Selection Bias in COVID-19 Test Negative Design Studies

Eric J Tchetgen Tchetgen
University of Pennsylvania
with
Xu Shi, University of Michigan
and Kendrick Li, St. Jude Children’s Research Hospital
TND Study of vaccine effectiveness (Jackson and Nelson, Vaccine 2013)

• Ideal TND Study Sample: Patients who
  • Have Covid-like symptoms and as a result present at a healthcare facility to get tested.
  • Cases = test-positive, controls = test-negative
  • VE = 1 - risk ratio (risk in vaccinated / risk in unvaccinated) obtained via logistic regression
When does TND work (Jackson and Nelson, Vaccine 2013 Shi et al, AJE, 2023)

• The Directed Acyclic Graph (DAG) below illustrates the rationale justifying TND and encodes relationships between Vaccination (V), True infection Status (I), Observed infection status (I*), Symptoms (S), Testing (T), and Healthcare-Seeking behavior (H),

• Implicitly conditions on measured confounders: Age, Gender, Socioeconomic status.

• TND works to the extent that it reduces confounding by H by enrolling only individuals who test
When does TND work? (Jackson and Nelson, Vaccine 2013; Shi et al., AJE, 2023)

• More formally, TND makes three key assumptions:
  • Tested patients have the same healthcare-seeking behavior (H=1 if T=1);
  • V does not have a direct effect on testing (no V->T);
  • V does not have a direct effect on symptoms in test-positive sample (no V->S);

• Under these assumptions, vaccinated vs unvaccinated are comparable wrt H by design, yielding an unbiased VE estimate via logistic regression.
Challenges in TND studies (Shi et al., AJE, 2023; Sullivan et al., AJE, 2016)

• TND is susceptible to several potential sources of bias:
  • Confounding bias: there may be unmeasured common causes \( U \) of vaccination, COVID infection and testing, e.g. occupation as healthcare worker, educator, resident of care facility, previous infection, etc.

  • Assumption that tested patients have the same healthcare-seeking behavior is seldom realistic. HSB is likely on a spectrum and cannot be accounted for fully by conditioning on testing -> residual confounding by HSB
Challenges in TND studies (Shi et al, AJE, 2023; Sullivan et al, AJE, 2016)

• TND sources of bias continued:
  • More importantly, conditioning on testing may induce a particularly insidious form of selection bias known as collider stratification bias along the pathway V->T<-S->I. Collider bias can be made worse if as likely the case, vaccination has direct effect not only on testing but also on symptoms.

• Collider bias can render two factors that are independent in the population dependent in the TND sample.
Challenges in TND studies (Sullivan et al, AJE, 2016; Shi et al, AJE, 2023)

• TND sources of bias continued:
  • In fact, HSB are independent on S in the population, however both are positively associated with testing, then in TND sample HSB and S will be negatively correlated, that is, a person in the TND sample with low HSB is likely in the sample because they experience severe S. This, in turn, creates spurious association between V and I!
  • Furthermore, recent challenges that TND must face include widespread home testing and repeat testing which likely distort selection into TND studies.
References


