Distributed Hospital Comparer: A new end-to-end data aggregation approach for comparing hospital performance without sharing patient-level data

Presenter: Jiayi (Jessie) Tong
Ph.D. Candidate at the University of Pennsylvania

Co-lead: Jenna Reps¹, Yong Chen²

¹ Janssen R&D
² Department of Biostatistics, Epidemiology and Informatics (DBEI), the Perelman School of Medicine, University of Pennsylvania

November 07, 2023
Hospital profiling

‣ “involves a comparison of a health care provider’s structure, processes of care, or outcomes” – Normand and Shahian 2007

‣ Goal:
  • compare the quality of care between hospitals via structural measures
    – E.g., nursing ratios, presence of residency programs, availability of advanced technology, volume
Variation in US Hospital Mortality Rates for Patients Admitted With COVID-19 During the First 6 Months of the Pandemic

David A. Asch, MD, MBA; Natalie E. Sheils, PhD; Md Nazmul Islam, PhD, MBA; Yong Chen, PhD; Rachel M. Werner, MD, PhD; John Buresh, BS; Jalpa A. Doshi, PhD

It is unknown how much the mortality of patients with coronavirus disease 2019 (COVID-19) depends on the hospital that cares for them, and whether COVID-19 hospital mortality rates are improving.
Map of Collaborators

The OHDSI community brings together volunteers from around the world to establish open community data standards, develop open-source software, conduct methodological research, and apply scientific best practices to both answer public health questions and generate reliable clinical evidence.

Our community is ALWAYS seeking new collaborators. Do you want to focus on data standards or methodological research? Are you passionate about open-source development or clinical applications? Do you have data that you want to be part of global network studies? Do you want to be part of a global community that truly values the benefits of open science? Add a dot to the map below and JOIN THE JOURNEY!

OHDSI By The Numbers

- 3,266 collaborators
- 80 countries
- 21 time zones
- 6 continents
- 1 community
Challenges

- Patient-level data cannot be shared
- “case-mix” situation

A Solution

Distributed Hospital Comparer

GLMM: considered as the gold standard method by National Quality Forum
An OHDSI Study

OHDSI: Applying the Decentralized Generalized Linear Mixed Effects Model (dGEM) for Hospital Profiling of COVID-19 Mortality Data across OHDSI Network

Lead: Jessie Tong1, Jenna Rempel2, Yong Chen1

1 Department of Biostatistics, Epidemiology and Informatics (DBEI), the Perelman School of Medicine, University of Pennsylvania
2 Janssen R&D

May 2022
Implementation – R package

pda: Privacy-Preserving Distributed Algorithms

A collection of privacy-preserving distributed algorithms for conducting multi-site data analyses. The regression analyses can be linear regression for continuous outcome, logistic regression for binary outcome, Cox proportional hazard regression for time-to-event outcome, or Poisson regression for count outcome. The PDA algorithm runs on a lead site and only requires summary statistics from collaborating sites, with one or few iterations. For more information, please visit our software websites: <https://github.com/PennCIL/pda>, and <https://pdamethods.org>.

Version: 1.0-2
Imports: Rcpp (≥ 0.12.19), stats, http, rvest, jsonlite, data.table, survival
LinkingTo: Rcpp, RcppArmadillo
Suggests: imager
Published: 2020-12-10
Author: Chongliang Luo [aut, cre], Rui Daan [aut], Mackenzie Edmondson [aut], Jiayi Tong [aut], Yong Chen [aut], Penn Computing Inference Learning (PennCIL) lab [cph]
Maintainer: Chongliang Luo <luoci3009 at gmail.com>
License: Apache License 2.0
NeedsCompilation: yes
CRAN checks: pda results

Documentation:

Reference manual: pda.pdf

Downloads:

Package source: pda_1.0-2.tar.gz
Windows binaries: r-devel: pda_1.0-2.zip, r-release: pda_1.0-2.zip, r-oldrel: pda_1.0-2.zip
macOS binaries: r-release (arm64): pda_1.0-2.tar.gz, r-oldrel (arm64): pda_1.0-2.tar.gz, r-release (x86_64): pda_1.0-2.tar.gz, r-oldrel (x86_64): pda_1.0-2.tar.gz
Old sources: pda archive
Implementation – PDA-OTA

Welcome to PDA-OTA

PDA-OTA is a web-based interface for secure sharing of aggregated data for multi-site studies using privacy-preserving distributed algorithms. PDA-OTA, once built, will facilitate national and international collaborations requiring aggregated data sharing for collaborative modeling. PDA-OTA synchronizes project status, offers cloud-based SFTP, and generates model-specific tasks for streamlined implementations.

Sign in

- E-mail address
  E-mail

- Password
  Password

- I'm not a robot

Sign in

Sign in with Google

Forgot password?
Create a new account?
Module 1: dGEM (non-iterative decentralized GLMM)

1. Initial estimated common effects
2. Overall estimated common effects
3. Hospital-specific effects

Module 2: Counterfactual Modeling

4. Calibrated hospital-specific effects
5. Counterfactual risk
6. Hospital Ranking
Module 1: dGEM (non-iterative decentralized GLMM)

1. Initial estimated common effects
2. Overall estimated common effects
3. Hospital-specific effects

Module 2: Counterfactual Modeling

4. Calibrated hospital-specific effects
5. Counterfactual risk
6. Hospital Ranking

• Only one round
• No initialization

One-shot & Lossless Generalized Linear Mixed effect Model (OLGLMM)
An OHDSI Study

OLGLMM-COVID project: Applying the One-shot Lossless Generalized Linear Mixed Models (OLGLMM) for Hospital Profiling of COVID-19 Mortality Data across OHDSI Network

Lead: Jessie Tong¹, Jenna Reps², Yong Chen¹

¹ Department of Biostatistics, Epidemiology and Informatics (DBEI), the Perelman School of Medicine, University of Pennsylvania
² Janssen R&D

May 16, 2023
Control file to run OLGLMM

Participating List

<table>
<thead>
<tr>
<th>Name</th>
<th>Unit ID</th>
<th>File upload status</th>
<th>Unit status</th>
<th>Role</th>
<th>Upload date</th>
<th>Download / Preview</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mitsu Brand</td>
<td>03e79a44</td>
<td>1/1</td>
<td>1/1</td>
<td>Data contributor</td>
<td>2023-06-15 06:22:10</td>
<td></td>
</tr>
<tr>
<td>Jenna repo</td>
<td>39853bbf1</td>
<td>2/2</td>
<td>2/2</td>
<td>Data contributor</td>
<td>2023-06-12 14:00:28</td>
<td></td>
</tr>
<tr>
<td>Angela Leis</td>
<td>4da20dee</td>
<td>--</td>
<td>1/1</td>
<td>Project observer</td>
<td>--</td>
<td></td>
</tr>
<tr>
<td>Yu Huang</td>
<td>4f1b3e1f1</td>
<td>1/4</td>
<td>1/4</td>
<td>Data contributor</td>
<td>2023-06-19 21:47:05</td>
<td></td>
</tr>
<tr>
<td>Juan Manuel Ramirez</td>
<td>41f524e8</td>
<td>--</td>
<td>1/1</td>
<td>Project observer</td>
<td>--</td>
<td></td>
</tr>
<tr>
<td>Thomas Falconer</td>
<td>5a5519d8</td>
<td>--</td>
<td>1/1</td>
<td>Data contributor</td>
<td>2023-06-16 15:19:44</td>
<td></td>
</tr>
<tr>
<td>Ross Williams</td>
<td>5b56abf0</td>
<td>1/1</td>
<td>1/1</td>
<td>Data contributor</td>
<td>2023-06-16 12:56:08</td>
<td></td>
</tr>
<tr>
<td>Miguel Angel Mayer</td>
<td>8h20de20</td>
<td>1/1</td>
<td>1/1</td>
<td>Data contributor</td>
<td>2023-06-15 13:37:12</td>
<td></td>
</tr>
<tr>
<td>Jessie Tong</td>
<td>a77f9d2f</td>
<td>0/1</td>
<td>0/1</td>
<td>Data contributor</td>
<td>--</td>
<td></td>
</tr>
</tbody>
</table>
Thank you!