Test selection bias due to selective testing for exposures in pregnancy:

An example from SARS-CoV-2 and preeclampsia

Chelsea Messinger

MD-PhD Candidate, Epidemiology

HARVARD T.H. CHAN
SCHOOL OF PUBLIC HEALTH

Society for Pediatric and Perinatal Epidemiologic Research (SPER) Annual Meeting
Chicago, IL | June 2022

Outline

1. General overview of test selection bias
2. Theory behind test selection bias in causal studies of COVID-19 in pregnancy
3. Specific example using preeclampsia as outcome of interest

Test selection bias

A type of collider stratification bias that results from conditioning on (i.e., requiring) testing for an exposure (e.g., infection) when systematic testing is not performed

Typically, we think of bias with respect to causal studies, but test selection can be a problem in descriptive stats/studies as well

Test selection in in descriptive COVID epi

Massachusetts, case data from 5/1/22[1]

Let's estimate the point prevalence of COVID-19 in MA on May 1st

Cases (positive molecular test)	1,703
Molecular tests reported to DPH	15,782
Population of Massachusetts[2]	6,984,723[2]

Which number would you use in the denominator?

$$prevalence = \Pr(COVID = 1)$$
$$= \frac{N_{cases}}{N_{total\ pop}}$$
$$= \frac{1{,}703}{?}$$

[1]COVID-19 Response Reporting. Massachusetts Department of Public Health. https://www.mass.gov/info-details/covid-19-response-reporting. Accessed June 5, 2022.
[2]U.S. Census Bureau QuickFacts Massachusetts. https://www.census.gov/quickfacts/MA. Accessed June 5, 2022.

Test selection in in descriptive COVID epi

Massachusetts, case data from 5/1/22[1]

Let's estimate the point prevalence of COVID-19 in MA on May 1st

Cases (positive molecular test)	1,703
Molecular tests reported to DPH	15,782
Population of Massachusetts[2]	6,984,723[2]

Which number would you use in the denominator?

$$\frac{1{,}703}{15{,}782} = 10.8\%. \qquad \frac{1{,}703}{6{,}984{,}723} = 0.024\%$$

→Tested would be an overestimate of true prevalence due to higher pre-test odds of COVID-19 among the tested

→General population would be underestimate due to misclassification of asymptomatic infected as unexposed (numerator too small, denominator too big)

[1]COVID-19 Response Reporting. Massachusetts Department of Public Health. https://www.mass.gov/info-details/covid-19-response-reporting. Accessed June 5, 2022.
[2]U.S. Census Bureau QuickFacts Massachusetts. https://www.census.gov/quickfacts/MA. Accessed June 5, 2022.

Features of COVID-19 relevant for test selection bias

- Not routinely screened for (generally or in pregnancy)
- Large asymptomatic proportion
- Selective testing of specific populations of patients

Who gets tested?

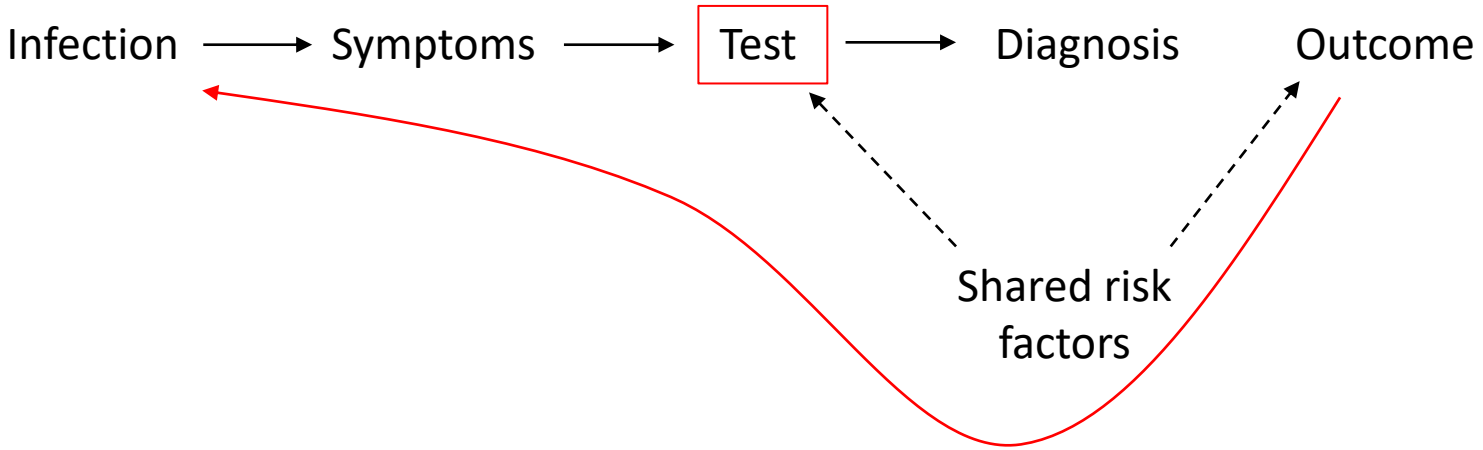
Patients with:

1. Symptoms of infection, or;
2. No symptoms, but other indications for testing (targeted screening), or;
3. No symptoms, in population-based screening

Causal questions and test selection bias

In the absence of systematic or universal testing for an infection, requiring a test result for inclusion as exposed/unexposed can introduce test selection bias in causal studies if the outcome of interest – or factors associated with the outcome – trigger testing.

→ Directed acyclic graph (DAG)



COVID & pregnancy: Framing causal questions

Two primary questions regarding COVID-19 in pregnancy:

1) **Among SARS-CoV-2 infected patients, what is the effect of pregnancy on severe COVID-19 outcomes?**
- Proxy outcomes include hospitalization, ICU admission, etc.
- Reference group: Infected nonpregnant patients

2) **Among pregnant women, what is the effect of SARS-CoV-2 infection on adverse maternal and neonatal outcomes?**
- Reference group: Uninfected pregnant women

In the absence of systematic/universal SARS-CoV-2 testing, both study types confront the challenge of selecting an appropriate reference group

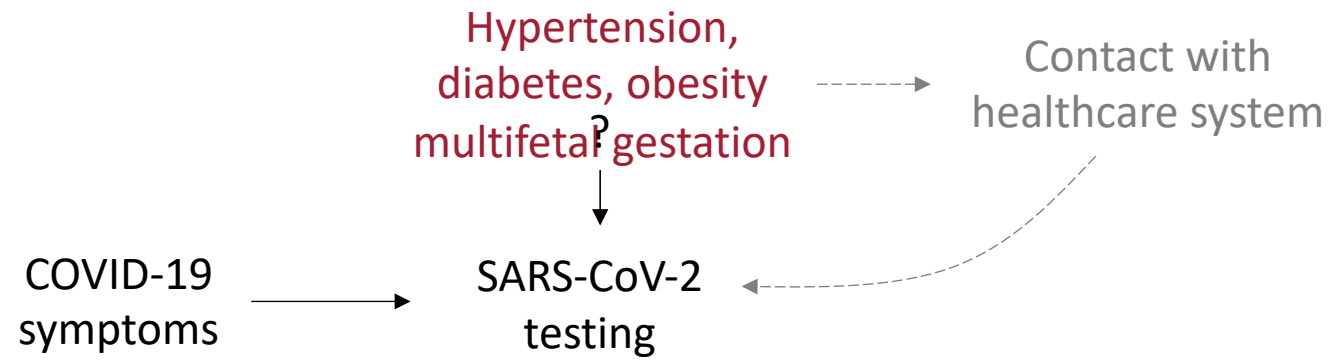
→ Test selection bias

Selective SARS-CoV-2 testing in pregnancy

Recall: Testing performed in people who have:

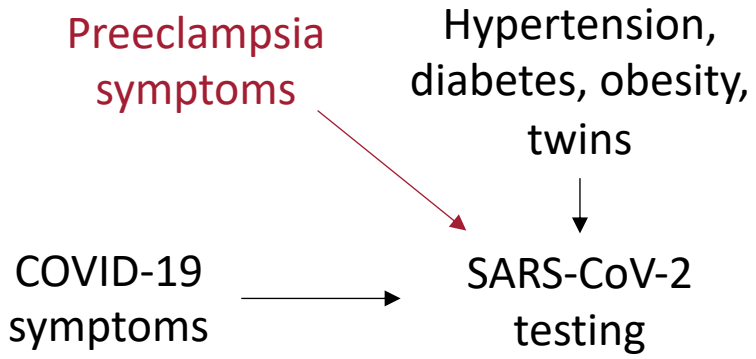
1. Symptoms of infection, or;
2. No symptoms, but other indications for testing (targeted screening), or;
3. ~~No symptoms, in population-based screening~~

- In asymptomatic patients, risk factors for severe COVID-19 may trigger more regular testing (targeted screening)
  - Some risk factors overlap with those for conditions like preeclampsia



Selective SARS-CoV-2 testing in pregnancy

- Additionally, there is overlap between symptoms of preeclampsia and symptoms of COVID-19
  - Nausea/vomiting
  - Headache
  - Abdominal pain
  - General malaise



Asymptomatic patients drive test selection bias

- **Key Point**: If a test is required for diagnosis, asymptomatic patients who are tested and diagnosed are likely to have had other reasons for testing that may be associated with the outcome
- In pregnancy, asymptomatic proportion is 54% - 68%[3]

  Concept check: What would you want to know about how this data was collected before comparing to the nonpregnant population?

  → Who was tested?

  → Systematic review of 68 studies of women in **universal screening programs**

[3]Allotey J, Stallings E, Bonet M, et al. Clinical manifestations, risk factors, and maternal and perinatal outcomes of coronavirus disease 2019 in pregnancy: living systematic review and meta-analysis. *BMJ*. 2020;370:m3320.

Study design: requiring a test result

Testing can be required for entry into:

- The exposed group only (patients with diagnosed COVID-19), or;
- Both the exposed and unexposed group (patients without diagnosed COVID-19)

Both designs involve conditioning on testing and can introduce bias; direction of the bias depends on who a test is required for

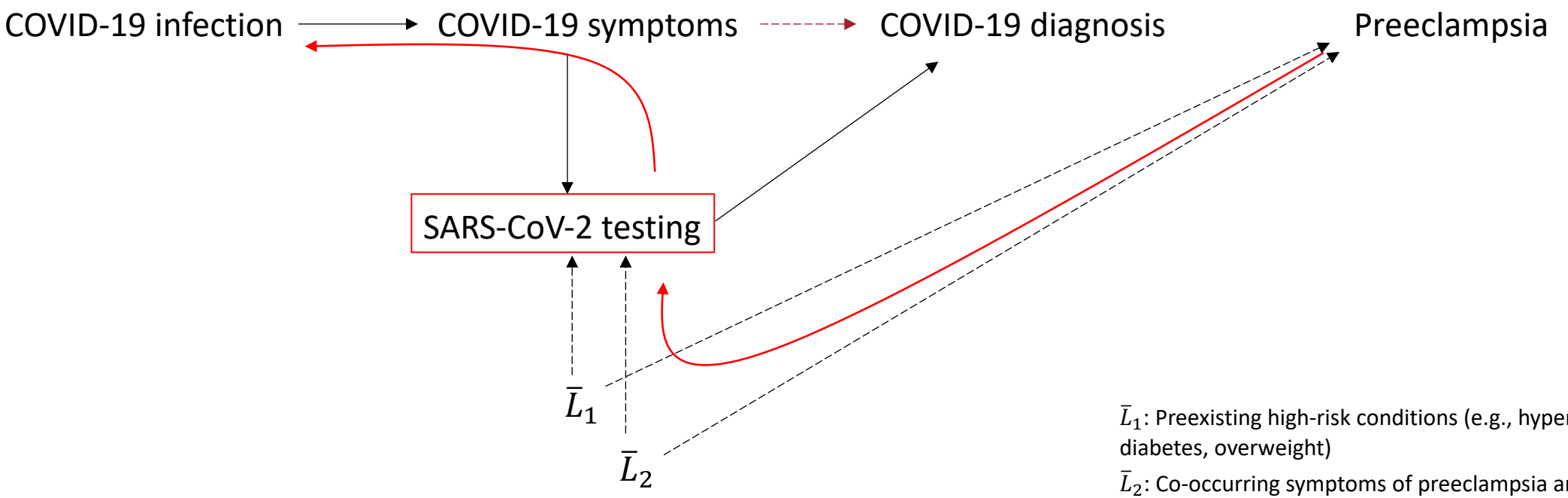
Specific example: COVID-19 and preeclampsia

Study Question:

Among pregnant women, what is the effect of SARS-CoV-2 infection on the incidence of preeclampsia?

- Study population: Pregnant women
- Exposed group: SARS-CoV-2 infected pregnant women
- Reference group: Uninfected pregnant women
- Outcome: Preeclampsia (all types)

Conditioning on testing → test selection bias



$\bar{L}_1$: Preexisting high-risk conditions (e.g., hypertension, diabetes, overweight)

$\bar{L}_2$: Co-occurring symptoms of preeclampsia and COVID-19

Hypothetical study: Require test for exposed only

Aim: Assess the effect of COVID-19 on preeclampsia

- Study population: Pregnant women age ≥18 years old
- Exposed group: Women with positive COVID-19 test (PCR or antigen) at any time in pregnancy
- Unexposed group: Women without positive COVID-19 test in pregnancy enrolled at same gestational age as exposed cases

The authors find that women with preeclampsia were mostly diagnosed with COVID-19 from 33 to 37 weeks' gestation, while women without preeclampsia had a steady rate of diagnosis over pregnancy. What is a possible non-causal explanation for these findings?

→ Preeclampsia often diagnosed around 33-37 weeks, and many deliver prematurely; increased contact with healthcare system → increased testing

Hypothetical study: Require test for exposed only

Aim: Assess the effect of COVID-19 on preeclampsia

- Study population: Pregnant women age ≥18 years old
- Exposed group: Women with positive COVID-19 test (PCR or antigen) at any time in pregnancy
- Unexposed group: Women without positive COVID-19 test in pregnancy enrolled at same gestational age as exposed cases

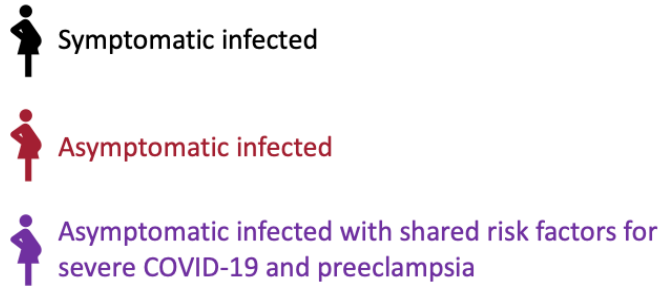
After adjusting for known confounders, the authors find a positive association between COVID-19 and preeclampsia. What could explain these findings? Assume perfect follow-up.

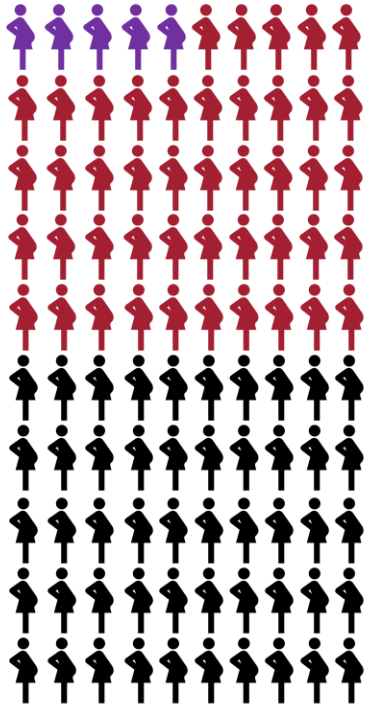
→ True association

→ Test selection bias due to selective testing of asymptomatic women with preeclampsia risk factors and/or symptoms

Direction of the bias

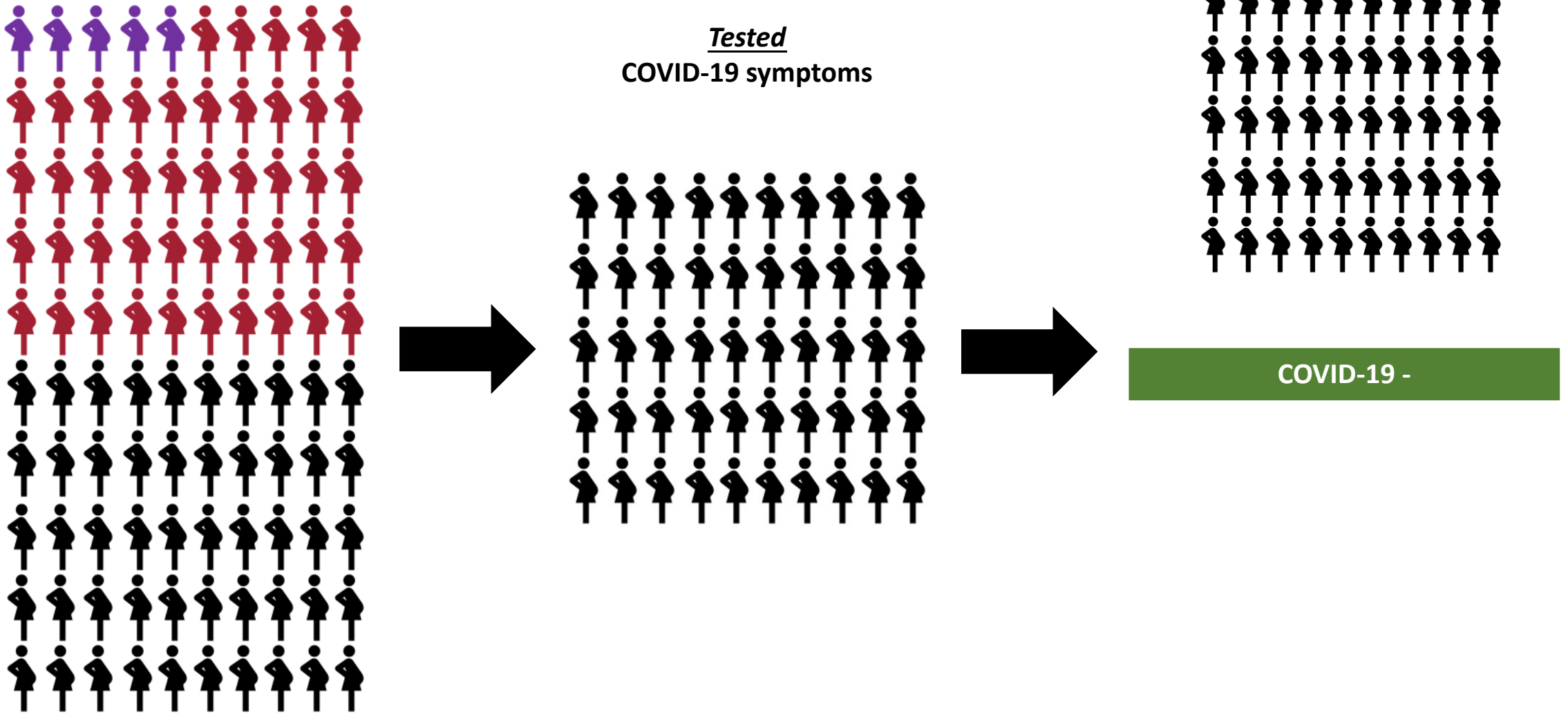
- Hypothetical study: require test for exposed only
- Example:
  - Take 100 **infected** pregnant women
  - Assume (conservative) 50% asymptomatic[3,4]
  - Assume 10% of asymptomatic have at least one shared risk factor for PE and severe COVID, e.g., hypertension or diabetes that triggers targeted screening





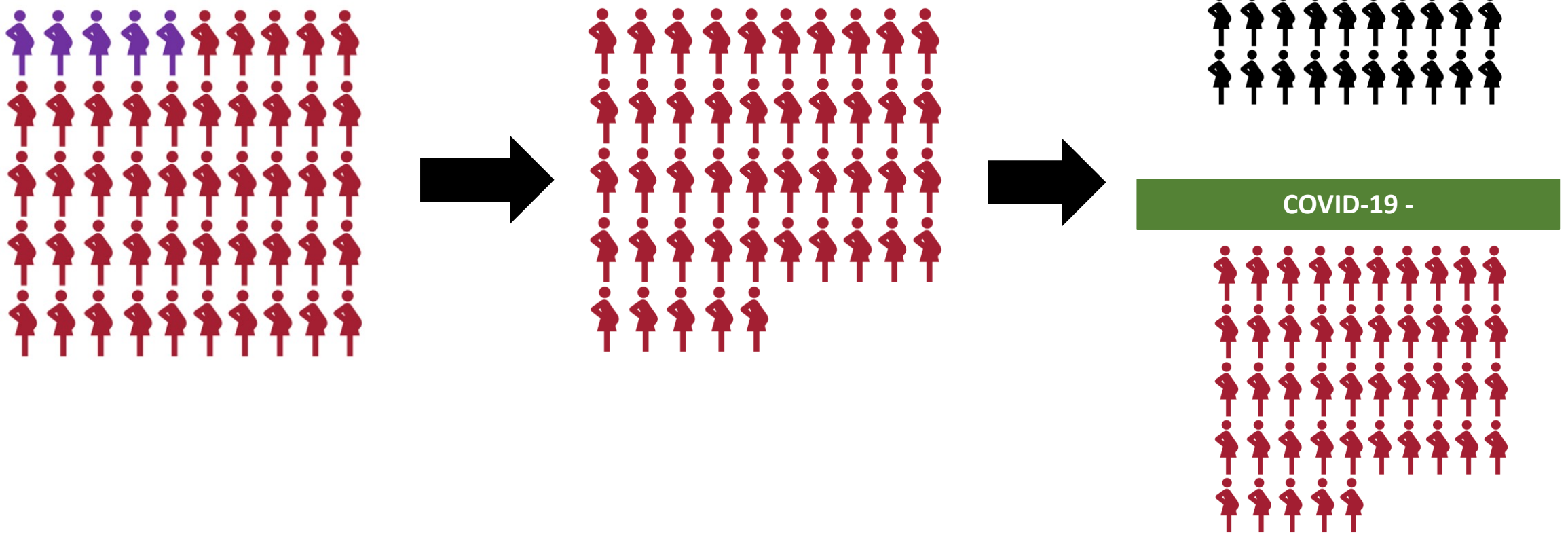
Symptomatic infected Asymptomatic infected Asymptomatic infected with common risk factors

Require testing for exposed only



Symptomatic infected | Asymptomatic infected | Asymptomatic infected with common risk factors

Require testing for exposed only



Symptomatic infected Asymptomatic infected Asymptomatic infected with common risk factors

Require testing for exposed only

Tested

Shared risk factors for PE and severe COVID-19 that trigger targeted screening

Require testing for exposed only

- Women with risk factors for and/or symptoms of preeclampsia are preferentially selected into the exposed group
- These women are more likely to develop preeclampsia

Which direction would you expect the bias to be in (upward/downward)?

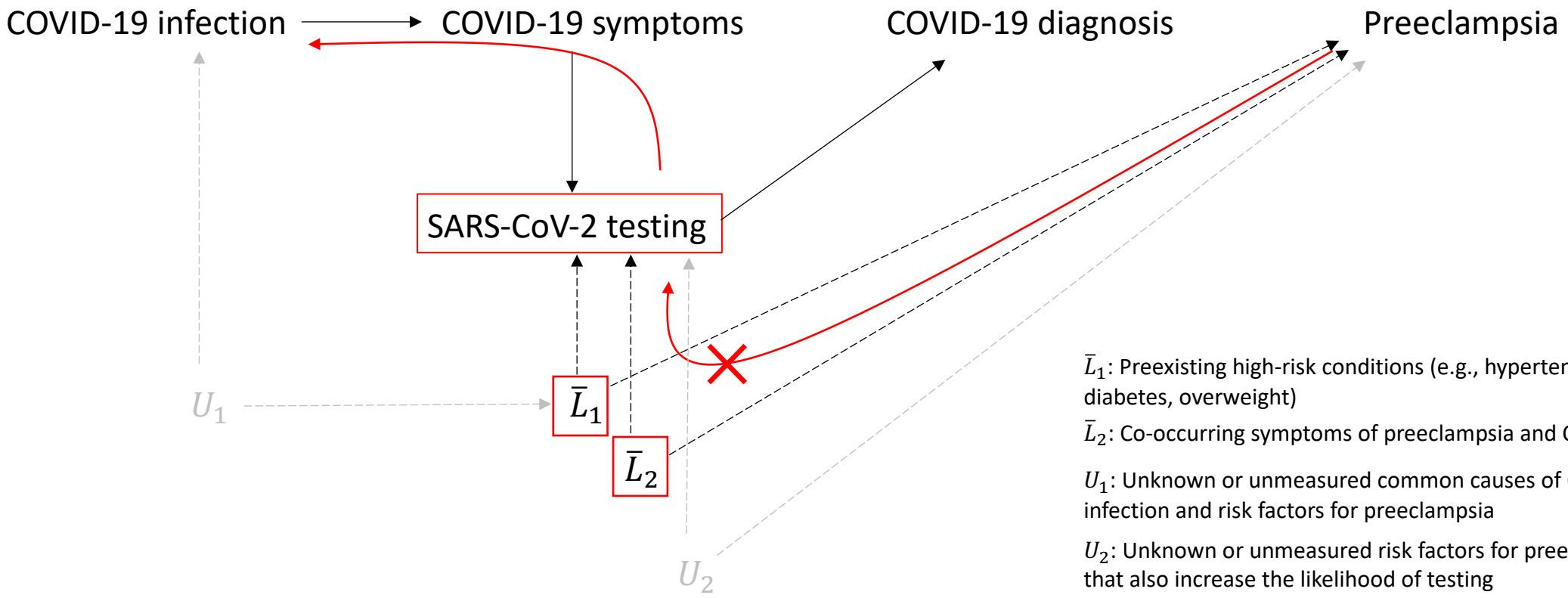
→ Upward

Flawed solution: require test for unexposed

- Sample reference group from among those with a documented negative test, *outside of population-based screening programs*
- Problem: Asymptomatic women with an available test result are likely to have had other reasons for being tested (targeted screening)
  - If these reasons are also risk risk factors for preeclampsia, test selection bias will still occur
  - Women will shared risk factors for preeclampsia and SARS-CoV-2 testing will now also be selected into the reference group
  - Bias can go either way depending on degree of selection into reference vs. exposure group

Potential solutions

1. Adjust for risk factors for outcome that also prompt testing



$\bar{L}_1$: Preexisting high-risk conditions (e.g., hypertension, diabetes, overweight)

$\bar{L}_2$: Co-occurring symptoms of preeclampsia and COVID-19

U_1: Unknown or unmeasured common causes of COVID-19 infection and risk factors for preeclampsia

U_2: Unknown or unmeasured risk factors for preeclampsia that also increase the likelihood of testing

Potential solutions

2. Enroll a random, population-based, asymptomatic, negative-test–confirmed reference group from the *same study population* and at the *same gestational age* as participants with confirmed COVID-19 infection
   - Can enroll prospectively by testing on day of enrollment
     - Limitations: labor-intensive, $$
   - Can enroll retrospectively from population of women in population-based screening program so long as it is the same study population as exposed
     - Limitations: Generalizability

Takeaways

1. Test selection bias is a type of collider stratification bias that results from conditioning on testing for an exposure when systematic testing is not performed
2. In causal questions about maternal/fetal outcomes, test selection bias can occur when risk factors for/symptoms of the outcome of interest (preeclampsia) also trigger testing for the exposure (COVID-19)
3. Test selection bias can occur in both prospective and retrospective studies
4. Researchers must be aware of how testing is related to exposure classification for both exposed and unexposed groups, and ideally account for sources of bias in the design stage of studies
5. Test selection bias can be extended to a wide variety of other infections in pregnancy (e.g., CMV/Zika and microcephaly[4]) as well as more generally to any clinical exposure that requires a diagnostic test and is selectively tested for

[4]Messinger C, Lipsitch M, Bateman B, et al. Association Between Congenital Cytomegalovirus and the Prevalence at Birth of Microcephaly in the United States. *JAMA Pediatr.* 2020;174(12):1159-1167.

Acknowledgements

Thanks to Sonia Hernández-Díaz, MD, DrPH for her helpful input on this content.

References

1. COVID-19 Response Reporting. Massachusetts Department of Public Health. https://www.mass.gov/info-details/covid-19-response-reporting. Accessed June 5, 2022.
2. U.S. Census Bureau QuickFacts Massachusetts. https://www.census.gov/quickfacts/MA. Accessed June 5, 2022.
3. Allotey J, Stallings E, Bonet M, et al. Clinical manifestations, risk factors, and maternal and perinatal outcomes of coronavirus disease 2019 in pregnancy: living systematic review and meta-analysis. *BMJ*. 2020;370:m3320.
4. Messinger C, Lipsitch M, Bateman B, et al. Association Between Congenital Cytomegalovirus and the Prevalence at Birth of Microcephaly in the United States. *JAMA Pediatr.* 2020;174(12):1159-1167.
5. Messinger C, Hernández-Díaz S, McElrath T. Further Observations on Pregnancy Complications and COVID-19 Infection. *JAMA Pediatr*. 2021;175(11):1184-1185.

Contact:

e: chelsea_messinger@hms.harvard.edu
@ChelsMessinger

Appendix

COVID & pregnancy: Framing causal questions

Two primary questions regarding COVID-19 in pregnancy:

1) **Among SARS-CoV-2 infected patients, what is the effect of pregnancy on severe COVID-19 outcomes?**
- Proxy outcomes include hospitalization, ICU admission, etc.
- Reference group: Infected nonpregnant patients

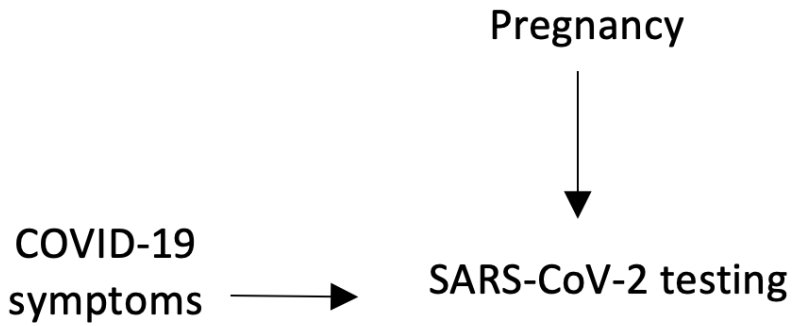
2) **Among pregnant women, what is the effect of SARS-CoV-2 infection on adverse maternal and neonatal outcomes?**
- Reference group: Uninfected pregnant women

In the absence of universal SARS-CoV-2 testing, both study types confront the challenge of selecting an appropriate reference group

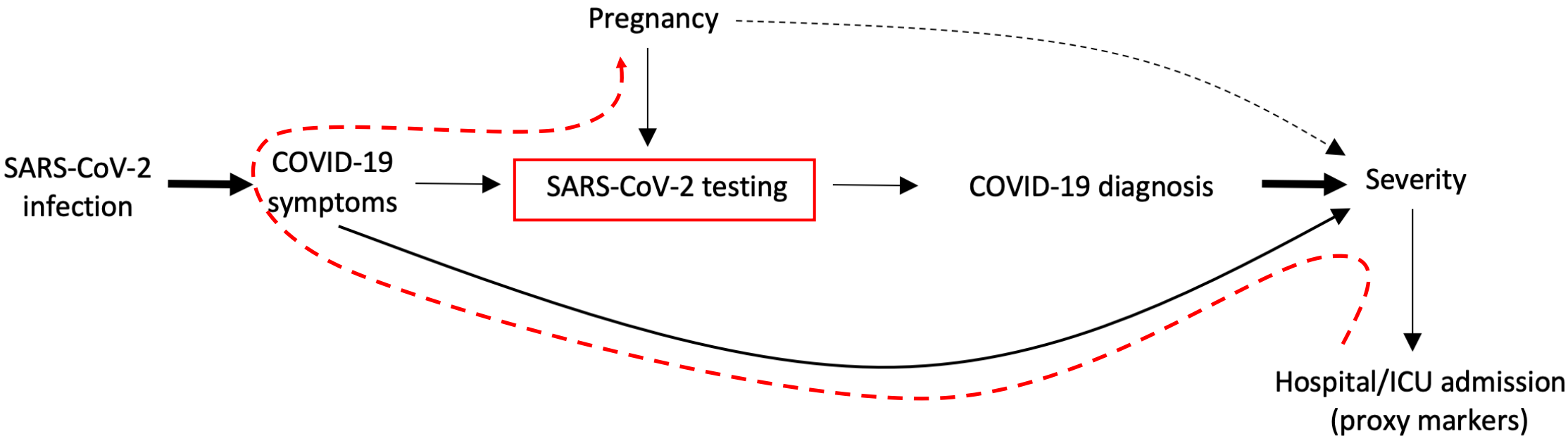
→ Test selection bias

Selective SARS-CoV-2 testing

- Testing for SARS-CoV-2 is most performed in people who have:
  1) Symptoms of infection, or:
  2) No symptoms, but other indications for testing (targeted screening), or:
  3) No symptoms, in population-based screening
- Pregnancy is a trigger for targeted screening, even in asymptomatic patients



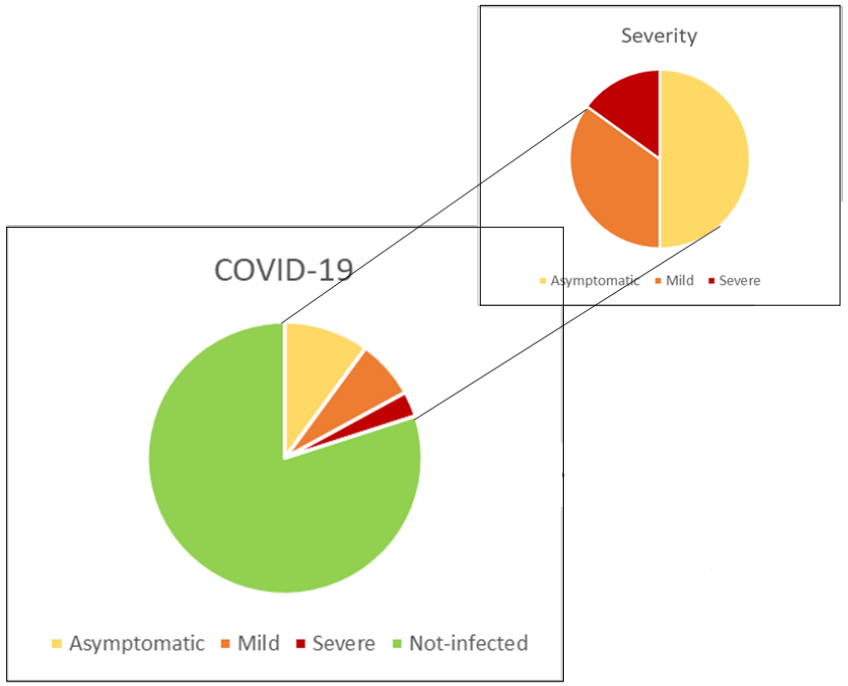
Test selection bias



Direction of the bias?

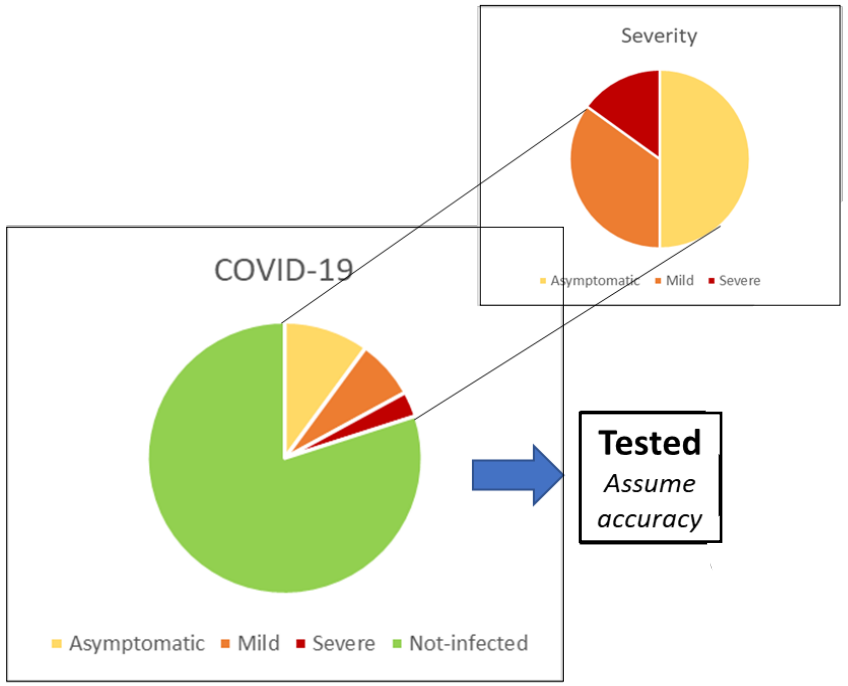
- Symptoms → severity
- Greater degree of undersampling of asymptomatic infected in nonpregnant population compared to pregnant population
- Among those with a test result available, would expect higher proportion of symptomatic infected in nonpregnant sample compared to pregnant sample → higher proportion of severe outcomes
- Results in downward bias of the effect estimate

$$\downarrow Risk\ Ratio = \frac{\Pr(severe = 1 | pregnancy = 1, COVID = 1, test = 1)}{\Pr(severe = 1 | pregnancy = 0, COVID = 1, test = 1)} \uparrow$$



True population distribution

Figure from paper under review; used with permission from S. Hernández-Díaz



True population distribution

Figure from paper under review; used with permission from S. Hernández-Díaz

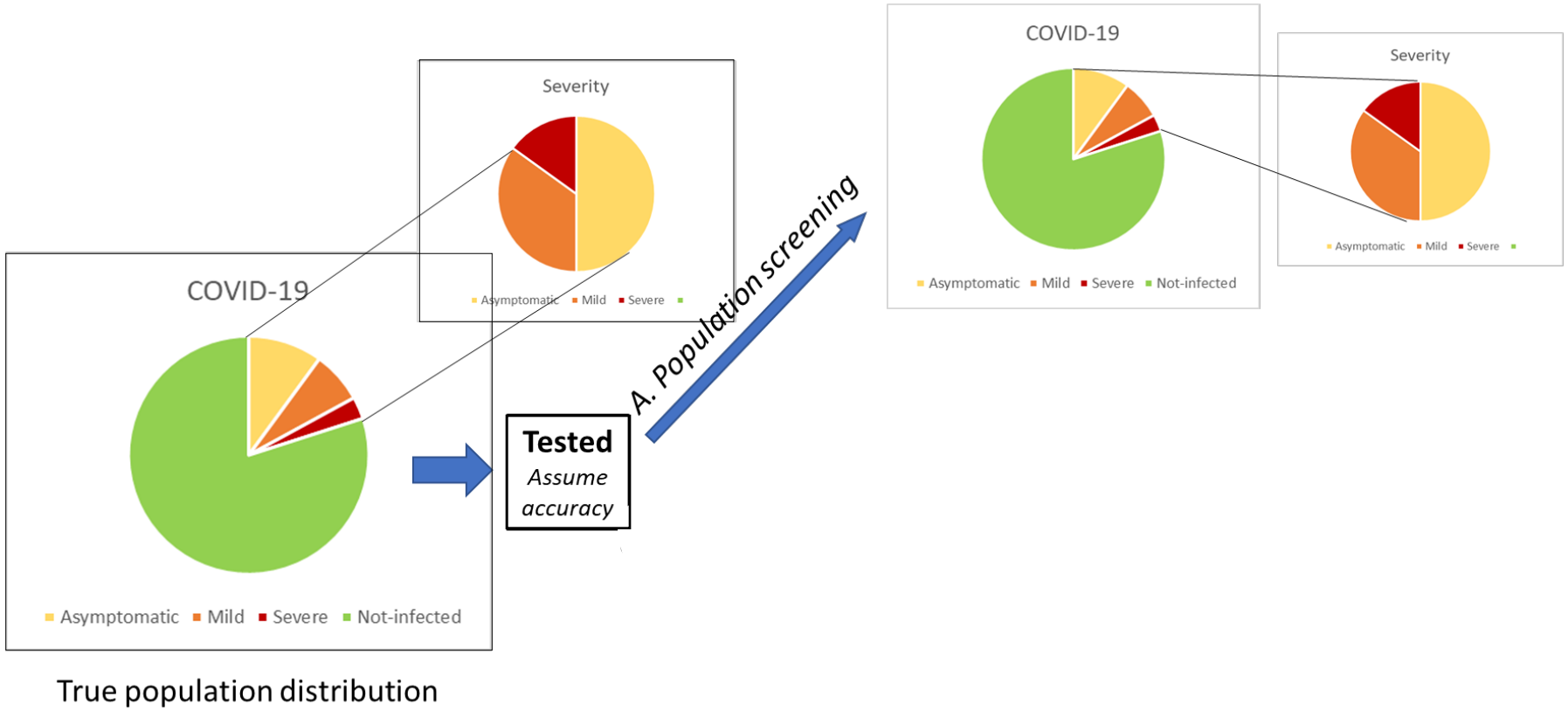
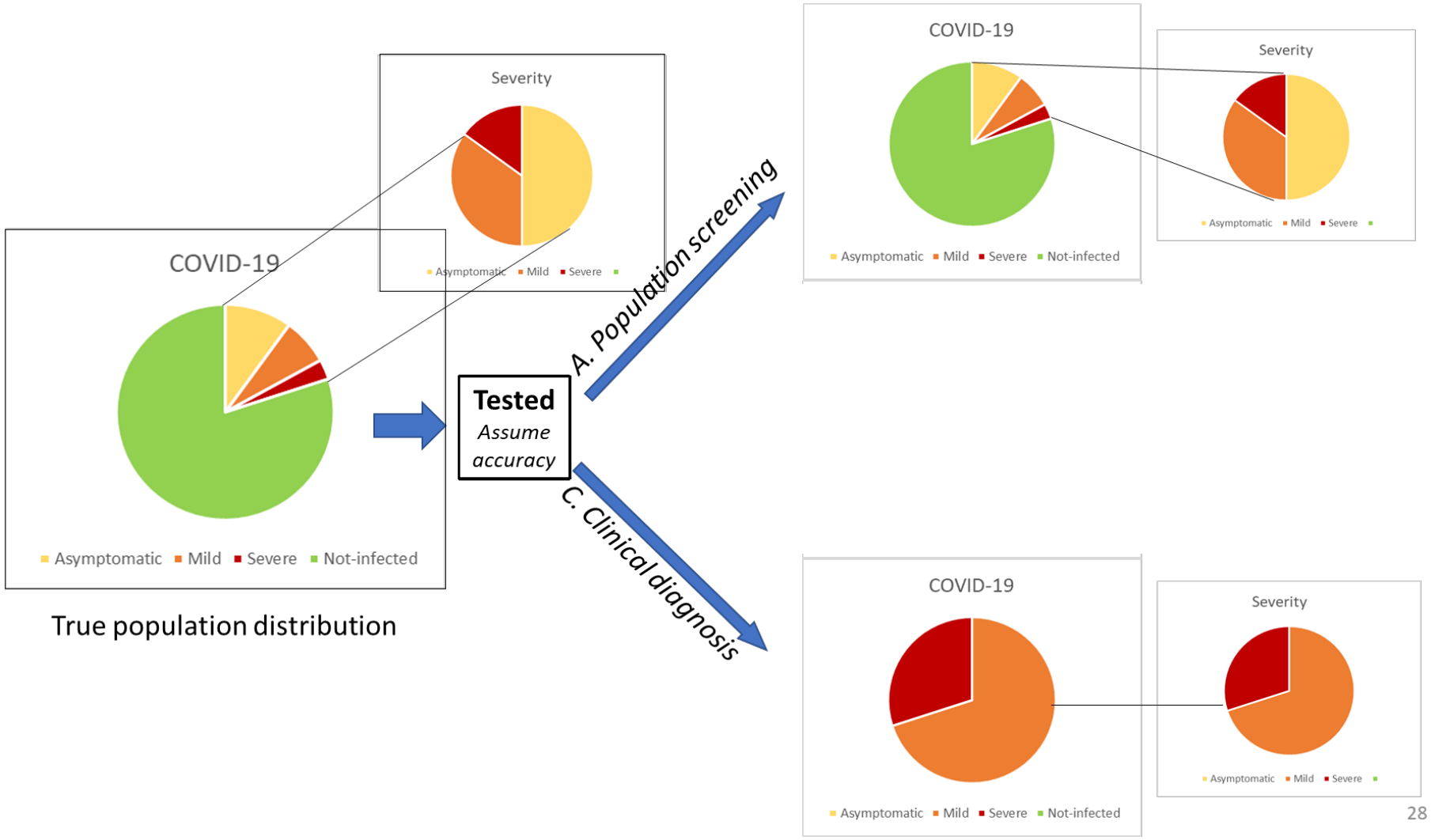


Figure from paper under review; used with permission from S. Hernández-Díaz



28

Figure from paper under review; used with permission from S. Hernández-Díaz

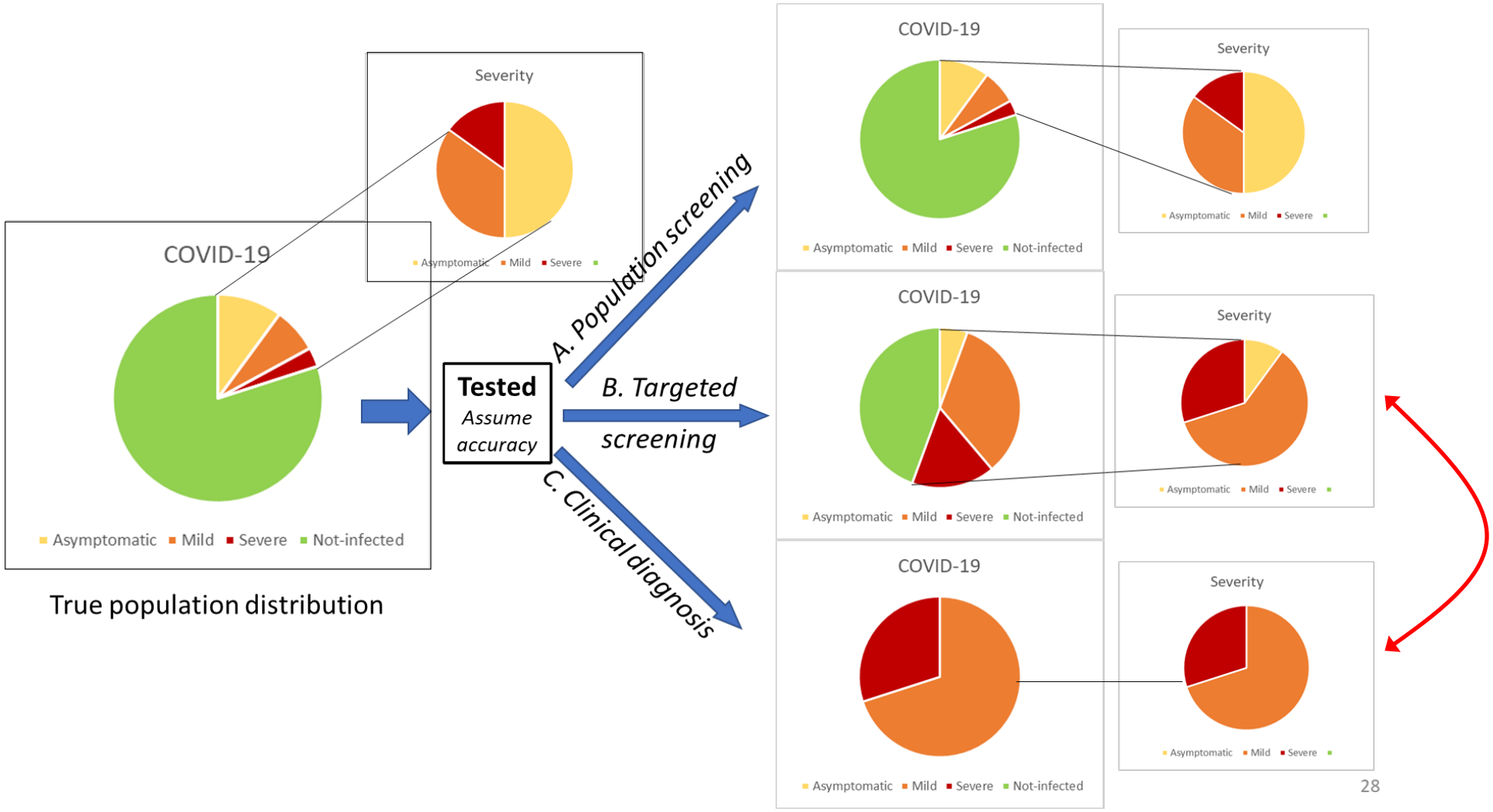
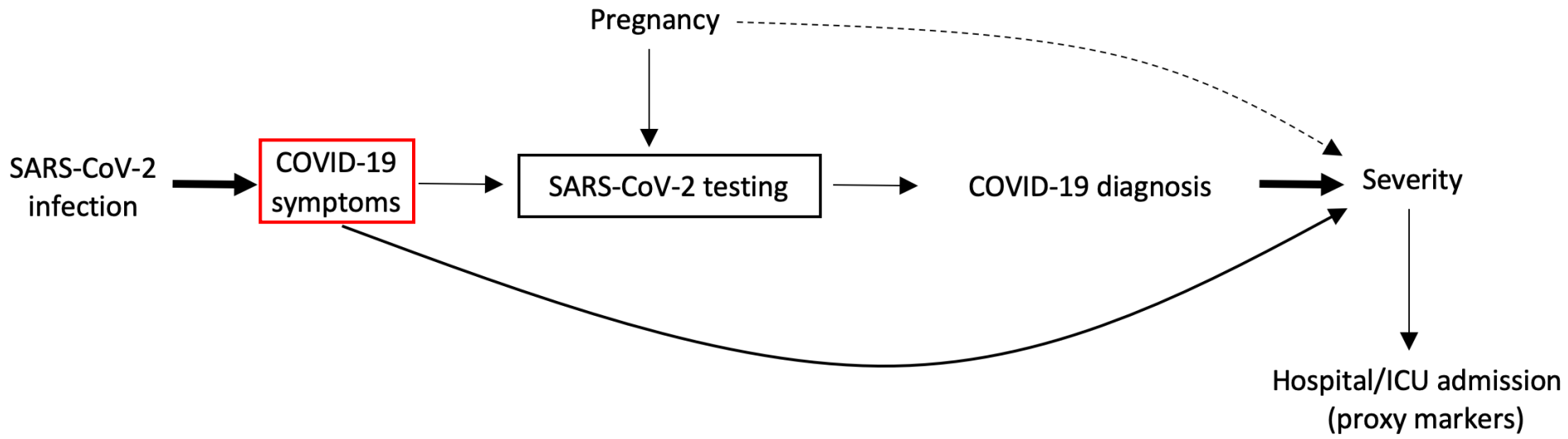


Figure from paper under review; used with permission from S. Hernández-Díaz

Potential solutions

1. Condition on COVID symptoms: compare outcomes within subset of tested population who are symptomatic
   - Yields conditional effect of pregnancy on severity (conditional on having symptoms), not total effect



Potential solutions

2. Enroll random sample of both pregnant and nonpregnant participants who receive regular testing
   - E.g. healthcare workers of childbearing age who are routinely tested for work
   - Limitation: generalizability