability
  - Increasing amount of evidence needed to receive benefits
  - Exclude “diseases of aging”

- Stigmatization of injured workers
  - Drug testing of those reporting injuries
  - Publicity designed to link WC and fraud
Conclusion

- Capture-recapture can be useful in reducing surveillance bias
- Accurate linking, identical scope are both important and can be difficult
- Source dependence is probably important
- Cannot identify source dependence with only 2 sources