COLUMBIA
COLUMBIA UNIVERSITY
DEPARTMENT OF
BIOMEDICAL INFORMATICS

Factors Influencing Background Incidence Rates: Systematic Empirical Evaluation Across an International Network of Observational Databases

Anna Ostropolets

Columbia University Medical Center

Acknowledgement

Thanks to all co-authors and collaborators:

Xintong Li, Rupa Makadia, Gowtham Rao, Peter R. Rijnbeek, Talita Duarte-Salles, Martijn Schuemie, Anthony G. Sena, Azza Shaoibi, Marc A. Suchard, Patrick B. Ryan, Daniel Prieto-Alhambra and George Hripcsak

Factors Influencing Background Incidence Rate Calculation: Systematic Empirical Evaluation Across an International Network of Observational Databases. Frontiers in Pharmacology; 2022

Drug (vaccine) safety studies

Observed rate of adverse event

FDA's Adverse Event Reporting System
Adverse events following exposure in EHR and claims data

vs

Expected rate of adverse event

Background (baseline, historic) rates of adverse events in EHR and claims data

Background incidence rates

New cases

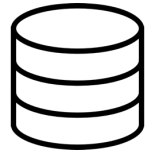
Time at risk (TAR)

$$\text{Background incidence rate} = \frac{\text{New cases during specified time period}}{\text{Person-time of the at-risk population}}$$

Population

Estimates of baseline incidence of ***stroke*** in different studies: **4.6 – 679** per 100,000 person-years depending on the population, time-at-risk, data source.

Systematic experiment



New cases **x** Population **x** Time at risk (TAR) **x** 12 data sources

15 adverse events (Brighton list)
2 new case identification strategies

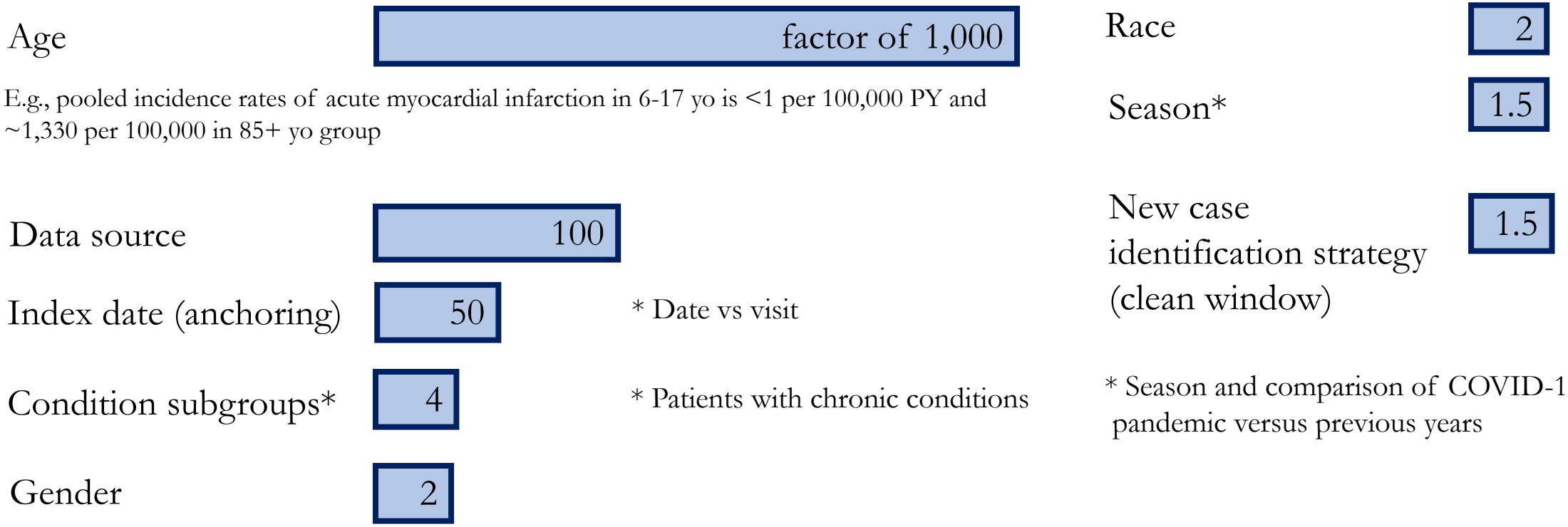
2017-2020
8 age groups
3 condition groups
2 sexes
2 races

4 index dates
2 TAR starts
5 TAR durations
4 years
4 seasons

10 research questions, compute **incidence rates** and incidence rate **ratios**, random effect model **meta-analysis**

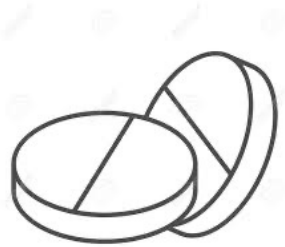
Main findings

Magnitude of influence



Selecting an **index date** in a cohort or case-crossover study

Exposed group/time

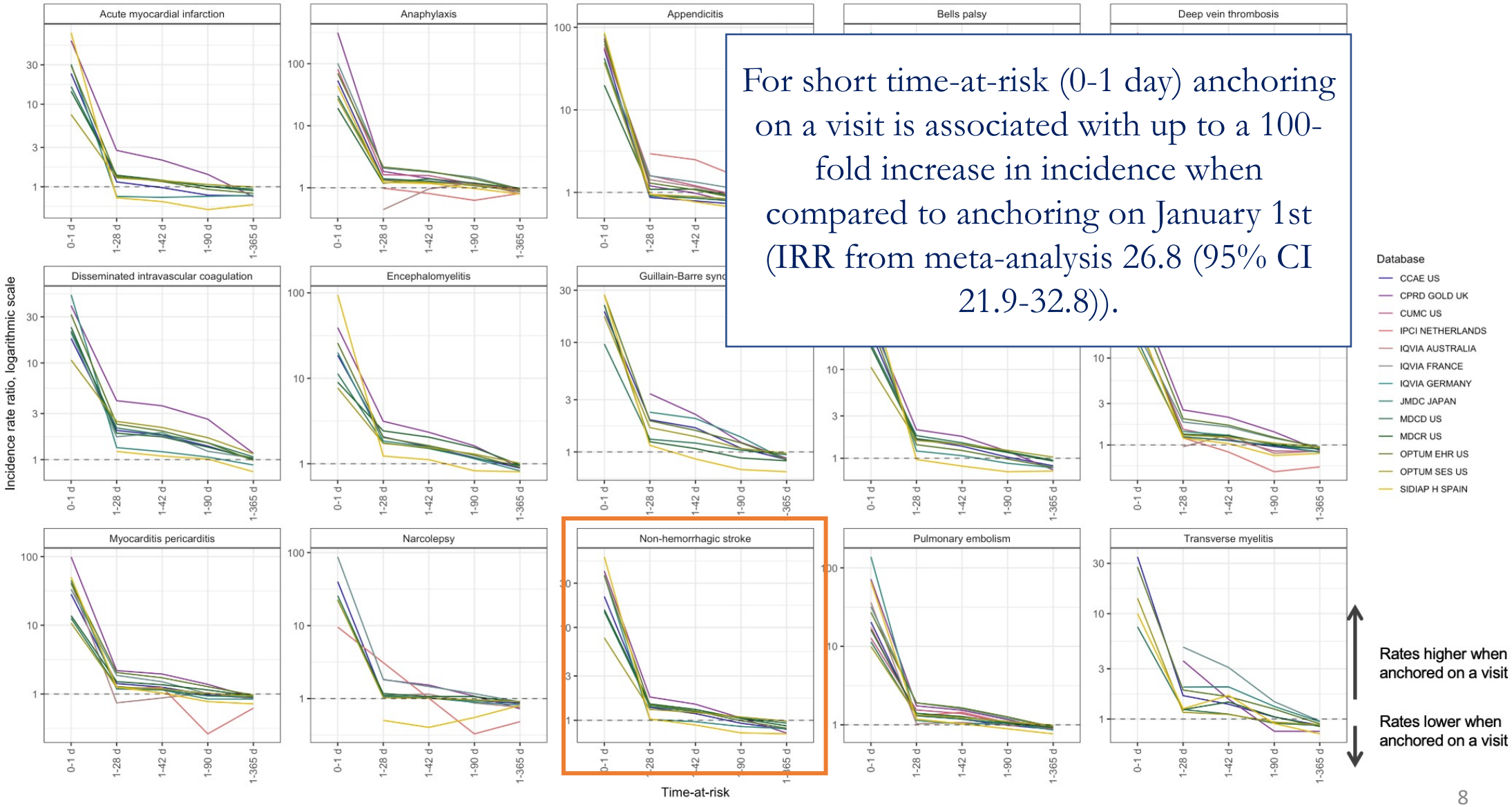


Drug exposure date
Vaccination date
Procedure date

Unexposed group/time

January 1st?
A random date?
A visit?
Something else?

Anchoring on a visit increases incidence rates for short and medium time-at-risk intervals for all conditions

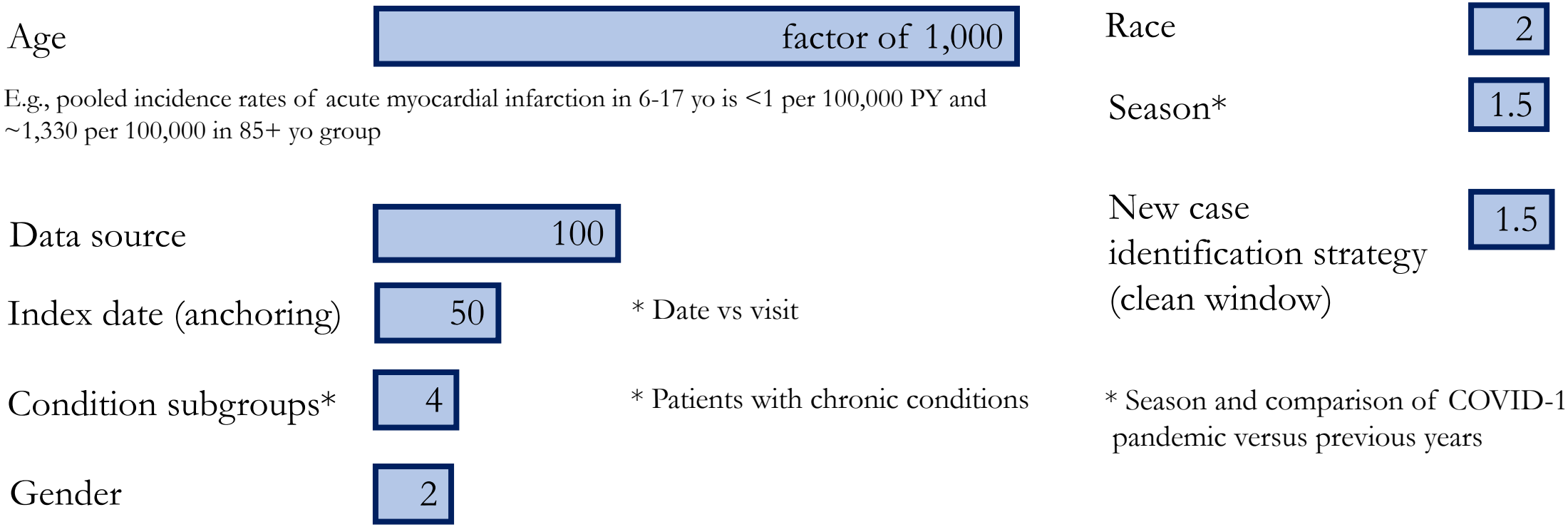


For short time-at-risk (0-1 day) anchoring on a visit is associated with up to a 100-fold increase in incidence when compared to anchoring on January 1st (IRR from meta-analysis 26.8 (95% CI 21.9-32.8)).

Comparison of anchoring on a random visit versus anchoring on January 1st in patients with a visit in the next year, incidence rate ratio.

Main findings

Magnitude of influence



Key findings

- Age has the largest impact on incidence with incidence rates varying up to a factor of 1,000 across age groups.
- Anchoring has a great impact on incidence rates (and patient characteristics). Anchoring time-at-risk interval on any type of healthcare encounter yielded higher incidence when compared to anchoring on a random date, especially for the short time-at-risk.
- Temporal and seasonal trends, gender, race and clean window choice have moderate influence

Key implications

- As population characteristics have high impact on baseline rates, population used for background rates calculation should represent the population for observed rate calculations

- As background rates don't have a definitive point-in-time index date for time-at-risk interval, we must select an index date or event (anchor) that serves as a counterfactual for exposure (vaccination) based on the background knowledge and empirical assessment