



# OHDSI Reproducibility Challenge

Anna Ostropolets  
PhD Student,  
Columbia University  
On behalf of the OHDSI  
Reproducibility Challenge Team



# Thank you, Yasser Albogami



Diabetes Care Volume 44, June 2021

1



## Glucagon-Like Peptide 1 Receptor Agonists and Chronic Lower Respiratory Disease Exacerbations Among Patients With Type 2 Diabetes

Yasser Albogami,<sup>1,2,3</sup> Kenneth Cusi,<sup>4</sup>  
Michael J. Daniels,<sup>5</sup> Yu-Jung J. Wei,<sup>1,2</sup> and  
Almut G. Winterstein<sup>1,2</sup>

Diabetes Care 2021;44:1–9 | <https://doi.org/10.2337/DC20-1794>

### OBJECTIVE

Emerging data from animal and human pilot studies suggest potential benefits of glucagon-like peptide 1 receptor agonists (GLP-1RA) on lung function. We aimed to assess the association of GLP-1RA and chronic lower respiratory disease (CLRD) exacerbation in a population with comorbid type 2 diabetes (T2D) and CLRD.

### RESEARCH DESIGN AND METHODS

A new-user active-comparator analysis was conducted with use of a national claims database of beneficiaries with employer-sponsored health insurance spanning 2005–2017. We included adults with T2D and CLRD who initiated GLP-1RA or dipeptidyl peptidase 4 inhibitors (DPP-4i) as an add-on therapy to their antidiabetes regimen. The primary outcome was time to first hospital admission for CLRD. The secondary outcome was a count of any CLRD exacerbation associated with an inpatient or outpatient visit. We estimated incidence rates using inverse probability of treatment weighting for each study group and compared via risk ratios.

### RESULTS

The study sample consisted of 4,150 GLP-1RA and 12,540 DPP-4i new users with comorbid T2D and CLRD. The adjusted incidence rate of first CLRD admission during follow-up was 10.7 and 20.3 per 1,000 person-years for GLP-1RA and DPP-4i users, respectively, resulting in an adjusted hazard ratio of 0.52 (95% CI 0.32–0.85). For the secondary outcome, the adjusted incidence rate ratio was 0.70 (95% CI 0.57–0.87).

### CONCLUSIONS

GLP-1RA users had fewer CLRD exacerbations in comparison with DPP-4i users. Considering both plausible mechanistic pathways and this real-world evidence, potential beneficial effects of GLP-1RA may be considered in selection of an antidiabetes treatment regimen. Randomized clinical trials are warranted to confirm our findings.

<sup>1</sup>Department of Pharmaceutical Outcomes & Policy, College of Pharmacy, University of Florida, Gainesville, FL

<sup>2</sup>Center for Drug Evaluation and Safety, University of Florida, Gainesville, FL

<sup>3</sup>Department of Clinical Pharmacy, College of Pharmacy, King Saud University, Riyadh, Saudi Arabia

<sup>4</sup>Division of Endocrinology, Diabetes and Metabolism, University of Florida, Gainesville, FL

<sup>5</sup>Department of Statistics, University of Florida, Gainesville, FL

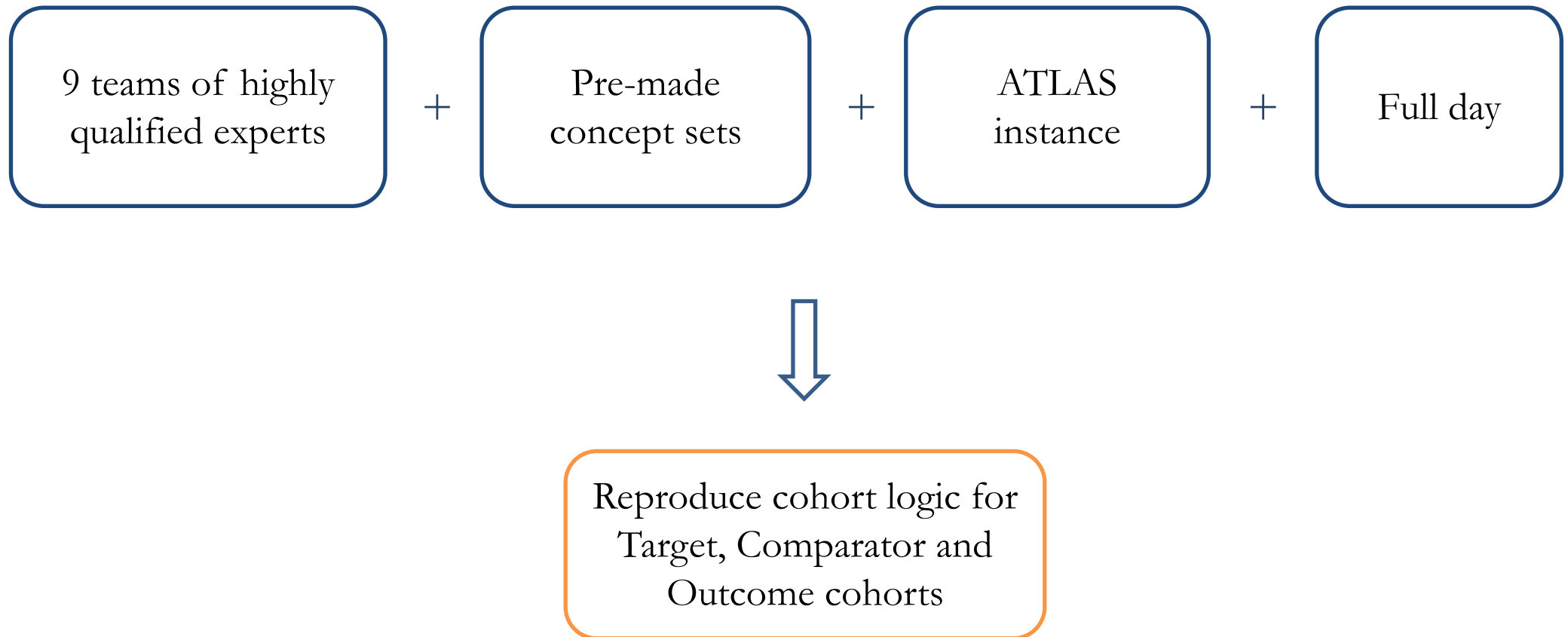
Corresponding author: Yasser Albogami, [y.albogami@ksu.edu.sa](mailto:y.albogami@ksu.edu.sa)

Received 17 July 2020 and accepted 1 March 2021



# OHDSI Reproducibility Experiment

OHDSI 2021 Symposium





# Focus: Target cohort – GLP1-RA users

## Study Sample

The study sample comprised patients aged >17 years who have had at least one inpatient or two outpatient encounters with T2D and CLRD, defined based on the presence of diagnoses or medication dispensing (Supplementary Tables 1 and 2) during the year before index date. Exposure assignment and start of follow-up (index date) were defined based on the first observed prescription for GLP-1RA and dipeptidyl peptidase 4 inhibitors (DPP-4I) (whichever came first) in the database. Pa-

tients whose first prescription fill did not follow at least 365 days' continuous health plan enrollment were excluded to achieve a new-user design. We included patients who used either study drug class as an add-on to other antidiabetes agents, e.g., metformin. Patients using insulin prior to the index date or with a diagnosis of type 1 diabetes were excluded. Patients with a history of cystic fibrosis, lung cancer, pulmonary embolism, and pulmonary hypertension, which may independently worsen lung

function, were excluded, as were women with pregnancy at index date. Patients with thyroid carcinoma who would likely be channeled to DPP-4I agents because of the related GLP-1RA contraindication were also excluded. Finally, patients with clinical conditions (except chronic obstructive pulmonary disease [COPD] and asthma) that frequently require chronic systemic corticosteroid treatment, which is protective against lung exacerbations, were excluded.





# Master implementation

Concept sets were discussed with the author and provided to the teams

## 1. Cohort entry

### Cohort Entry Events

Events having any of the following criteria:

a drug exposure of **GLP-1RA without insulin**  
✗ occurrence start is: between 2006-01-01 and 2017-12-31

with continuous observation of at least 365 days before and 0 days after event index date

Limit initial events to: earliest event per person.

## 3. Cohort exit

### Cohort Exit

#### Event Persistence:

Event will persist until: end of a continuous drug exposure

#### Continuous Exposure Persistence:

Specify a concept set that contains one or more drugs. A drug era will be derived from all drug exposure events within the window to the final exposure event. If no exposure event end date is provided, then an exposure event ends greater than the drug era end date.

Concept set containing the drug(s) of interest:

**GLP-1RA without insulin**

- Persistence window: allow for a maximum of 63 days between exposure records when inferring
- Surveillance window: add 1 days to the end of the era of persistence exposure as an additional
- Use days supply and exposure end date for exposure duration. [change](#)

#### Censoring Events:

Exit Cohort based on the following criteria:

a drug exposure of **DPP-4I**

### Cohort Eras

- Specify era collapse gap size: 0 days
- Left censor cohort start dates to 2005-01-01 ✗
- Right censor cohort end dates to 2017-12-31 ✗

## 2. Exclusion and inclusion criteria (10)

1. no GLP1 exposure in 365d prior to index

2. age > 17

3. has no type 1 diabetes in 365d prior to index

4. has type 2 diabetes mellitus in 365d prior to index

5. has CLRD in prior 365d prior to or on index based 1 IP dx OR 2 OP dx days OR (1 OP dx AND 1 rx)

6. no conditions of exclusion in 365d prior to index

7. no pregnancy observations in 180d prior to or after index

8. no prior DPP4 exposure in 365d prior to or on index

9. no prior insulin exposure in 365d prior to or on index

10. must have other T2DM drug that starts before index and ends on or after index

with at least 1 using all occurrences of:  
a condition occurrence of Chronic lower respiratory dis...  
✗ occurrence start is: between 2006-01-01 and 2017-12-31  
✗ having all of the following criteria:

with at least 1 using all occurrences of:  
a visit occurrence of Inpatient or Inpatient/ER visit  
where event starts between  
All days Before 0 days Before index start date  
✗ and event ends between 0 days After All days After index start date  
The index date refers to the condition occurrence of Chronic lower respiratory disease (CLRD).  
☐ restrict to the same visit occurrence  
☐ allow events from outside observation period

where event starts between  
365 days Before 0 days After index start date [add additional constraint](#)  
The index date refers to the event from the Cohort Entry criteria.  
☐ restrict to the same visit occurrence  
☐ allow events from outside observation period

or with at least 2 using distinct Start Date  
a condition occurrence of Chronic lower respiratory dis...  
✗ occurrence start is: between 2006-01-01 and 2017-12-31  
where event starts between  
365 days Before 0 days After index start date [add additional constraint](#)  
The index date refers to the event from the Cohort Entry criteria.  
☐ restrict to the same visit occurrence  
☐ allow events from outside observation period

or having all of the following criteria:

with at least 1 using all occurrences of:  
a condition occurrence of Chronic lower respiratory dis...  
✗ occurrence start is: between 2006-01-01 and 2017-12-31  
where event starts between  
365 days Before 0 days After index start date [add additional constraint](#)  
The index date refers to the event from the Cohort Entry criteria.  
☐ restrict to the same visit occurrence  
☐ allow events from outside observation period

and with at least 1 using all occurrences of:  
a drug exposure of Drugs for CLRD  
✗ occurrence start is: between 2006-01-01 and 2017-12-31  
where event starts between  
365 days Before 0 days After index start date [add additional constraint](#)  
The index date refers to the event from the Cohort Entry criteria.  
☐ restrict to the same visit occurrence  
☐ allow events from outside observation period

Delete Group



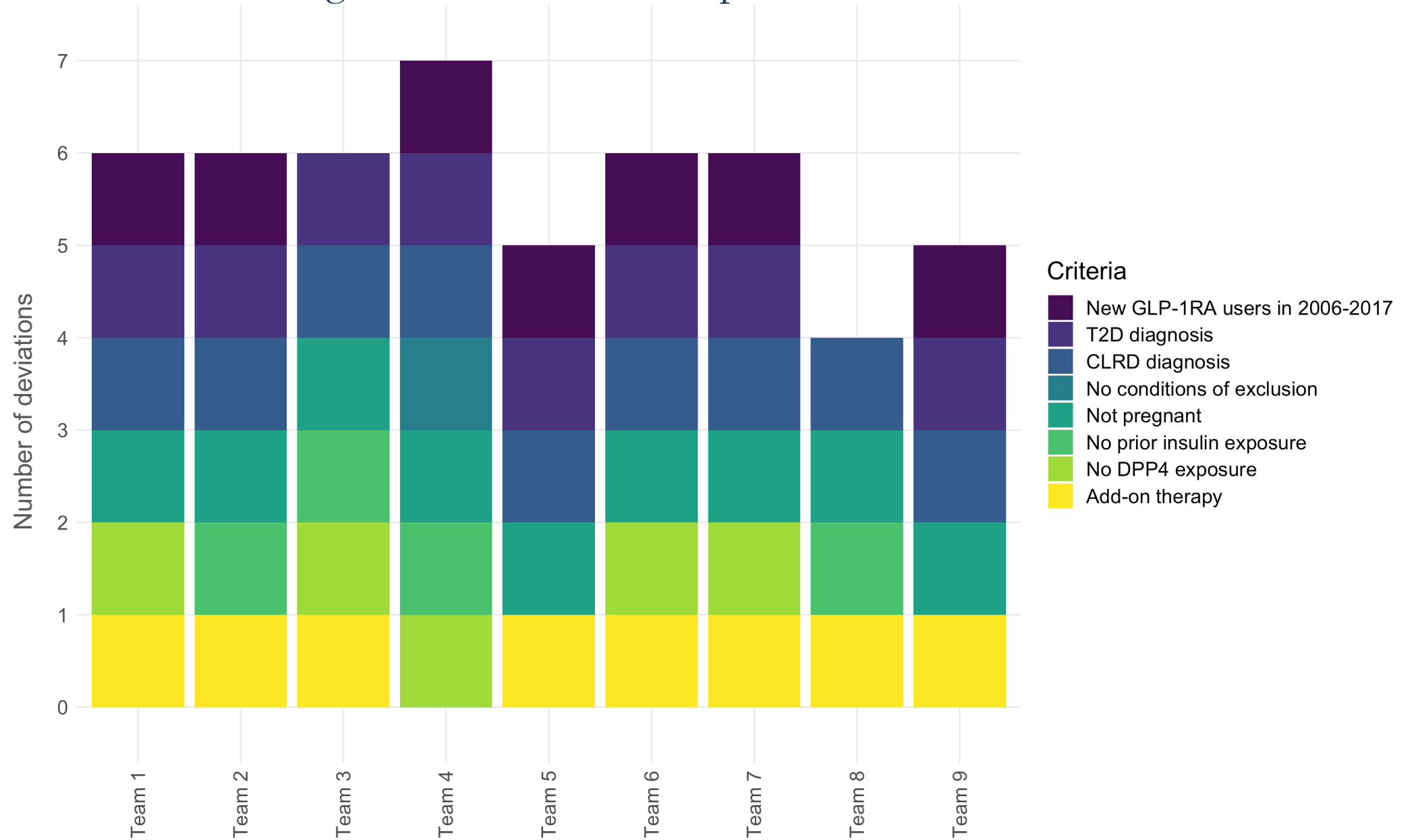
## Teams' implementations

9 different GLP1-RA user cohorts

All teams were confident they were able to translate the logic of the paper into a cohort definition



On average, the teams did not reproduce 60% of the criteria





# Only the simplest criteria are easy to reproduce

Easy to reproduce

Complex criteria with multiple sub-elements (conditions, drugs, visits)

Not a simple criteria but a complex phenotype

Variable temporal logic

New GLP-1RA users in 2006-2017

365 days of prior observation

Age > 17 at the index date

T2D diagnosis

CLRD diagnosis

No conditions of exclusion

Not pregnant

No prior insulin exposure

No DPP4 exposure

Add-on therapy

Number of teams with deviations

0 1 2 3 4 5 6 7 8 9

78%

0%

0%

89%

100%

11%

100%

44%

55%

89%

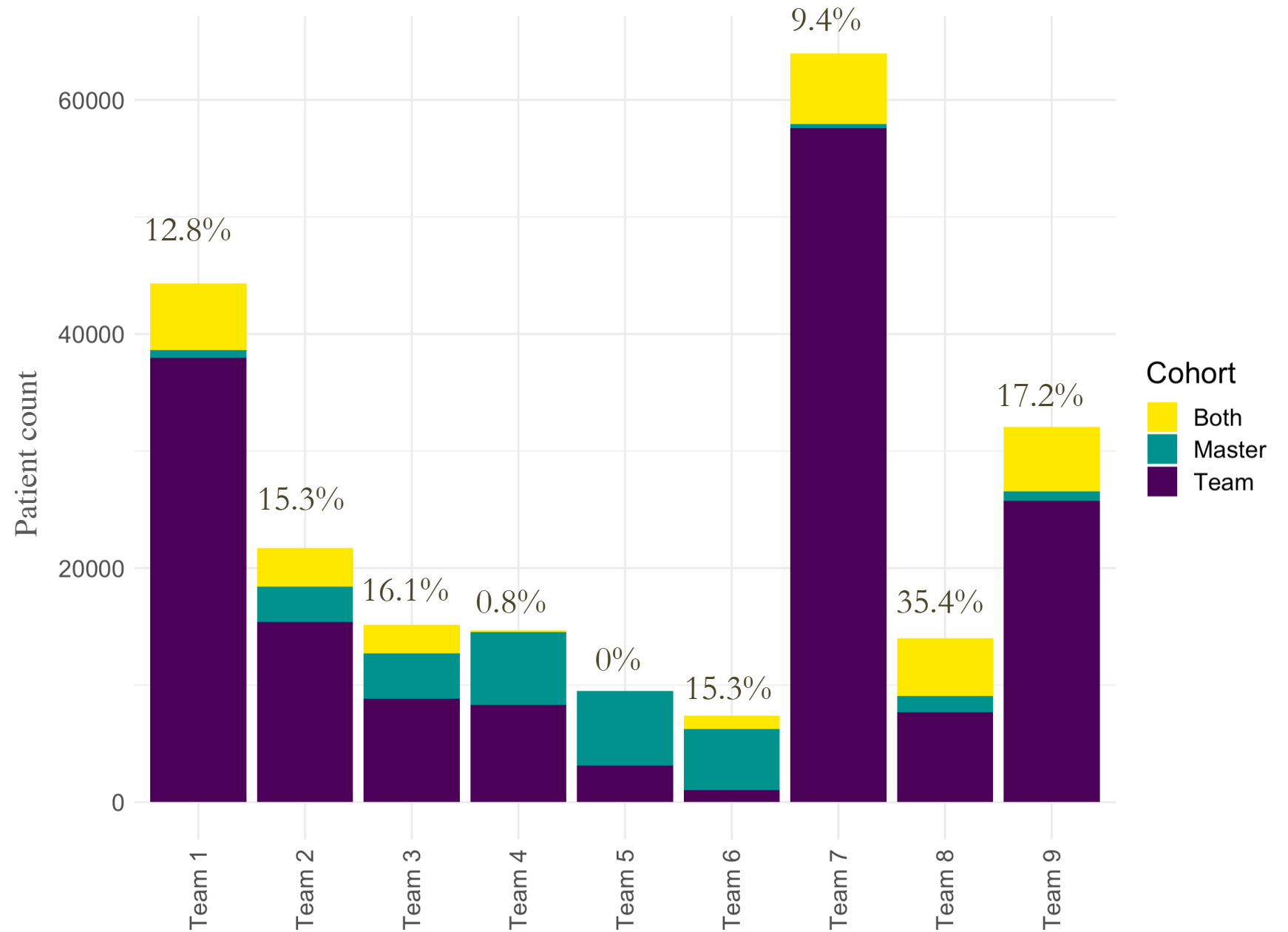




Influence on patient selection



Average agreement between the master cohort and teams' implementations (patient overlap) was low and varied from 0% to 35.4%



\* Agreement was calculated as number of patients in both cohorts divided by the number of patients in either cohort



# Compared to the master, teams' cohorts are different in baseline prevalence of CLRD and antidiabetic drug exposure



\* Blank spaces correspond to standardized difference of means < 0.1



# Conclusions

Researchers tend to underestimate the amount of information needed to reproduce a study

Design choices cannot be reproduced using free-text alone.  
Supplements and figures may introduce additional ambiguity

Sharing analytical code supported by a common data model allows reproducing a study in an unambiguous way