A meta-analysis of SARS-CoV-2 prevalence

using the Stan probabilistic programming language

Bob Carpenter

Center for Computational Mathematics, Flatiron Institute

May 2020

What is Prevalence?

- A condition's **prevalence** is the proportion of the population that has it
- We need to **estimate** prevalence of individuals
  - with SARS-Cov-2 **virus**,
  - with COVID-19 **disease**,
  - who have developed **antibodies** to SARS-Cov-2, and
  - who are **infectious**.

Why is Estimation Challenging?

- Conditions form a **scale**
  - how much virus? which symptoms? how infections?
- Measurements are **noisy**
  - **error**: inaccurate tests, varying accuracy across sites, human judgement, ...
  - **sampling**: extrapolate from sample to population
- Population **heterogeneity**
  - **population**: sex, age, medical conditions, ...
  - **behavior**: social distancing, protective measures, food, travel, ...
  - **testing**: availability, assignment, self selection, ...
  - **geo-political**: location, (local) government, climate, ...
  - **temporal**: everything changing over time