Using Surveillance Data for Research Requires Caution: An example with Clostridium difficile infection ascertainment

Wednesday, May 2nd, 2018
Bobby Warren

Sponsoring organization: Duke Infection Control Outreach Network (DICON)
Sponsoring supervisor: Dr. Deverick Anderson
Overview

• *C. difficile* background
• Healthcare facility-associated infections (HAIs) vs. Community-acquired Infections (CAIs)
• Acquisition definitions
• Cohort study
• Discussion
• Questions
Clostridium difficile

- Anaerobic
- Spore-forming
- Persistent
- Fecal-oral route
- Colonization\(^1\)

C. difficile infections (CDI)

- Antibiotics disrupt normal gut flora
- *C. difficile* proliferates, produces toxins
- Pseudomembranous colitis, severe diarrhea and dehydration²
- Can be classified as healthcare-associated or community-acquired.

Healthcare-associated infections (HAIs)

• Also known as nosocomial infections
• Acquired during treatment in healthcare setting
• *C. difficile* is the leading cause
  – 453,000 cases, 29,000 deaths in 2011
  – Still rising
Community-acquired infections (CAIs)

- Acquired from non-healthcare setting
- *C. difficile* infections historically understood as healthcare-associated (HAI)
  - 20-50% of CDIs are now community-acquired (CAI)\(^4\)
Research questions

• Big picture
  – CA-CDI really rising?
  – CA-CDI origin?

• Specific question
  – Are the current acquisition definitions accurately defining CDI cases?
Laboratory Identified Event Reporting (LabID)

- Simple and efficient
- Facility-level CDI rates
- NOT designed for epidemiologic purposes
LabID Definitions

• **Community-acquired (CAI):** Stool collected ≤ 3 days after admission from a patient who was not discharged from the same facility in the previous 4 weeks.  

• **Healthcare-facility associated (HAI):** Stool collected ≤ 3 days after admission from a patient who was discharged from the same facility in the previous 4 weeks.  

Proposed definitions

• **Community-acquired (CAI)**: Stool collected \(\leq 3\) days after admission from a patient who was not discharged from a **healthcare facility** in the previous 4-weeks.

• **Healthcare-facility associated (HAI)**: Stool collected \(\leq 3\) days after admission from a patient who was discharged from a **healthcare facility** in the previous 4-weeks.
Retrospective cohort study

• CDI cases from 2011-2017 from DICON
• Duke Infection Control Outreach Network (DICON)
  – 46 hospitals in southeast US
  – Share surveillance data and consultative services
  – LabID definitions
• Deployed proposed definitions retroactively
Exclusion criteria

- CDI cases from 2011-2017 in DICON database
- 4-week (28 day) cutoff
  - No previous discharge date → excluded
  - Invalid entry → excluded
Retroactive classification

- If $\leq 28$ day since last admission $\rightarrow$ HAI
- If $> 28$ days since last admission
  - Admitted from home $\rightarrow$ CAI
  - Admitted from healthcare facility $\rightarrow$ HAI
    - Nursing home
    - Hospice
    - Home health
    - Etc.
Redefining acquisition

DICON 23,313

- Known Previous Discharge Date 10,912
- No Previous Discharge Date 12,265
- Invalid entry (0) 103
- Invalid entry (-) 33

HAI

- ≤ 28 days since last discharge 4785
- > 28 days since last discharge 6127

HAI

- Admitted from HCF 1494
- Admitted from home 4404
- Admitted from other/unknown 229

CAI
## Results

- 23,313 CDI $\rightarrow$ 10,683
  - Majority excluded for missing discharge date
- 23.6% miscategorized overall
  - 7.0% of HAIs as CAIs
  - 43.7% of CAIs as HAIs

<table>
<thead>
<tr>
<th>LabID Surveillance</th>
<th>HAI (n = 6279)</th>
<th>CAI (n = 4404)</th>
<th>Total (n = 10683)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CAI</td>
<td>352 (5.6)</td>
<td>1387 (31.5)</td>
<td>1739 (16.3)</td>
</tr>
<tr>
<td>HAI</td>
<td>4659 (74.2)</td>
<td>1075 (24.4)</td>
<td>5734 (53.7)</td>
</tr>
<tr>
<td>Indeterminant</td>
<td>771 (12.3)</td>
<td>1620 (36.8)</td>
<td>2391 (22.4)</td>
</tr>
<tr>
<td>Not Followed</td>
<td>497 (7.9)</td>
<td>322 (7.3)</td>
<td>819 (7.6)</td>
</tr>
<tr>
<td><strong>Miscatergorized %</strong></td>
<td><strong>7</strong></td>
<td><strong>43.7</strong></td>
<td><strong>23.6</strong></td>
</tr>
</tbody>
</table>
## Results

### Optimized Definitions

<table>
<thead>
<tr>
<th>LabID Surveillance</th>
<th>HAI</th>
<th>CAI</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>LabID Surveillance</td>
<td>(n = 6279)</td>
<td>(n = 4404)</td>
<td>(n = 10683)</td>
</tr>
<tr>
<td>CAI</td>
<td>352 (5.6)</td>
<td>1387 (31.5)</td>
<td>1739 (16.3)</td>
</tr>
<tr>
<td>HAI</td>
<td>4659 (74.2)</td>
<td>1075 (24.4)</td>
<td>5734 (53.7)</td>
</tr>
<tr>
<td>Indeterminant</td>
<td>771 (12.3)</td>
<td>1620 (36.8)</td>
<td>2391 (22.4)</td>
</tr>
<tr>
<td>Not Followed</td>
<td>497 (7.9)</td>
<td>322 (7.3)</td>
<td>819 (7.6)</td>
</tr>
<tr>
<td>Miscategorized %</td>
<td>7</td>
<td>43.7</td>
<td>23.6</td>
</tr>
</tbody>
</table>

### Optimized Definitions

<table>
<thead>
<tr>
<th>LabID Surveillance</th>
<th>≤ 28 days since last admission (HAI)</th>
<th>&gt; 28 Days but from Healthcare setting (HAI)</th>
<th>&gt; 28 days, from home (CAI)</th>
<th>Total (n = 10683)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CAI</td>
<td>42 (0.9)</td>
<td>310 (20.7)</td>
<td>1387 (31.5)</td>
<td>1739 (16.3)</td>
</tr>
<tr>
<td>HAI</td>
<td>4223 (88.3)</td>
<td>436 (29.2)</td>
<td>1075 (24.4)</td>
<td>5734 (53.7)</td>
</tr>
<tr>
<td>Indeterminant</td>
<td>145 (3.0)</td>
<td>626 (41.9)</td>
<td>1620 (36.8)</td>
<td>2391 (22.4)</td>
</tr>
<tr>
<td>Not Followed</td>
<td>375 (7.8)</td>
<td>122 (8.2)</td>
<td>322 (7.3)</td>
<td>819 (7.6)</td>
</tr>
<tr>
<td>Miscategorized %</td>
<td><strong>1.0</strong></td>
<td><strong>41.6</strong></td>
<td>43.7</td>
<td>23.6</td>
</tr>
</tbody>
</table>
## Results

### Optimized Definitions

<table>
<thead>
<tr>
<th>LabID Surveillance</th>
<th>≤ 28 days since last admission (HAI)</th>
<th>&gt; 28 Days but from Healthcare setting (HAI)</th>
<th>&gt; 28 days, from home (CAI)</th>
<th>Total (n = 10683)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CAI</td>
<td>42 (0.9)</td>
<td>310 (20.7)</td>
<td>1387 (31.5)</td>
<td>1739 (16.3)</td>
</tr>
<tr>
<td>HAI</td>
<td>4223 (88.3)</td>
<td>436 (29.2)</td>
<td>1075 (24.4)</td>
<td>5734 (53.7)</td>
</tr>
<tr>
<td>Indeterminant</td>
<td>145 (3.0)</td>
<td>626 (41.9)</td>
<td>1620 (36.8)</td>
<td>2391 (22.4)</td>
</tr>
<tr>
<td>Not Followed</td>
<td>375 (7.8)</td>
<td>122 (8.2)</td>
<td>322 (7.3)</td>
<td>819 (7.6)</td>
</tr>
<tr>
<td>Miscategorized %</td>
<td>1.0</td>
<td>41.6</td>
<td>43.7</td>
<td>23.6</td>
</tr>
</tbody>
</table>

### Admitted From a Healthcare Facility

<table>
<thead>
<tr>
<th>LabID Surveillance</th>
<th>Home Health (n = 102)</th>
<th>Hospice (n = 6)</th>
<th>Hospital (n = 93)</th>
<th>Nursing Home (n = 1117)</th>
<th>Other Extended Care Facility (n = 175)</th>
<th>Total (n = 1493)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CAI</td>
<td>17 (16.7)</td>
<td>1 (16.7)</td>
<td>30 (32.3)</td>
<td>220 (19.7)</td>
<td>41 (23.4)</td>
<td>309 (20.7)</td>
</tr>
<tr>
<td>HAI</td>
<td>21 (20.6)</td>
<td>3 (50.0)</td>
<td>44 (47.3)</td>
<td>316 (28.3)</td>
<td>52 (29.7)</td>
<td>436 (29.2)</td>
</tr>
<tr>
<td>Indeterminant</td>
<td>53 (52.0)</td>
<td>2 (33.3)</td>
<td>18 (19.4)</td>
<td>476 (42.6)</td>
<td>77 (44.0)</td>
<td>626 (41.9)</td>
</tr>
<tr>
<td>Not Followed</td>
<td>11 (10.8)</td>
<td>0 (0)</td>
<td>1 (1.1)</td>
<td>105 (9.4)</td>
<td>5 (2.9)</td>
<td>122 (8.2)</td>
</tr>
<tr>
<td>Miscategorized %</td>
<td>44.7</td>
<td>25.0</td>
<td>40.5</td>
<td>41.0</td>
<td>44.1</td>
<td>41.6</td>
</tr>
</tbody>
</table>
Discussion

• Frequent misclassification
• 41.6% (310 of 746) of cases that:
  – Not admitted to index hospital in past 28 days
  – Admitted from a healthcare setting
  – Likely HAIs

• LabID surveillance definitions will continue to misclassify cases unless changed.
Discussion

• LabID definitions are ideal for single hospital surveillance.

• Heavily flawed when used for analyses that go beyond that hospital.

• Takeaway:
  – When using aggregate surveillance data know:
    • Why and how each data point is collected.
Limitations

• Community hospitals in southeast US
  – Not generalizable

• CDIs were required to have a discharge date
  – Increased exposure to healthcare facilities
  – Higher HAIs
  – Large chunk of data lost (~50%)

• Unable to evaluate for errors in database.
Summary

• LabID definitions
  – Underestimate HA-CDIs
  – Overestimate CA-CDIs

• Caution is needed when using LabID surveillance definitions
  – And any other surveillance data
Research questions revisited

• Big picture
  – CA-CDI really rising?
  – CA-CDI origin?

• Specific question
  – Are the current acquisition definitions accurately defining cases?
Thanks

- **DICON**
  - Dr. Deverick Anderson
  - Dr. Nicholas Turner
  - Rachel Addison

- **Fowler Lab**
  - Dr. Vance Fowler
  - Dr. Batu Sharma
  - Brenna Hansen

- **Graduate School**
  - Dr. Heidi Harkins

- **SILS**
  - Dr. Fei Yu
References


Questions?
<table>
<thead>
<tr>
<th>Demographic Characteristics</th>
<th>CDI Cases (n = 23313)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Ethnicity</strong></td>
<td></td>
</tr>
<tr>
<td>Caucasian</td>
<td>11549 (49.5)</td>
</tr>
<tr>
<td>Unknown</td>
<td>6277 (26.9)</td>
</tr>
<tr>
<td>African American</td>
<td>4894 (21)</td>
</tr>
<tr>
<td>American Indian</td>
<td>279 (1.2)</td>
</tr>
<tr>
<td>Asian</td>
<td>61 (0.3)</td>
</tr>
<tr>
<td>Hispanic</td>
<td>142 (0.6)</td>
</tr>
<tr>
<td>Other</td>
<td>111 (0.5)</td>
</tr>
<tr>
<td><strong>Age Distribution</strong></td>
<td></td>
</tr>
<tr>
<td>0-10</td>
<td>350 (1.5)</td>
</tr>
<tr>
<td>11-20</td>
<td>134 (0.6)</td>
</tr>
<tr>
<td>21-30</td>
<td>693 (3.0)</td>
</tr>
<tr>
<td>31-40</td>
<td>1063 (4.6)</td>
</tr>
<tr>
<td>41-50</td>
<td>1922 (8.2)</td>
</tr>
<tr>
<td>51-60</td>
<td>3383 (14.5)</td>
</tr>
<tr>
<td>61-70</td>
<td>4834 (20.7)</td>
</tr>
<tr>
<td>71-80</td>
<td>5370 (23)</td>
</tr>
<tr>
<td>81-90</td>
<td>4424 (19)</td>
</tr>
<tr>
<td>91-100</td>
<td>1117 (4.8)</td>
</tr>
<tr>
<td>100+</td>
<td>23 (0.1)</td>
</tr>
<tr>
<td><strong>Sex</strong></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>13811 (59.2)</td>
</tr>
<tr>
<td>Male</td>
<td>9058 (38.9)</td>
</tr>
<tr>
<td>Unknown/Blank</td>
<td>444 (1.9)</td>
</tr>
</tbody>
</table>