Department of Biostatistics, Epidemiology and Informatics

Synthesizing Evidence for Rare Events: a Novel Zero-Inflated Bivariate Model to Integrate Studies with Double-Zero Outcomes

Lu Li, Ph.D. candidate at the University of Pennsylvania

Advisor: Dr. Yong Chen

Joint work with Drs. Lifeng Lin, Haitao Chu, Yong Chen

2023 OHDSI Symposium

PennCIL

A Computing · Inference · Learning lab at University of Pennsylvania

Real-world case study

Cochrane Library

Cochrane Database of Systematic Reviews

Probiotics for the prevention of Clostridium difficile-associated diarrhea in adults and children (Review)

Goldenberg JZ, Yap C, Lytvyn L, Lo CKF, Beardsley J, Mertz D, Johnston BC

- Explores the potential use of probiotics as a treatment for Clostridium difficile-associated diarrhea (CDAD) caused by antibiotic use
- Whether probiotics cause any side effects when used to prevent CDAD

Analysis 1.24. Comparison 1 Probiotics versus control, Outcome 24 Adverse Events: complete case.

Study or subgroup	Experimental n/N	Control n/N	Risk Ratio M-H, Random, 95% CI	Weight	Risk Ratio M-H, Random, 95% CI
Allen 2013	294/1470	284/1471		13.09%	1.04[0.9,1.2]
Arvola 1999	0/61	0/58			Not estimable
Beausoleil 2007	21/44	20/45		6.6%	1.07[0.68,1.68]
Bravo 2008	3/41	4/45		1.08%	0.82[0.2,3.46]
Cindoruk 2007	41/62	62/62		12.37%	0.66[0.56,0.79]
Duman 2005	3/196	4/180		1.02%	0.69[0.16,3.04]
Ehrhardt 2016	18/146	12/146		3.75%	1.5[0.75,3]
Fominykh 2013	0/80	0/40			Not estimable
Gao 2010	1/171	2/84		0.41%	0.25[0.02,2.67]
Hickson 2007	0/57	0/56			Not estimable
Imase 2008	1/12	3/7		0.55%	0.19[0.02,1.53]
Klarin 2008	0/22	0/22			Not estimable
Koning 2008	15/19	17/19		10.02%	0.88[0.67,1.17]
Kotowska 2005	0/119	0/127			Not estimable
Lonnermark 2010	3/80	3/83		0.92%	1.04[0.22,4.99]
McFarland 1995	0/93	12/92		0.3%	0.04[0,0.66]
Miller 2008a	2/95	4/94		0.81%	0.49[0.09,2.64]
Miller 2008b	4/156	0/155		0.28%	8.94[0.49,164.71]
Nord 1997	9/11	10/12		7.92%	0.98[0.67,1.43]
Ouwehand 2014	14/304	12/144		3.36%	0.55[0.26,1.16]
Pozzoni 2012	41/106	35/98		8.29%	1.08[0.76,1.55]
Psaradellis 2010	90/216	103/221		11.61%	0.89[0.72,1.1]
Ruszczynski 2008	0/120	0/120			Not estimable
Safdar 2008	2/23	5/16		0.99%	0.28[0.06,1.26]
Selinger 2013	14/117	16/112		3.95%	0.84[0.43,1.63]
Shan 2013	0/139	0/144			Not estimable
Shimbo 2005	5/18	14/17		3.15%	0.34[0.16,0.73]
Siitonen 1990	2/8	3/8		1%	0.67[0.15,2.98]
Sullivan 2004	0/18	0/18			Not estimable
Surawicz 1989	0/116	0/64			Not estimable
Thomas 2001	37/133	52/134		8.52%	0.72[0.51,1.01]
Wong 2014	0/76	0/82			Not estimable
Total (95% CI)	**4329**	**3976**		**100%**	**0.83[0.71,0.97]**

PennCIL
A Computing · Inference · Learning lab at University of Pennsylvania

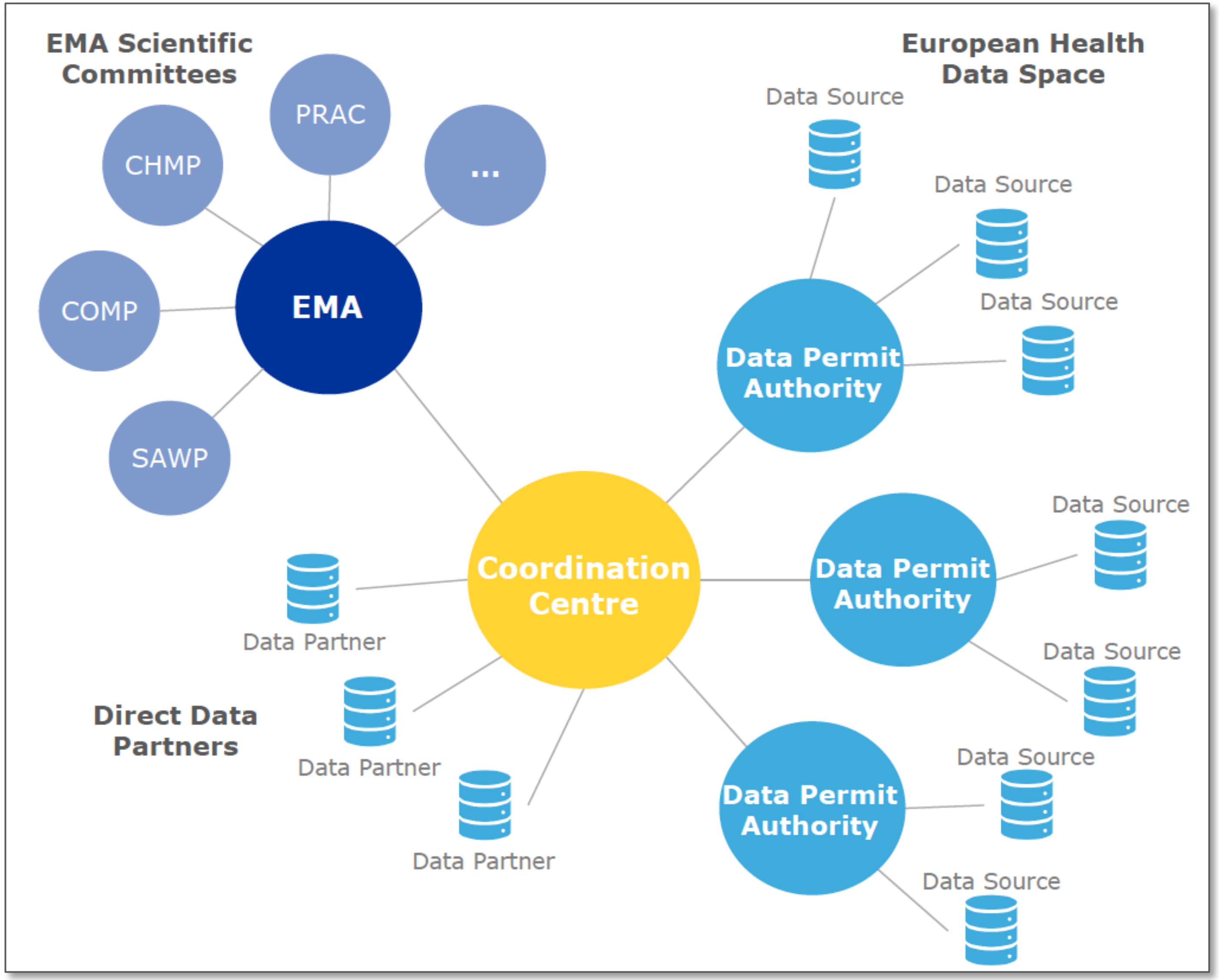
Real-world case study

Analysis 1.24. Comparison 1 Probiotics versus control, Outcome 24 Adverse Events: complete case.

Study or subgroup	Experimental n/N	Control n/N	Risk Ratio M-H, Random, 95% CI	Weight	Risk Ratio M-H, Random, 95% CI
Allen 2013	294/1470	284/1471		13.09%	1.04[0.9,1.2]
Arvola 1999	0/61	0/58			Not estimable

Double-zero-event study (DZS)

DARWIN EU*



* Data Analysis and Real-world Interrogation Network

Should we drop them?

Non-event
Event

Relative Risk

ELSEVIER

Check for updates

Journal of Clinical Epidemiology

Journal of Clinical Epidemiology 123 (2020) 91–99

ORIGINAL ARTICLE

Exclusion of studies with no events in both arms in meta-analysis impacted the conclusions

Chang Xu[a], Ling Li[a], Lifeng Lin[b], Haitao Chu[c], Lehana Thabane[d,e], Kang Zou[a], Xin Sun[a,*]

[a]Chinese Evidence-Based Medicine Center & Cochrane China, West China Hospital, Sichuan University, Chengdu, China
[b]Department of Statistics, Florida State University, Tallahassee, FL, USA
[c]Division of Biostatistics, School of Public Health, University of Minnesota, Minneapolis, MN, USA
[d]Department of Health Research Methods, Evidence, and Impact, McMaster University, 1280 Main Street West, Hamilton, Ontario L8S 4K1, Canada
[e]Biostatistics Unit, Father Sean O'Sullivan Research Centre, St Joseph's Healthcare, 3rd Floor, Martha Wing, Room H-325, 50 Charlton Avenue East, Hamilton, Ontario L8N 4A6, Canada

Accepted 26 March 2020; Published online 1 April 2020

Stat Methods Med Res. Author manuscript; available in PMC 2013 Dec 1.
PMCID: PMC3348438
Published in final edited form as:
NIHMSID: NIHMS351256
Stat Methods Med Res. 2012 Dec; 21(6): 621–633.
PMID: 21177306
Published online 2010 Dec 21. doi: 10.1177/0962280210393712

Author Manuscript

Bivariate Random Effects Models for Meta-Analysis of Comparative Studies with Binary Outcomes: Methods for the Absolute Risk Difference and Relative Risk

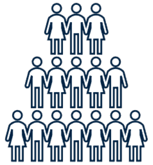
Haitao Chu,[1,*] Lei Nie,[2] Yong Chen,[3] Yi Huang,[4] and Wei Sun[5]

PennCIL
A Computing · Inference · Learning lab at University of Pennsylvania

Existing approach to incorporate DZS:

Bivariate Generalized Linear Mixed Model (BGLMM)

- A bivariate random effects model that ***jointly*** analyzes the risks in treatment and control groups



Group (for the i-th study)	Treatment	Control
Number of events	Y_{i1}	Y_{i0}
Sample size	N_{i1}	N_{i0}
Event risk	P_{i1}	P_{i0}
Fixed effects	μ_1	μ_0
Random effects	ν_{i1}	ν_{i0}

parameters of interest

$$Y_{ik} \sim \textbf{Binomial}\,(N_{ik}, P_{ik});\ g(P_{ik}) = \mu_k + \nu_{ik}$$

$$P(Y_{i0} = y_{i0}, Y_{i1} = y_{i1}) = \prod_{k=0}^{1} (P_{ik})^{y_{ik}} (1 - P_{ik})^{N_{ik} - y_{ik}}$$

Limitation:
BGLMM treats **all DZS similarly** to the other studies

Chu H, Nie L, Chen Y, Huang Y, Sun W. Bivariate random effects models for meta-analysis of comparative studies with binary outcomes: methods for the absolute risk difference and relative risk. Statistical Methods in Medical Research. 2012;21(6):621–633

Should we treat DZS similarly to the other studies?

Short answer:
NO

Rationale:
Sample size of the studies are informative

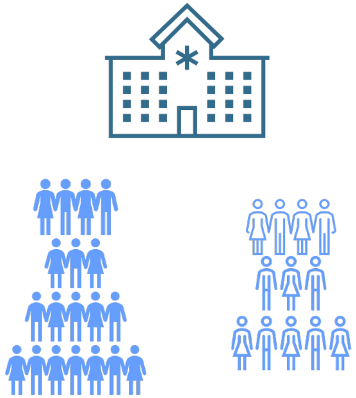
Conclusion:
We **should not** treat all the double zero studies the same as the other studies

Assuming an event rate of 1%:

$\frac{0}{57} / \frac{0}{53}$

$0.99^{110} \approx \frac{33}{100}$

Could happen by chance.



$\frac{0}{527} / \frac{0}{473}$

$0.99^{1,000} \approx \frac{4}{100,000}$

Very unlikely.

To differentiate DZS: Zero-Inflated Models

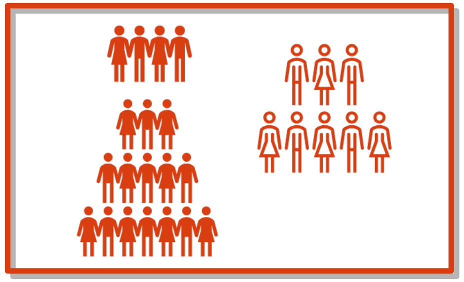
- Zero-inflated models separate observed zeros into two **distinct categories.**

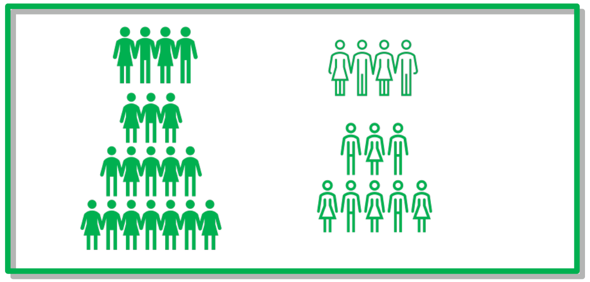
$$Y_{ik} \sim \begin{cases} \text{Binomial}\ (N_{ik}, P_{ik}), & \text{with probability } 1-\pi \\ 0, & \text{with probability } \pi \end{cases}$$

$1-\pi$: at-risk population

π: low-risk population

"at-risk" or "chance" zeros correspond to a latent group of individuals who are at risk for an event but have a recorded count of zeros.





"structural" zeroes represent individuals who are not susceptible to a specific event, thereby having no chance of a positive count.

Lambert D. Zero-inflated Poisson regression, with an application to defects in manufacturing. Technometrics. 1992;34(1):1–14.

Recap BGLMM:

$$P(Y_{i0} = y_{i0}, Y_{i1} = y_{i1}) = \prod_{k=0}^{1} (P_{ik})^{y_{ik}} (1 - P_{ik})^{N_{ik} - y_{ik}}$$

$$g(P_{ik}) = \mu_k + \nu_{ik}$$

Proposed method

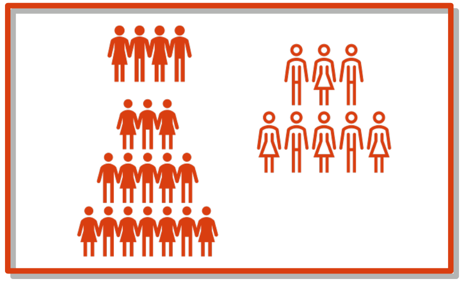
- **Zero-Inflated Bivariate Generalized Linear Mixed Model (ZIBGLMM)**

Advantages:

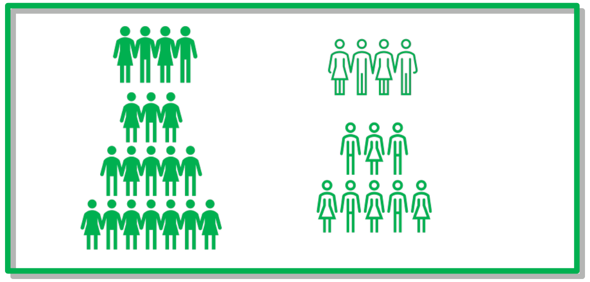
1. No studies are dropped from the analysis.

2. Models population heterogeneity through a data-driven approach.

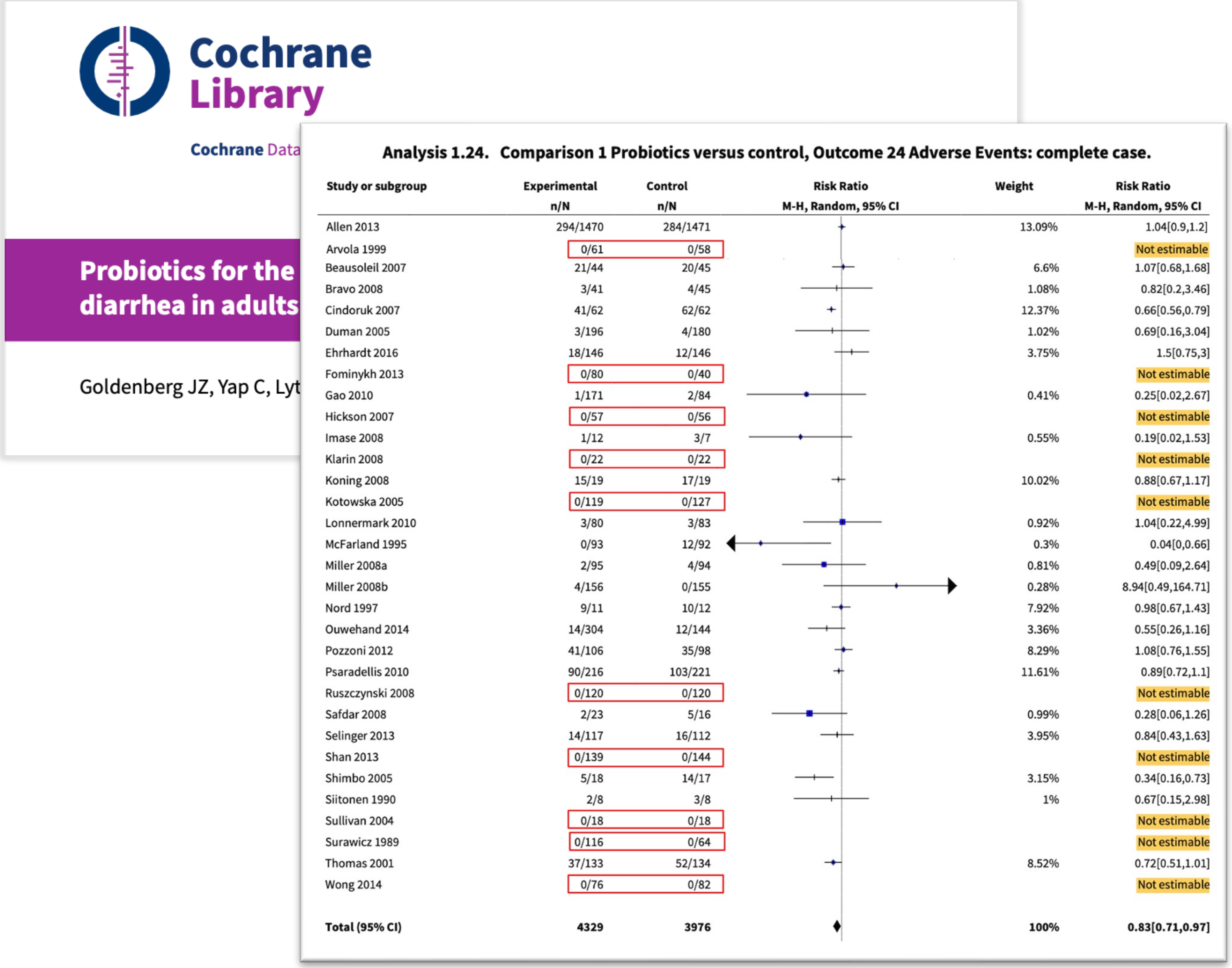
$1 - \pi$: at-risk population



π: low-risk population



Revisit the case study



Analysis 1.24. Comparison 1 Probiotics versus control, Outcome 24 Adverse Events: complete case.

Study or subgroup	Experimental n/N	Control n/N	Weight	Risk Ratio M-H, Random, 95% CI
Allen 2013	294/1470	284/1471	13.09%	1.04[0.9,1.2]
Arvola 1999	0/61	0/58		Not estimable
Beausoleil 2007	21/44	20/45	6.6%	1.07[0.68,1.68]
Bravo 2008	3/41	4/45	1.08%	0.82[0.2,3.46]
Cindoruk 2007	41/62	62/62	12.37%	0.66[0.56,0.79]
Duman 2005	3/196	4/180	1.02%	0.69[0.16,3.04]
Ehrhardt 2016	18/146	12/146	3.75%	1.5[0.75,3]
Fominykh 2013	0/80	0/40		Not estimable
Gao 2010	1/171	2/84	0.41%	0.25[0.02,2.67]
Hickson 2007	0/57	0/56		Not estimable
Imase 2008	1/12	3/7	0.55%	0.19[0.02,1.53]
Klarin 2008	0/22	0/22		Not estimable
Koning 2008	15/19	17/19	10.02%	0.88[0.67,1.17]
Kotowska 2005	0/119	0/127		Not estimable
Lonnermark 2010	3/80	3/83	0.92%	1.04[0.22,4.99]
McFarland 1995	0/93	12/92	0.3%	0.04[0,0.66]
Miller 2008a	2/95	4/94	0.81%	0.49[0.09,2.64]
Miller 2008b	4/156	0/155	0.28%	8.94[0.49,164.71]
Nord 1997	9/11	10/12	7.92%	0.98[0.67,1.43]
Ouwehand 2014	14/304	12/144	3.36%	0.55[0.26,1.16]
Pozzoni 2012	41/106	35/98	8.29%	1.08[0.76,1.55]
Psaradellis 2010	90/216	103/221	11.61%	0.89[0.72,1.1]
Ruszczynski 2008	0/120	0/120		Not estimable
Safdar 2008	2/23	5/16	0.99%	0.28[0.06,1.26]
Selinger 2013	14/117	16/112	3.95%	0.84[0.43,1.63]
Shan 2013	0/139	0/144		Not estimable
Shimbo 2005	5/18	14/17	3.15%	0.34[0.16,0.73]
Siitonen 1990	2/8	3/8	1%	0.67[0.15,2.98]
Sullivan 2004	0/18	0/18		Not estimable
Surawicz 1989	0/116	0/64		Not estimable
Thomas 2001	37/133	52/134	8.52%	0.72[0.51,1.01]
Wong 2014	0/76	0/82		Not estimable
Total (95% CI)	**4329**	**3976**	**100%**	**0.83[0.71,0.97]**

- 10 of 32 studies are double-zero-event studies (DZS), with sample size ranging from 18 to 144.
- Concluded that probiotics reduce the risk of AE by 17%:
  - RR 0.83 (95% CI 0.71 to 0.97)
- Using our proposed method (ZIBGLMM):
  - RR 0.70 (95% CI 0.55 to 0.88)
- Conclusion:
  - Including the DZS could potentially result in estimates that **differ by a large degree** (>0.1).
  - Using ZIBGLMM offers a more **comprehensive analysis** of the available data.

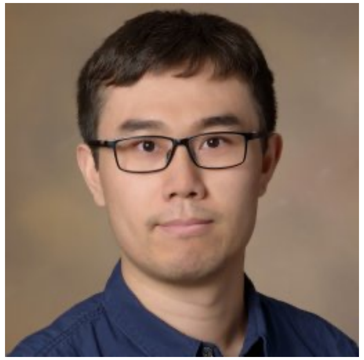
Summary

- Zeros in double-zero-event studies (DZS) may arise due to heterogeneity in the population.
- ZIBGLMM offers a more comprehensive analysis of the available data.

For **OHDSI**, ZIBGLMM is useful especially for **larger network studies** and for studies involving **rare events**.

PennCIL
A Computing · Inference · Learning lab at University of Pennsylvania

Lu Li



Lifeng Lin, Ph.D.

Haitao Chu, Ph.D.

Yong Chen, Ph.D.

Acknowledgments:

Jiayi Tong,

Qiong Wu, Ph.D.,

Dazheng Zhang,

Bingyu Zhang,

Jiajie Chen, Ph.D.,

Tianyu Zhang,

and other lab members for your help with the project.

Correspondence to:

Yong Chen, Ph.D.

Email: ychen123@upenn.edu

Lu Li

Email: luli1@sas.upenn.edu

Poster: # 506

Fri 10/20 4:15 – 5:00 pm

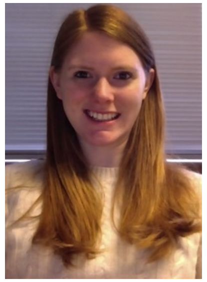
View our code at:

ASSURE | Active Safety Surveillance Using Real-world Evidence

Overview of ASSURE
OHDSI Symposium 2023


Kevin Haynes


Justin Bohn


Jenna Reps


Gowtham Rao


Mitch Conover

Global Epidemiology Organization

OFFICE OF THE CHIEF MEDICAL OFFICER

Standardizing regulatory-grade real-world evidence generation

Standardized design

- Indication
- Target
- Outcome
- Comparator
- Time-at-risk

Standardized analytics

OHDSI — Observational Health Data Sciences and Informatics

Open community data standards | Methodological research | Open-source development

JNJ/Epi standardized data network (OMOP CDM)

- US private claims x3
- US EHR x2
- US Hospital
- US Medicaid
- US Medicare suppl
- Germany EHR
- France EHR
- Australia EHR
- Japan claims

Standardized evidence

- Characterization
  - Disease natural history
  - Treatment utilization
  - Outcome Incidence
  - Time-to-event
- Population-level estimation
  - Comparative cohort
  - Self-controlled case series
- Patient-level prediction

Innovation

Transforming RWE generation from bespoke studies taking months to a **systematic process** taking days, while enabling transparent reproducibility and ensuring **scientific best practices** in causal inference and machine learning

Use cases

Current focus:

- Safety signal detection and evaluation
- Enhanced surveillance

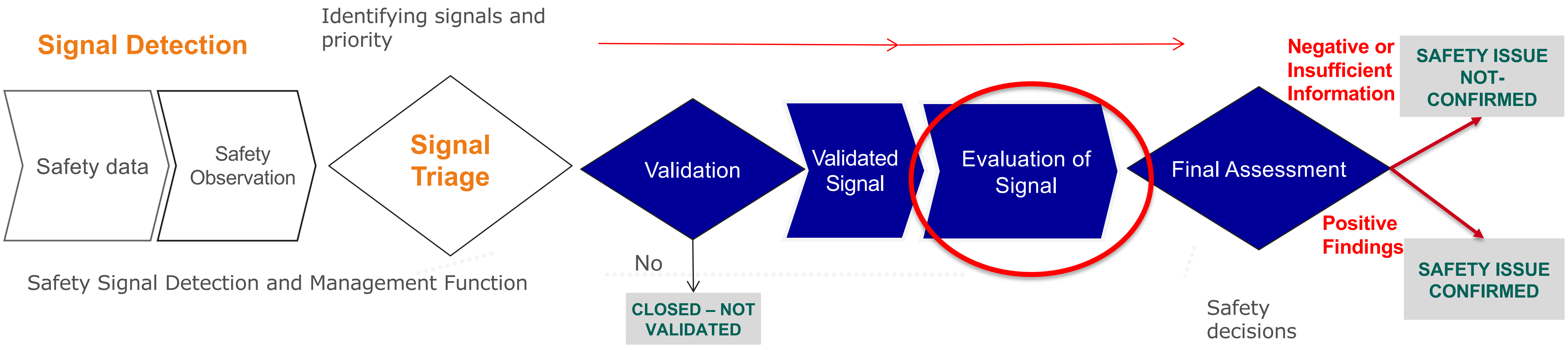
Future opportunities:

- Comparative effectiveness
- Disease interception

Results delivered in 2023

- 23 Requests
- Impact on regulatory decision making

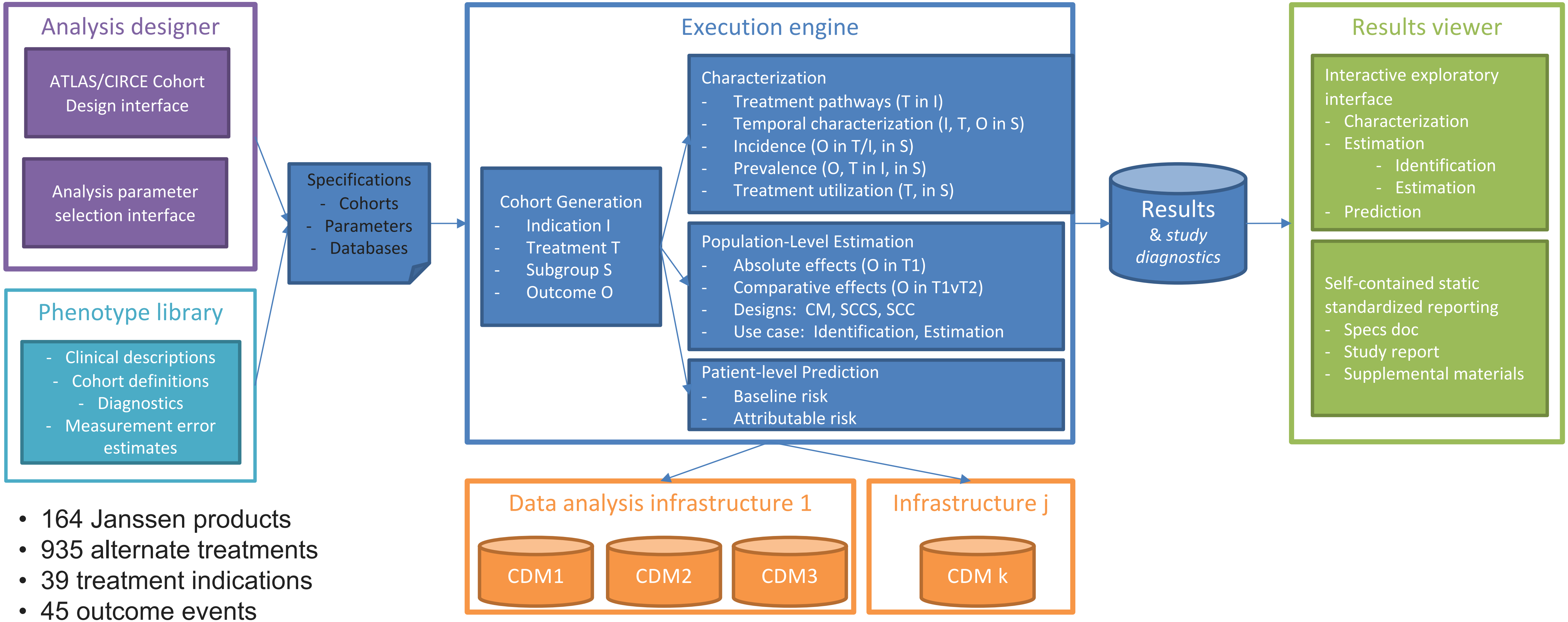
Where does ASSURE fit into the life of a safety signal?



- Early awareness of signals enables preparation and validation of input specifications
- Standardization enables evidence generation within a short timeline

Standardized inputs → Standardized analytics → Standardized databases → Standardized results

ASSURE Analyses: Inputs and Outputs



- 164 Janssen products
- 935 alternate treatments
- 39 treatment indications
- 45 outcome events

A Day in the Life of the ASSURE Team

1. Treatment/Comparator/Indication/Outcome
   - Comparator Selection Tool
2. Phenotype Development
   - Disease Advisory Board
3. Analytic Design and Implementation
   - Negative Control Selection
   - Time at Risk Selection
4. Result Interpretation
   - Shiny Dashboard
5. Documentation and Communication
   - Standardized Output

```r
tcis <- list(
  list(
    targetId = 13771,
    comparatorId = 13774,
    indicationId = NULL,
    genderConceptIds = c(8507, 8532), # use valid
    minAge = 18, # Age 18+. Can be NULL
    maxAge = NULL, # No max age. Can be NULL
    excludedCovariateConceptIds = c(1154029,
                                    1103640
    )

sccsTi <- list(
  list(
    targetId = 13771,
    indicationId = NULL, # NO INDICATION REQUIRED
    genderConceptIds = c(8507, 8532), # use valid
    minAge = 18, # Age 18+. Can be NULL
    maxAge = NULL # No max age. Can be NULL
  ))

outcomes <- tibble(
  cohortId = c(12308),
  cleanWindow = c(90)
)
```

```r
negativeConceptSetId <- 5749
timeAtRisks <- tibble(
  label = c("On-treatment"),
  riskWindowStart  = c(1),
  startAnchor = c("cohort start"),
  riskWindowEnd  = c(0),
  endAnchor = c("cohort end"),
)
# Try to avoid intent-to-treat TARs for SCCS:
sccsTimeAtRisks <- tibble(
  label = c("On-treatment"),
  riskWindowStart  = c(1),
  startAnchor = c("cohort start"),
  riskWindowEnd  = c(0),
  endAnchor = c("cohort end"),
)
# Try to use fixed-time TARs for PLP:
plpTimeAtRisks <- tibble(
  riskWindowStart  = c(1),
  startAnchor = c("cohort start"),
  riskWindowEnd  = c(365),
  endAnchor = c("cohort start"),
)
studyStartDate <- "" # YYYYMMDD
studyEndDate <- "" # YYYYMMDD
```

```r
# Probably don't change below this line ------------------------------------------
```

2023 OHDSI Global Symposium

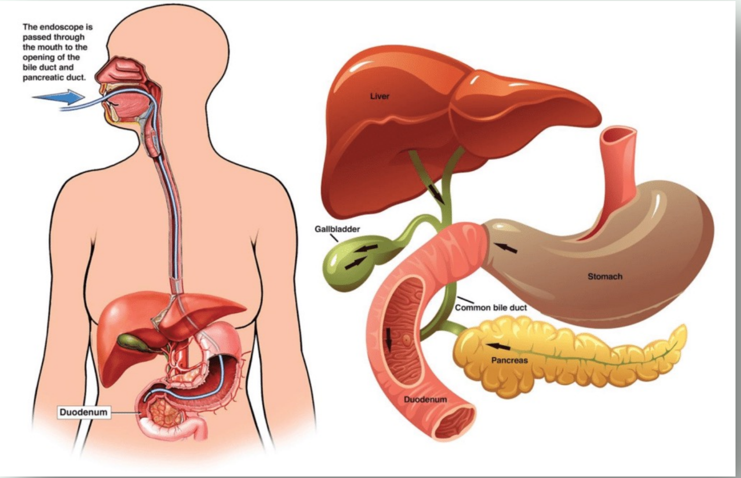
Patient's outcomes after endoscopic retrograde cholangiopancreatography (ERCP) using reprocessed duodenoscope accessories: a descriptive study using real-world data

Jessica Mayumi Maruyama
Eduardo Sleiman Beljavskis
Laila Colações
Lisandry Aquino
Renata Martins
Sarah Rodrigues
Suellen dos Santos
Julio Cesar Barbour Oliveira

Boston Scientific

PRECISION DATA
BRIDGING PEOPLE AND DATA

1. Background



Source: https://www.sages.org/

ERCP: Significant impact on management and prognosis of biliary and pancreatic diseases

Concerns related to duodenoscope-related infections due to material reprocessing

Study objective using an OMOP CDM harmonized dataset from Brazil:

- To compare the % of readmissions post-ERCP between Single-use (SUG) and Non-single-use (NSUG) institutions

2. Methods

Data source: Hospital and Ambulatory Information System from Brazilian Administrative Database, mapped to OMOP CDM v 5.4. A deterministic linkage algorithm was developed to connect hospitals with outpatient records using the key information of zip code, date of birth, and gender.

Study period: January 2020 to January 2023

- ✓ Patients with no history of cancer
- ✓ ERCP procedure, excluding due to sepsis, acute pancreatitis, or cholangitis

- ✓ Readmission within 30 day
- ✓ Causes for readmission: sepsis, acute pancreatitis, or cholangitis

3. Methods

Identification of ERCP procedures:

Specific SUS coding system, named Table of the Procedure, Medication, Orthotics, Prosthetics, and Special Materials Management System of the SUS (SIGTAP)

Identification of SUG and NSUG hospitals:

3 Single-use institutions

15 Non-single use institutions

Statistical analysis: Atlas

4. Results

Table 1. Descriptive information of total and readmitted patients in SUG and NSUG groups

	SUG		NSUG	
	Total	Readmitted patients	Total	Readmitted patients
N	669	20	887	43
Male (%)	30.9	50.0	34.0	37.0
Mean age (SD)	55.0 (19.0)	55.0 (17.9)	55.0 (19.0)	51.0 (14.9)

Note. SUG – single-use group; NSUG – non-single-use group; SD – standard deviation;
Readmitted patients included patients who were hospitalized within 30 days after a patient's ERCP due to sepsis, acute pancreatitis, or cholangitis.

In comparison to the readmitted patients from SUG, the readmitted patients from NSUG had a **higher proportion of female individuals and patients with a lower mean age**

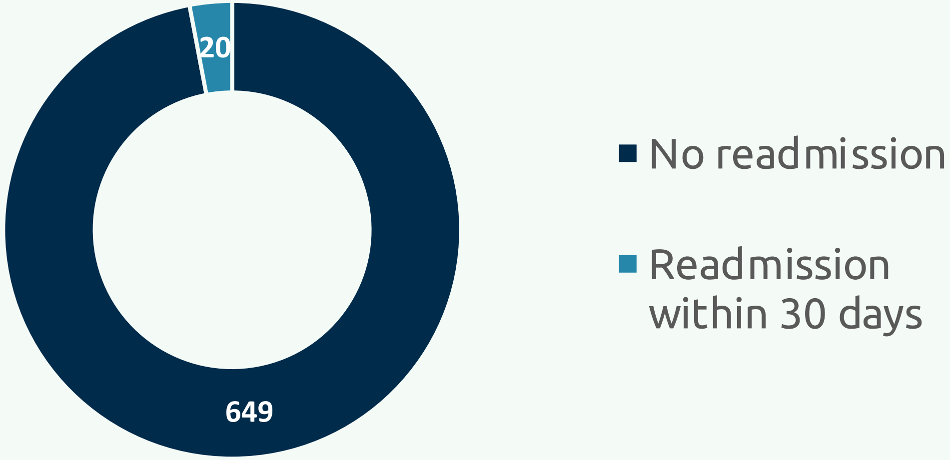
5. Results

Non-Single-Use (NSUG)

Readmission:
4.8% (43)

Single-Use (SUG)

Readmisson:
2.9% (20)

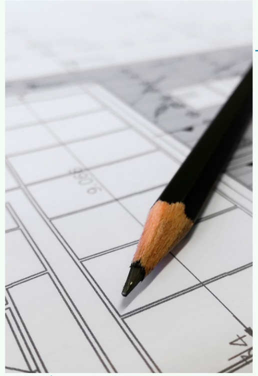


Difference between NSUG Group and SUG Group:
The NSUG group had a percentage of readmissions approximately 65% higher compared to the SUG group

6. Conclusion and next steps

Real-world data from Brazilian administrative dataset

Higher % of readmissions in NSUG institutions compared to SUG institutions



Next step: estimation study adjusting for confounders and unbalanced data

Inform clinical decision-making and optimal ERCP management practices

Connect with us:

hello@precisiondata.com.br
Telefone: 55 11 982497427

PRECISION DATA
BRIDGING PEOPLE AND DATA

precisiondata.ai

Department of Biostatistics, Epidemiology and Informatics

Does COVID-19 Increase Racial/Ethnic Differences in Prevalence of PASC/Long COVID in Children and Adolescents?

— Findings from Difference-in-Differences Analyses using an EHR-Based Cohort from the RECOVER Program

Bingyu Zhang

PhD student, University of Pennsylvania

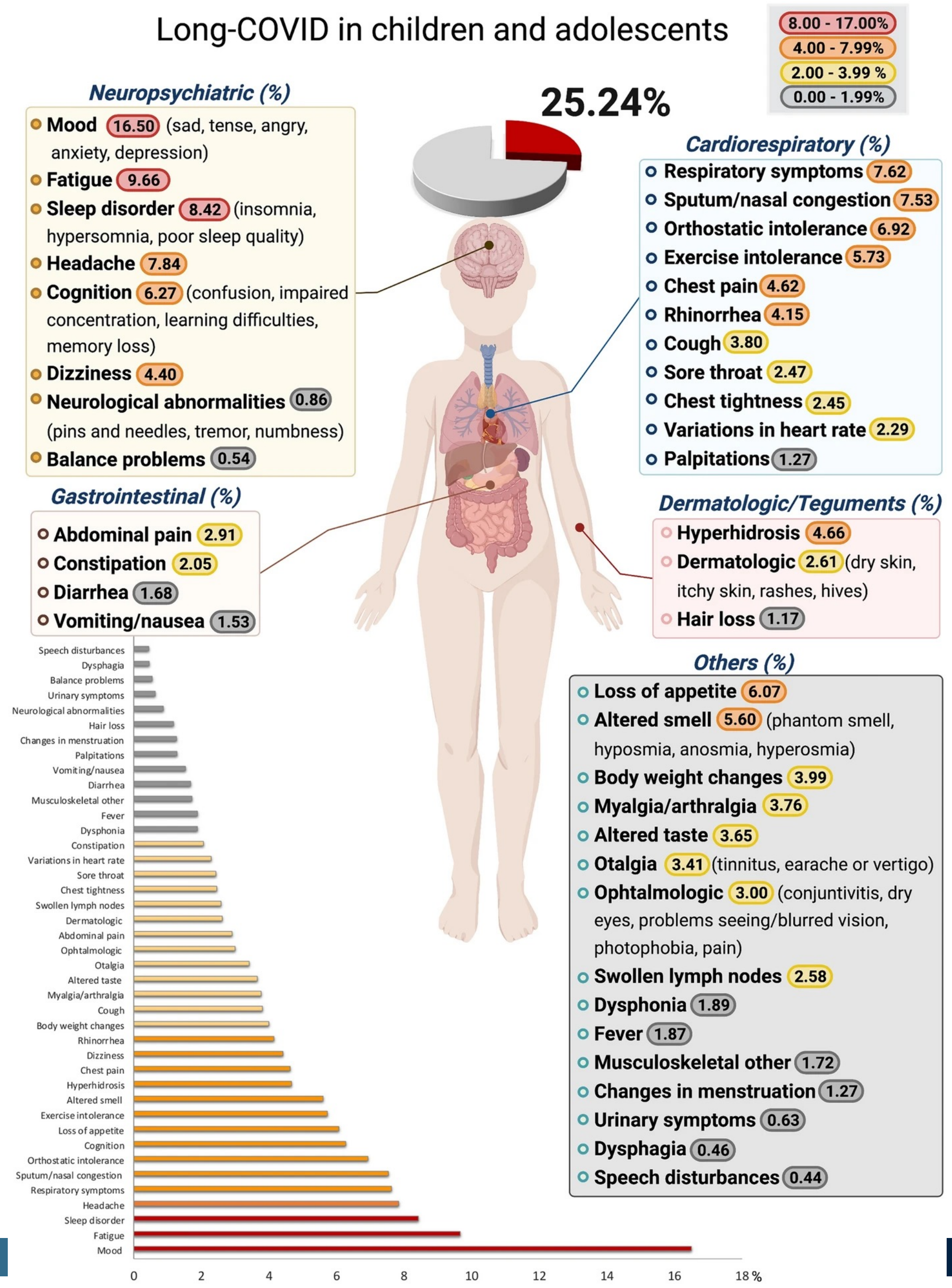
Advisor: Dr. Yong Chen

2023 OHDSI Symposium, October 20

PennCIL

A Computing · Inference · Learning lab at University of Pennsylvania

What is PASC?



https://recovercovid.org/long-covid
https://covid19community.nih.gov/what-you-need-to-know-about-long-covid

Lopez-Leon, S., Wegman-Ostrosky, T., Ayuzo del Valle, N.C. et al. Long-COVID in children and adolescents: a systematic review and meta-analyses. Sci Rep 12, 9950 (2022)

RECOVER: Researching COVID to Enhance Recovery

- The National Institutes of Health (NIH) created the RECOVER Initiative to learn about the long-term effects of COVID
- The goal of RECOVER is to rapidly improve our understanding of and ability to predict, treat, and prevent PASC

- PI for **pediatric** RECOVER:
  - Christopher Forrest (Children's Hospital of Philadelphia)
- PI for **adult** RECOVER:
  - Rainu Kaushal (Weill Cornell)
- **Biostatistics** Core Director:
  - Yong Chen
  - for PCORnet Pediatric RECOVER

https://recovercovid.org

Selected Publications on PASC within RECOVER

RECOVER
Researching COVID to Enhance Recovery

> Lancet Digit Health. 2022 Jul;4(7):e532-e541. doi: 10.1016/S2589-7500(22)00048-6. Epub 2022 May 16.

Identifying who has long COVID in the USA: a machine learning approach using N3C data

Emily R Pfaff [1], Andrew T Girvin [2], Tellen D Bennett [3], Abhishek Bhatia [4], Ian M Brooks [5], Rachel R Deer [6], Jonathan P Dekermanjian [7], Sarah Elizabeth Jolley [8], Michael G Kahn [9], Kristin Kostka [10], Julie A McMurry [11], Richard Moffitt [12], Anita Walden [11], Christopher G Chute [13], Melissa A Haendel [11]; N3C Consortium

> Nat Commun. 2023 Apr 7;14(1):1948. doi: 10.1038/s41467-023-37653-z.

Data-driven analysis to understand long COVID using electronic health records from the RECOVER initiative

Chengxi Zang [1], Yongkang Zhang [1], Jie Xu [2], Jiang Bian [2], Dmitry Morozyuk [1], Edward J Schenck [3], Dhruv Khullar [1], Anna S Nordvig [4], Elizabeth A Shenkman [2], Russell L Rothman [5], Jason P Block [6], Kristin Lyman [7], Mark G Weiner [1], Thomas W Carton [7], Fei Wang [8], Rainu Kaushal [1]

> JAMA Pediatr. 2022 Oct 1;176(10):1000-1009. doi: 10.1001/jamapediatrics.2022.2800.

Clinical Features and Burden of Postacute Sequelae of SARS-CoV-2 Infection in Children and Adolescents

Suchitra Rao [1], Grace M Lee [2], Hanieh Razzaghi [3], Vitaly Lorman [3], Asuncion Mejias [4], Nathan M Pajor [5], Deepika Thacker [6], Ryan Webb [3], Kimberley Dickinson [3], L Charles Bailey [3], Ravi Jhaveri [7], Dimitri A Christakis [8] [9], Tellen D Bennett [1], Yong Chen [10], Christopher B Forrest [3]

> Nat Commun. 2023 May 22;14(1):2914. doi: 10.1038/s41467-023-38388-7.

Long COVID risk and pre-COVID vaccination in an EHR-based cohort study from the RECOVER program

M Daniel Brannock [#] [1], Robert F Chew [#] [2], Alexander J Preiss [#] [2], Emily C Hadley [#] [2], Signe Redfield [3], Julie A McMurry [4], Peter J Leese [5], Andrew T Girvin [6], Miles Crosskey [7], Andrea G Zhou [8], Richard A Moffitt [9] [10], Michele Jonsson Funk [5], Emily R Pfaff [5], Melissa A Haendel [4], Christopher G Chute [11]; N3C; RECOVER Consortia

> MMWR Morb Mortal Wkly Rep. 2022 Apr 8;71(14):517-523. doi: 10.15585/mmwr.mm7114e1.

Cardiac Complications After SARS-CoV-2 Infection and mRNA COVID-19 Vaccination - PCORnet, United States, January 2021-January 2022

Jason P Block, Tegan K Boehmer, Christopher B Forrest, Thomas W Carton, Grace M Lee, Umed A Ajani, Dimitri A Christakis, Lindsay G Cowell, Christine Draper, Nidhi Ghildayal, Aaron M Harris, Michael D Kappelman, Jean Y Ko, Kenneth H Mayer, Kshema Nagavedu, Matthew E Oster, Anuradha Paranjape, Jon Puro, Matthew D Ritchey, David K Shay, Deepika Thacker, Adi V Gundlapalli

> Nat Med. 2023 Jan;29(1):226-235. doi: 10.1038/s41591-022-02116-3. Epub 2022 Dec 1.

Data-driven identification of post-acute SARS-CoV-2 infection subphenotypes

Hao Zhang [1], Chengxi Zang [1], Zhenxing Xu [1], Yongkang Zhang [1], Jie Xu [2], Jiang Bian [2], Dmitry Morozyuk [1], Dhruv Khullar [1], Yiye Zhang [1], Anna S Nordvig [3], Edward J Schenck [4], Elizabeth A Shenkman [2], Russell L Rothman [5], Jason P Block [6], Kristin Lyman [7], Mark G Weiner [1], Thomas W Carton [7], Fei Wang [8], Rainu Kaushal [1]

PennCIL
A Computing · Inference · Learning lab at University of Pennsylvania

Racial/ethnic Differences in PASC Prevalence

Millions of people have had COVID-19 — and in many ways, people of color have been hit hardest.

Studies show that **some groups and communities are more likely to go to the hospital** for health issues related to COVID-19. This is because people don't have equal access to health care and information about COVID. And some people live or work in places where they are more likely to catch COVID-19.

AMERICAN INDIAN OR ALASKA NATIVE 2.5

BLACK OR AFRICAN AMERICAN 2.1

HISPANIC OR LATINO 1.9 times more likely to go to the hospital

ASIAN 0.7 times more likely to go to the hospital

WHITE

J Gen Intern Med. 2023 Apr; 38(5): 1127–1136.
Published online 2023 Feb 16. doi: 10.1007/s11606-022-07997-1
PMCID: PMC9933823
PMID: 36795327

Racial/Ethnic Disparities in Post-acute Sequelae of SARS-CoV-2 Infection in New York: an EHR-Based Cohort Study from the RECOVER Program

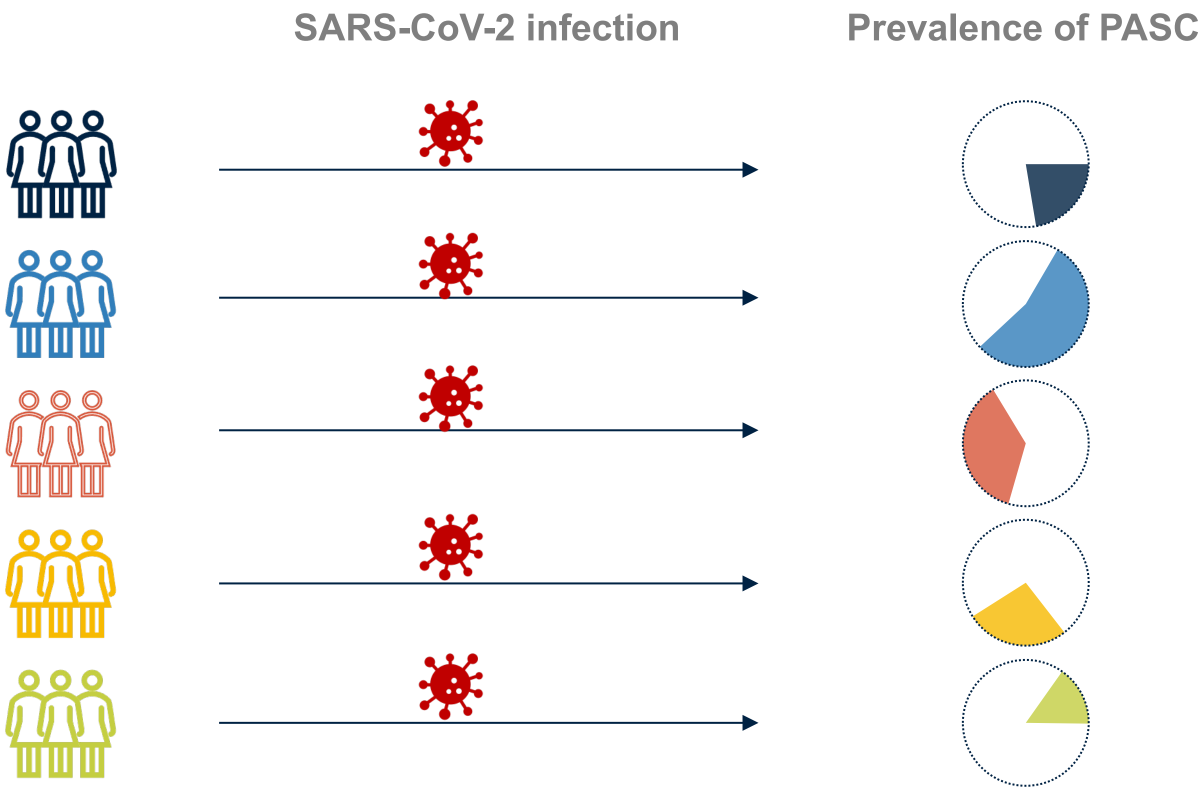
Dhruv Khullar, MD, MPP,[1,2] Yongkang Zhang, PhD,[1] Chengxi Zang, PhD,[1] Zhenxing Xu, PhD,[1] Fei Wang, PhD,[1] Mark G. Weiner, MD,[1] Thomas W. Carton, PhD,[3] Russell L. Rothman, MD, MPP,[4] Jason P. Block, MD, MPH,[5] and Rainu Kaushal, MD, MPH[1]

J Gen Intern Med

https://stacks.cdc.gov/view/cdc/105453

Clinical Question

- Does there exist **racial/ethnic differences** in **potential PASC symptoms and conditions** among children and adolescents after **SARS-CoV-2 infection**?



Typical Solutions

JOURNAL ARTICLE

Large-scale evidence generation and evaluation across a network of databases (LEGEND): assessing validity using hypertension as a case study

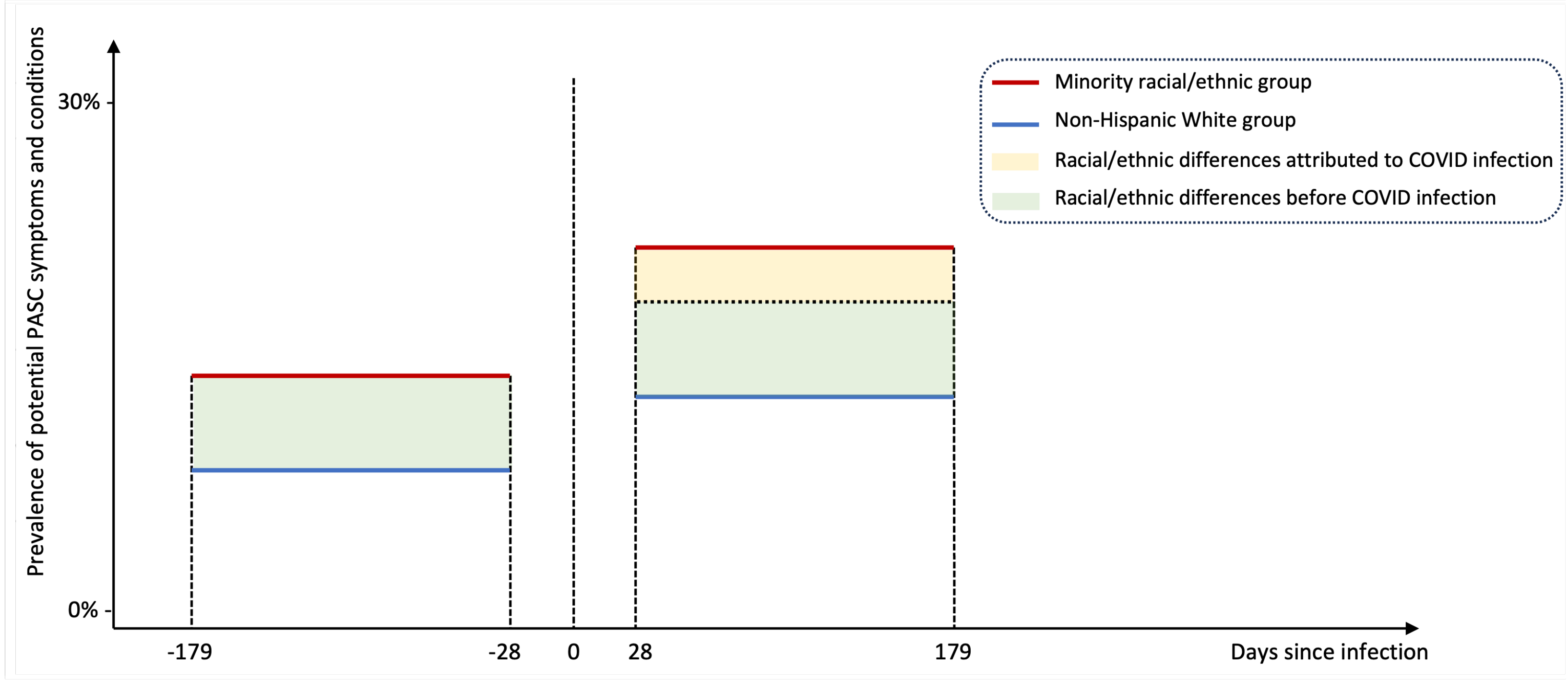
Martijn J Schuemie, Patrick B Ryan, Nicole Pratt, RuiJun Chen, Seng Chan You, Harlan M Krumholz, David Madigan, George Hripcsak, Marc A Suchard

Journal of the American Medical Informatics Association, Volume 27, Issue 8, August 2020, Pages 1268–1277, https://doi.org/10.1093/jamia/ocaa124

How many differences are attributable to COVID infection?

How Many Differences are Attributable to COVID Infection?

- Difference-in-differences approach



Proposed Solution

Standard regression model

- regression model adjusted for confounders

LEGEND pipeline

- **Step 1**: large-scale propensity score (LSPS) stratification/ matching/weighting
  - Fit LSPS model: Race/ethnicity ~ confounders
  - Stratify or match or weight on propensity scores
- **Step 2**: Outcome regression model
  - Regression model, with propensity score adjusted

Proposed method

- **Step 1**: large-scale propensity score (LSPS) stratification/ matching/weighting
  - Fit LSPS model: Race/ethnicity ~ confounders
  - Stratify or match or weight on propensity scores
- **Step 2**: Outcome regression model
  - **Difference-in-differences analyses to control pre-COVID racial/ethnic differences**
  - Regression model, with propensity score adjusted

PennCIL
A Computing · Inference · Learning lab at University of Pennsylvania

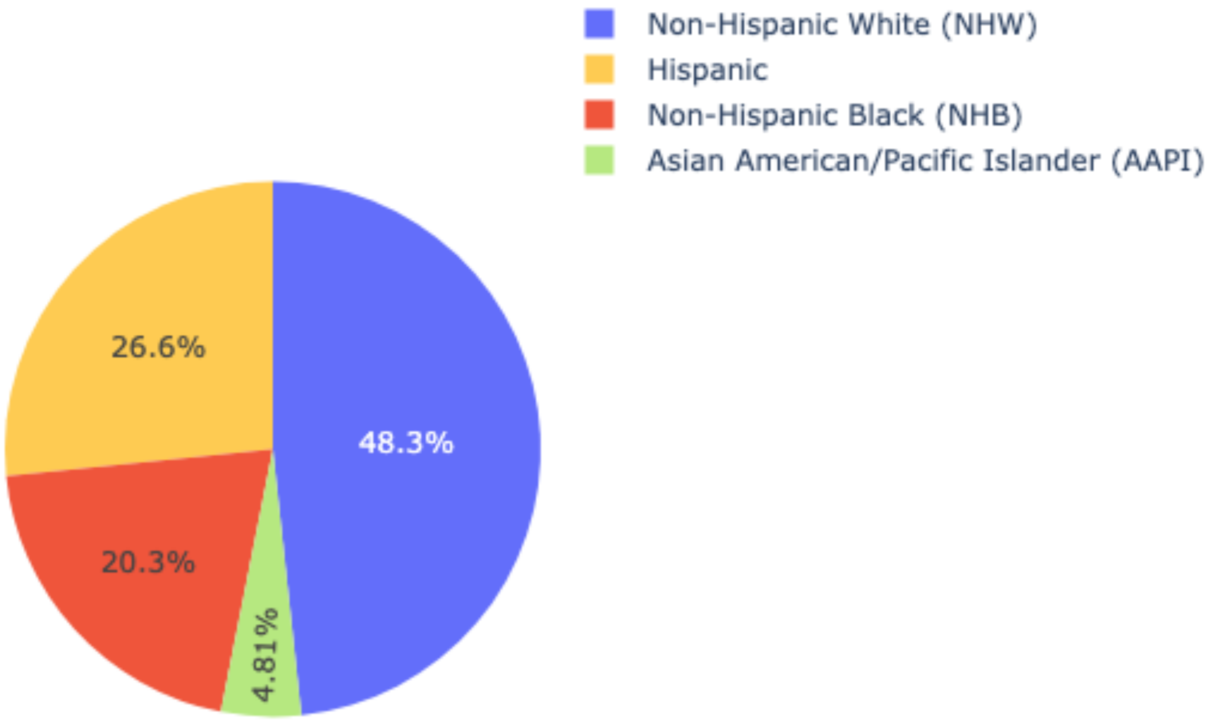
Study Cohort

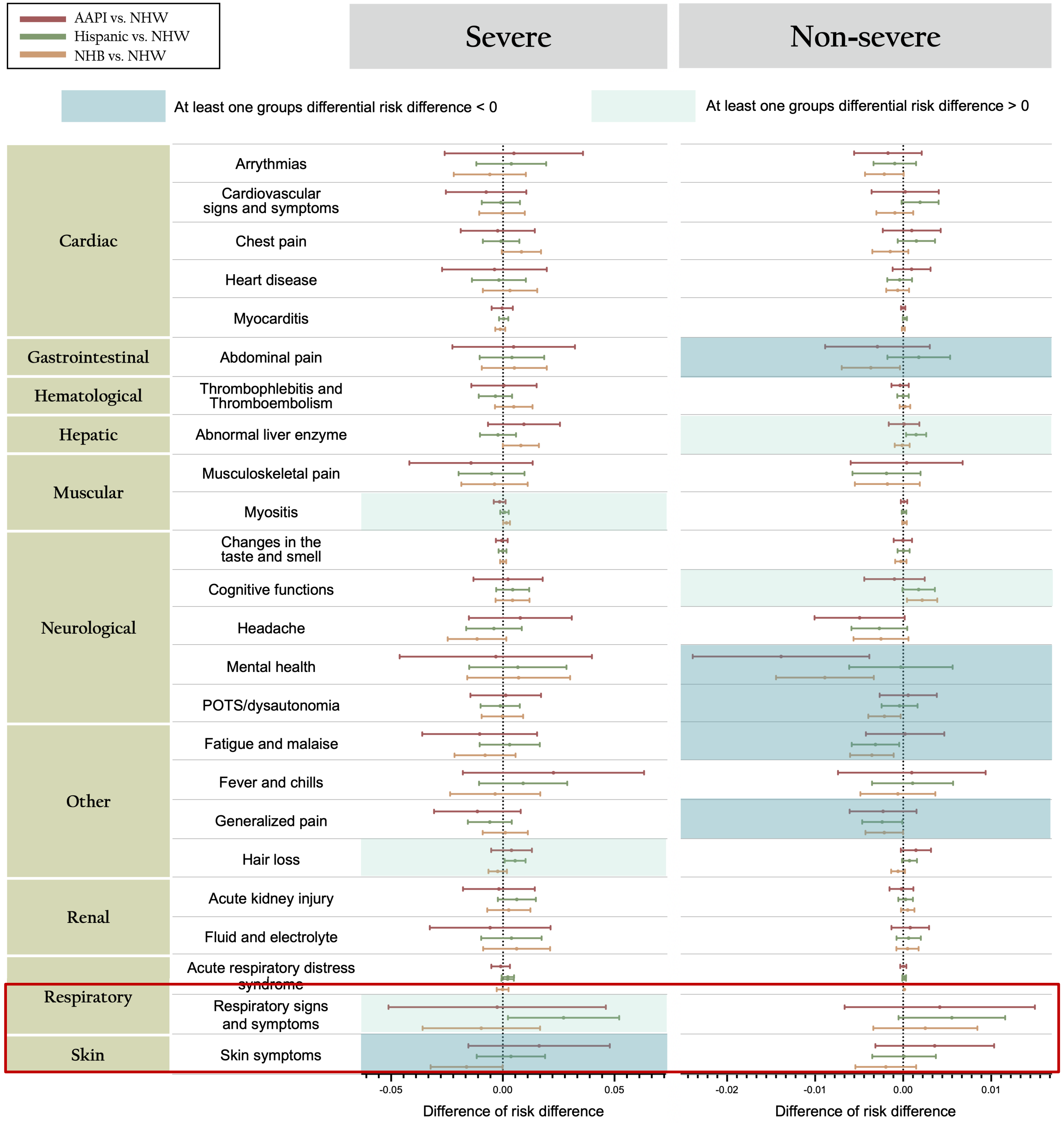
RECOVER
Researching COVID to Enhance Recovery

Inclusion Criteria

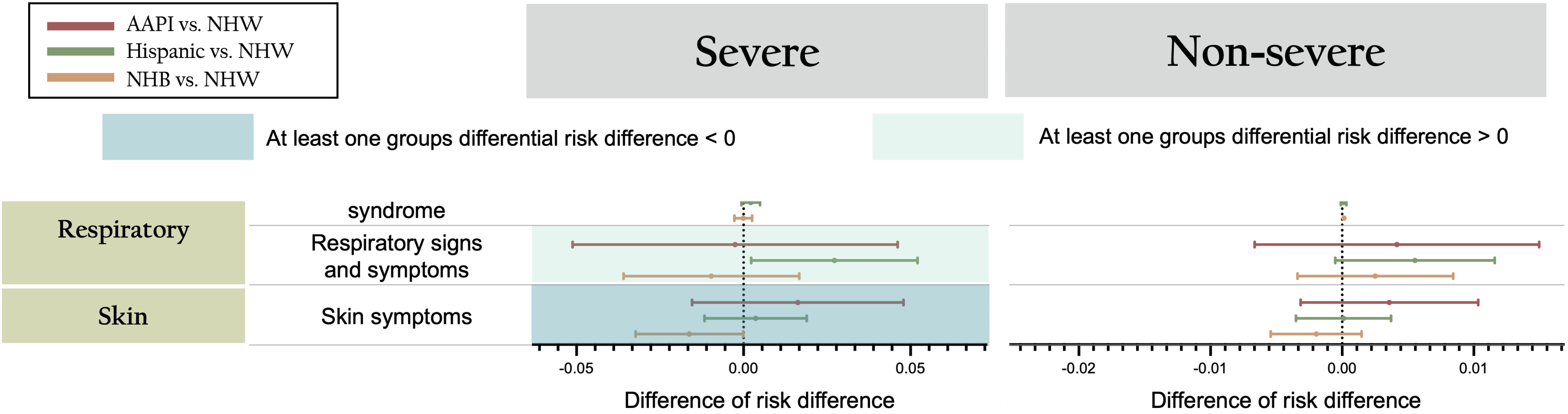
- Documented SARS-CoV-2 infection
- Age < 21 years
- Had at least one visit during the baseline period
- Had at least one visit during the follow-up period

225,723 patients across **13** institutions in the US





- Stratify by COVID-19 acute phase **severity**
- Three minority groups compare to Non-Hispanic White group respectively



AAPI vs. NHW
Hispanic vs. NHW
NHB vs. NHW
Severe
Non-severe
At least one groups differential risk difference < 0
At least one groups differential risk difference > 0
Respiratory
syndrome
Respiratory signs and symptoms
Skin
Skin symptoms
-0.05
0.00
0.05
Difference of risk difference
-0.02
-0.01
0.00
0.01
Difference of risk difference

Takeaways

- Help understand racial/ethnic differences in PASC after SARS-CoV-2 infection among children and adolescents

- Cover a broad spectrum of the US pediatric population

  LEGEND principle 1: Generate evidence at a large scale
  LEGEND principle 9: Generate evidence across a network of multiple databases

- Handle measured confounders using propensity score matching

- Control pre-COVID racial/ethnic differences using difference-in-differences analyses

  LEGEND principle 5: Generate evidence using best practices to minimize bias

- Future work
  - Explore methods to adjust for systematic bias

Acknowledgment

- Research Team
  - Dazheng Zhang^, Bingyu Zhang^, Qiong Wu, PhD, Ting Zhou, MD, Jiayi Tong, Yiwen Lu, Jiajie Chen, PhD, Deena J. Chisolm, PhD, Ravi Jhaveri, MD, Rachel C Kenney, PhD, Russell L Rothman, MD, MPP, Suchitra Rao, MD, David A. Williams, MD, Mady Hornig, MA, MD, Jeffrey S. Morris, PhD*, Christopher B. Forrest, MD, PhD*, and Yong Chen, PhD*
  - ^ co-first author
  - * senior author

OHDSI
OBSERVATIONAL HEALTH DATA SCIENCES AND INFORMATICS

RECOVER
Researching COVID to Enhance Recovery

👉 **Poster # 509**

👉 **Fri 10/20 4:15-5 pm**

Eye Care and Vision Research Workgroup: First Year Update

Michelle R. Hribar, PhD
Kerry E. Goetz, PhDc
Sally L. Baxter, MD, MSc
Eye Care and Vision Research Workgroup

OHDSI 2023 Global Symposium: October 20, 2023

Eye Care and Vision Research Workgroup
Our Journey

Getting Started

- OHDSI Eye Care and Vision Research Workgroup was started in spring 2022
  - Members of American Academy of Ophthalmology (AAO) Data Standards Workgroup identified need for ophthalmic data elements in the OMOP common data model
  - Ophthalmic concepts in source terminologies had not been updated consistently in over a decade
- Goals
  - Create access to large diverse datasets of ophthalmic and systemic data
  - Enable research in vision and systemic health

Initial Steps

- Created subgroups for tasks & subspecialties
    - Tasks: Concept mapping, visual acuity concept mapping, visual impairment phenotype, image integration, ETL scripts
    - Subspecialties: Glaucoma, retina, pediatrics/strabismus, uveitis
- Recruited colleagues to participate
- National Eye Institute (NEI) at National Institutes of Health (NIH) hired DATA Scholar to manage the project

Milestones

- Membership
    - 122 total, ~40 active
    - 13 trainees, 10 AI-READI (Bridge2AI) interns
    - Ophthalmologists, optometrists, informaticists, vision scientists
- Meetings
    - 17 Teams workgroup meetings
    - 3 in person meetings
    - ~42 subgroup meetings
    - Countless ad-hoc meetings

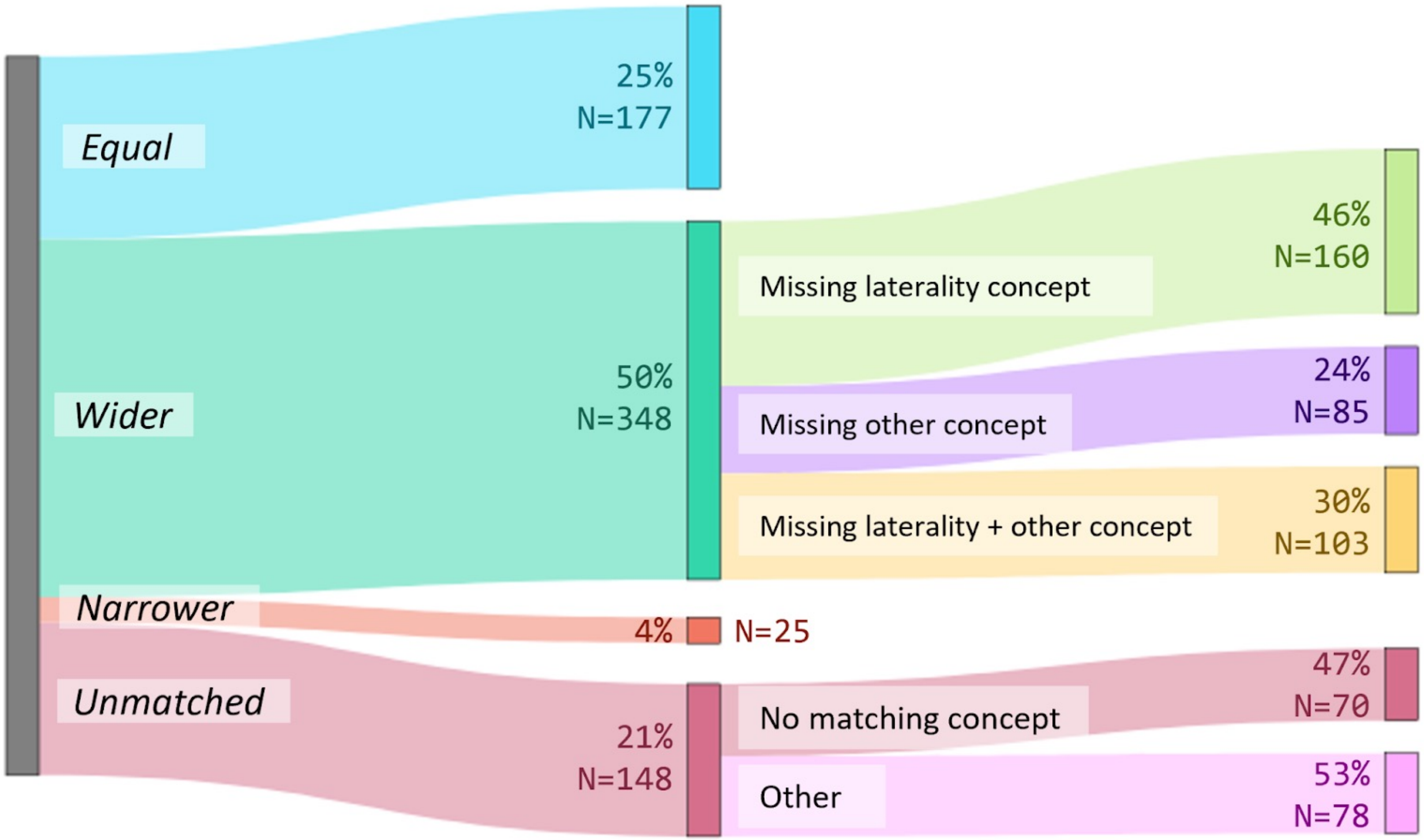
Milestones

- Collaborations
  - 9 OHDSI workgroups
  - 10 external groups including:
    - American Academy of Ophthalmology (AAO)
    - Association for Research in Vision and Ophthalmology (ARVO)
    - National Eye Institute
    - NIH Bridge2AI
    - NIH All of Us
    - SNOMED International and LOINC

Milestones

- Data Concepts
  - >3700 ophthalmic data elements analyzed & mapped
  - 11 retina condition codes submitted to SNOMED International
  - 224 visual acuity concepts submitted to LOINC
  - Glaucoma concepts currently in discussion with SNOMED International

Epic EHR Concept Matches



Cai C.X., Halfpenny W., Boland M.V., Lehmann H.P., Hribar M., Goetz K.E. & Baxter S.L., Advancing toward a common data model in ophthalmology: gap analysis of general eye examination concepts to standard OMOP concepts, Ophthalmology Science (2023), doi: https://doi.org/10.1016/j.xops.2023.100391.

Milestones

- Phenotypes
  - 3 visual impairment
  - 6 uveitis*
  - 3 new anti-VEGF users*
  - 1 blinding disease*
  - 5 diabetic retinopathy

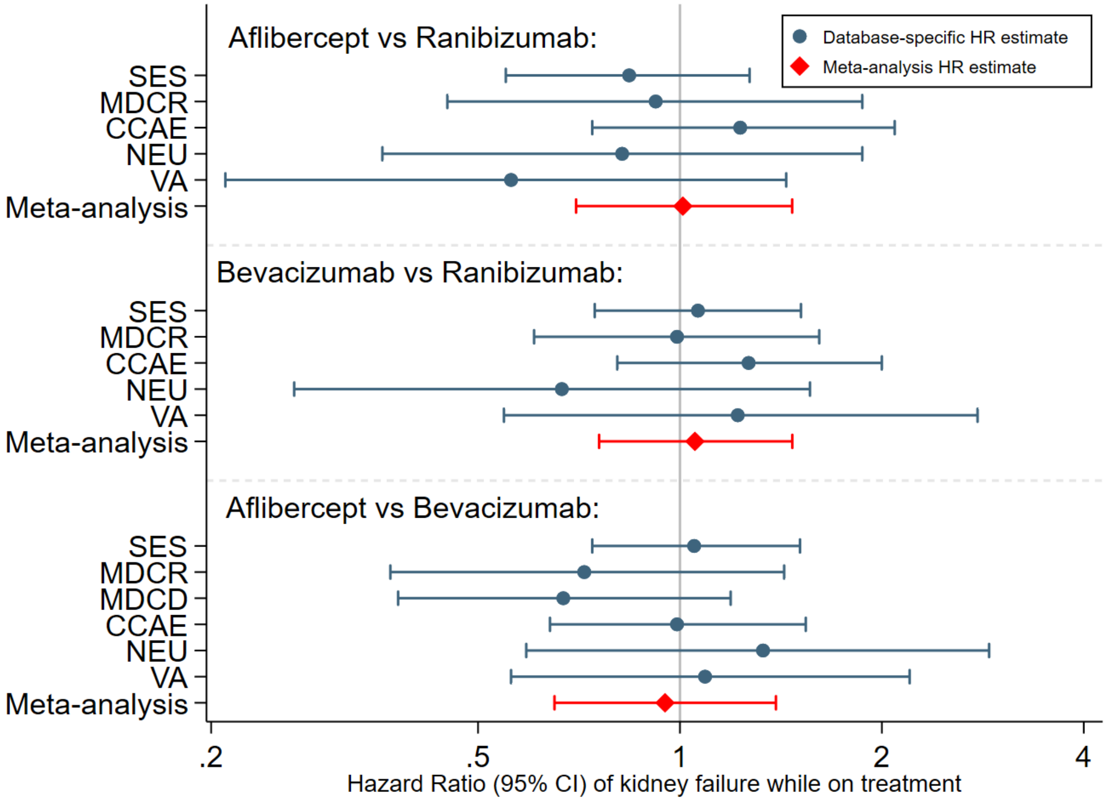
*Submitted to How Often

Milestones

- Publications
  - 9 papers, 4 EyeWiki pages
  - 5 more in progress
- Presentations
  - 18 talks, 5 posters
- Support
  - 1 NEI/NIH Data Scholar
  - 2 Grant submissions

Milestones

- SOS Challenge 2023
  - Led by Cindy X. Cai MD MS from Johns Hopkins University
  - Comparison of 3 anti-VEGF agents for risk of kidney injury when injected intravitreally
  - Results: no increased risk for kidney injury in any pairwise comparisons
  - Manuscript is in process



Next Steps

- Pilot at test sites
    - Image integration
    - Concept mappings (prioritized set)
- More eye care and vision research community outreach and education
- More network studies
- More funding support

First Year Challenges

Schedules

- Clinic schedules
- Time zones

Concept Modifiers

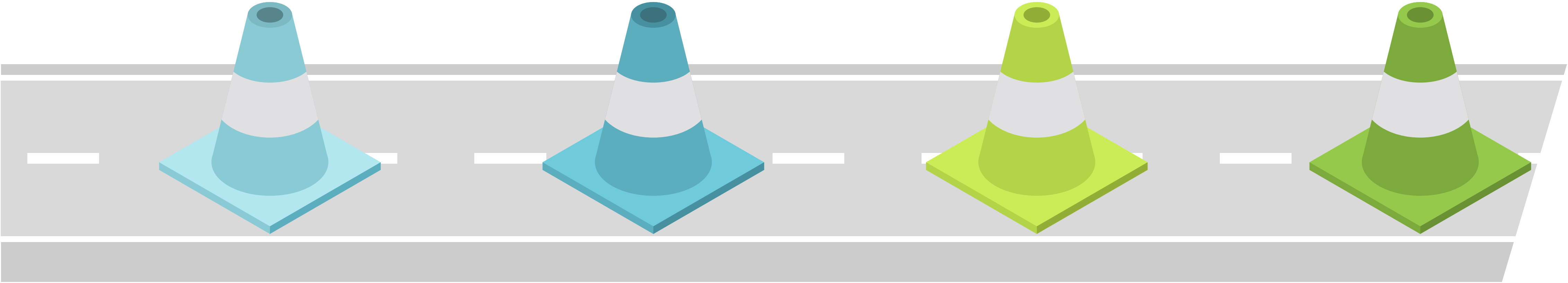
- Measurements often have multiple modifiers
- Pre-coordination results in thousands of concepts

Resources

- All volunteer effort
- DATA Scholar position is only 2 years

Diversity

- Members are from academic medical centers
- Need more diverse partners



Summary

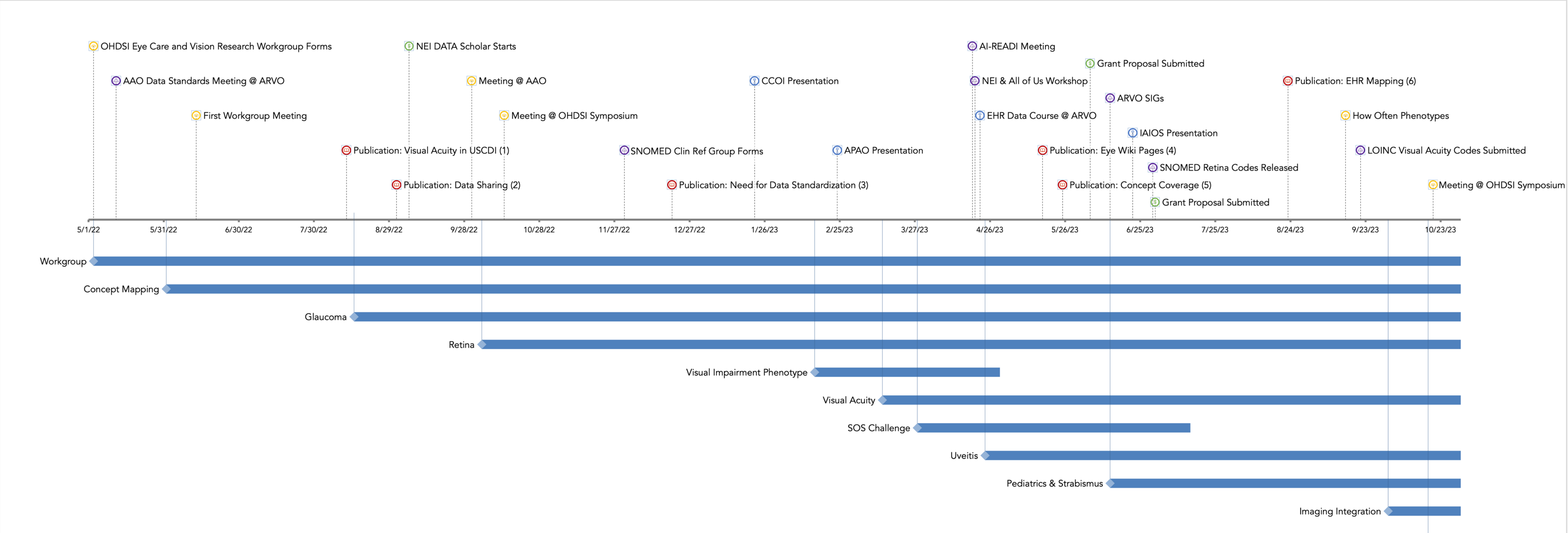
- Eye Care and Vision Research Workgroup had a productive year
- Working towards goal of including ophthalmic data and imaging in the OMOP common data model
- Still much more work to do—come join us!

Workgroup meeting is Sunday, Oct. 22 at 1 – 5 pm.

Thank you!

- OHDSI Community
  - Clair Blacketer, Paul Nagy, Elisse Katzman, Nathan Hall, Patrick Ryan, Craig Sachson, Anna Ostropolets
  - SOS Challenge collaborators
- Eye Care and Vision Research Workgroup
  - Co-leads: Kerry Goetz, Sally Baxter
  - Subgroup leads: Cindy Cai, Gayathri Srinivasan, Brian Stagg, Kavi Thakoor, Brian Toy
  - All of our wonderful members!
- National Eye Institute and National Institutes of Health

Eye Care and Vision Research Timeline



Poster 510