Comparability Assessment of Cohorts with and without Laboratory Values

Rupa Makadia, MS1,2, Jill Hardin, MS, PhD1,2, Laura Hester, PhD1,2, Chris Knoll, BCS1,2, Ajit Londhe, MPH1,2, Joel Swerdel, MPH1,2

1 Janssen Research & Development, LLC, Titusville, NJ 2 OHDSI collaborators, Observational Health Data Sciences and Informatics (OHDSI), New York, NY

Table 1. Condition cohort descriptions and statistics

<table>
<thead>
<tr>
<th>Measurement</th>
<th>Target Cohort (Absence) (n)</th>
<th>Comparator Cohort (Presence) (n)</th>
<th>Total n (All Domains)</th>
<th>Condition (n)</th>
<th>Drug (n)</th>
<th>Procedure (n)</th>
<th>(N) covariates w/ abs std difference &gt; 0.1 (All Domains)</th>
<th>covariate differences by domain w/ abs std difference &gt; 0.1 (N, %)**</th>
</tr>
</thead>
<tbody>
<tr>
<td>NH measure</td>
<td>1,726</td>
<td>2,307</td>
<td>4,034</td>
<td>1,916</td>
<td>3,023</td>
<td>1,296</td>
<td>4,562</td>
<td>4,176</td>
</tr>
<tr>
<td>Reactive protein</td>
<td>4,813</td>
<td>3,846</td>
<td>8,659</td>
<td>2,297</td>
<td>3,542</td>
<td>3,351</td>
<td>2,868</td>
<td>2,729</td>
</tr>
<tr>
<td>DLc measure</td>
<td>217,867</td>
<td>80,752</td>
<td>329,977</td>
<td>21,345</td>
<td>23,655</td>
<td>12,320</td>
<td>24,345</td>
<td>20,580</td>
</tr>
<tr>
<td>Total cholesterol</td>
<td>15,675</td>
<td>19,269</td>
<td>34,977</td>
<td>16,741</td>
<td>17,565</td>
<td>14,746</td>
<td>17,345</td>
<td>15,594</td>
</tr>
</tbody>
</table>

Figure 1. Plots of covariate prevalence for 4 cohorts

METHODS cont

- Plots were generated comparing cohorts with and without laboratory measurements and values prior to the index date.

RESULTS

- The comparisons of cohorts with a measurement and value vs. cohorts with a measurement and no value have more imbalanced covariates than comparisons between cohorts with laboratory measurements vs. cohorts without laboratory measurements.

- Cohorts of subjects with hepatitis B and Crohn’s disease had a higher proportion of covariates with (aSMD) means > 0.1 compared to hyperlipidemia and patients aged 50+ cohorts.

- There are differences among the number and domain of imbalanced covariates between the cohorts with and without laboratory measurements and the cohorts with laboratory measurements with and without values across each outcome.

LIMITATIONS

- Lab results are only available within certain datasets, therefore studies requiring lab data may not be generalizable to broader populations.

- Claims data only capture laboratory tests that are reimbursed and therefore some measurements (e.g. body weight) are more difficult to assess.

CONCLUSIONS

- A developed systematic framework to assess if use of laboratory measurement data is appropriate to represent subjects without measurements.

- When using measurements to define cohorts, we suggest thoroughly examining the underlying cohorts for comparability.

NEXT STEPS

- This study serves as the basis for developing a set of criteria to illustrate similarity between cohorts of subjects with and without measurements. Further research will include metrics and tests to evaluate the use of cohorts with and without measurements in analyses.

REFERENCES


CONFLICT OF INTEREST STATEMENT

The authors are full time employees of Janssen Research and Development, a unit of Johnson and Johnson. The work on this study was part of their employment. They also hold pension rights from the company and own stock and stock options.