Challenges and successes in estimating vaccine effectiveness from observational data

Jonathan Sterne
University of Bristol, UK
Declarations

I do not have any financial interests with any firms/entities that are related to the meeting topic.

I am co-lead of the Longitudinal Health and Wellbeing COVID-19 UK National Core Study (see https://www.ucl.ac.uk/covid-19-longitudinal-health-wellbeing/national-core-study-0).

This study aims to understand the health, social and economic impacts of the COVID-19 pandemic by uniting established population cohorts and national anonymised electronic health records to inform policy.
Randomized trials of COVID-19 vaccines

Among 10 cases of severe Covid-19 with onset after the first dose, 9 occurred in placebo recipients and 1 in a BNT162b2 recipient.

From 21 days after first dose, there were ten cases hospitalised for COVID-19, all in the control arm; two classified as severe COVID-19, including one death.
Most studies of effectiveness of COVID-19 vaccines are observational studies using routine data assembled during the rollout

• Randomized trials provide the best estimates of effectiveness in the real world, but…
  • A host of urgent questions could not be addressed in randomized trials
  • Far reaching policy decisions have been made using observational studies
  • Such studies aim to make causal inferences about the effects, and comparative effects, of vaccines and vaccination strategies
• To make causal inferences from observational data, think about the randomized trial whose result you would like to estimate
Practice of Epidemiology

Using Big Data to Emulate a Target Trial When a Randomized Trial Is Not Available

Miguel A. Hernán* and James M. Robins

*Correspondence to Dr. Miguel A. Hernán, Department of Epidemiology, 677 Huntington Avenue, Boston, MA 02115 (e-mail: miguel_hernan@hsp.harvard.edu).

Initially submitted December 9, 2014; accepted for publication September 8, 2015.

Ideally, questions about comparative effectiveness or safety would be answered using an appropriately designed and conducted randomized experiment. When we cannot conduct a randomized experiment, we analyze observational data. Causal inference from large observational databases (big data) can be viewed as an attempt to emulate a randomized experiment—the target experiment or target trial—that would answer the question of interest. When the goal is to guide decisions among several strategies, causal analyses of observational data need to be evaluated with respect to how well they emulate a particular target trial. We outline a framework for comparative effectiveness research using big data that makes the target trial explicit. This framework channels counterfactual theory for comparing the effects of sustained treatment strategies, organizes analytic approaches, provides a structured process for the criticism of observational studies, and helps avoid common methodologic pitfalls.

big data; causal inference; comparative effectiveness research; target trial
Features of randomized trials of vaccine effectiveness

• Define eligible participants
• Define intervention (vaccination) and comparator (no vaccination / vaccination against a different infection)
• Random assignment to vaccine or comparator
• Follow up for vaccine and comparator group starts on the day of assignment
  • The calendar date of assignment is comparable for the two groups
• Follow up continues for the same time, regardless of intervention group
Confounding occurs when there is a common cause (C) of both vaccination (V) and the outcome event (Y)
“Sequential” specification of a target trial
BNT162b2 mRNA Covid-19 Vaccine in a Nationwide Mass Vaccination Setting

Noa Dagan, M.D., Noam Barda, M.D., Eldad Kepten, Ph.D., Oren Miron, M.A., Shay Perchik, M.A., Mark A. Katz, M.D., Miguel A. Hernán, M.D., Marc Lipsitch, D.Phil., Ben Reis, Ph.D., and Ran D. Balicer, M.D.

ABSTRACT

BACKGROUND
As mass vaccination campaigns against coronavirus disease 2019 (Covid-19) commence worldwide, vaccine effectiveness needs to be assessed for a range of outcomes across diverse populations in a noncontrolled setting. In this study, data from Israel’s largest health care organization were used to evaluate the effectiveness of the BNT162b2 mRNA vaccine.

METHODS
All persons who were newly vaccinated during the period from December 20, 2020, to February 1, 2021, were matched to unvaccinated controls in a 1:1 ratio according to demographic and clinical characteristics. Study outcomes included documented cases of Covid-19, confirmed by a positive polymerase chain reaction test, during the 2-week period after vaccination.
“Sequential” specification of a target trial

On each day, a vaccinated individual is closely matched to an unvaccinated (control) individual

- Covariates (matching factors) are measured up to the day of vaccination
- Follow up continues until the control individual is vaccinated, at which time follow up for both individuals is censored.
  - Can conduct sensitivity analyses extending follow up (eg for a week) subsequent to vaccination of the control individual

- Control individuals can subsequently be included as a vaccinated individual in a new pair
- Can directly compare cumulative incidence in the two groups, or can adjust for additional covariates beyond those used for matching.
- We waste a lot of data (matching failures, censoring follow up of vaccinated individuals) but we compare similar individuals over the same time periods
In a public health emergency, we need to balance the need for rapid estimates of VE with the need to address potential biases.
Covid-19 Vaccine Effectiveness against the Omicron (B.1.1.529) Variant


BACKGROUND
A rapid increase in coronavirus disease 2019 (Covid-19) cases due to the Omicron (B.1.1.529) variant of severe acute respiratory syndrome coronavirus 2 has raised concerns about the effectiveness of current vaccines.

METHODS
We used a test-negative case–control design to estimate vaccine effectiveness among symptomatic disease caused by the omicron and delta (B.1.617.2) variants. Vaccine effectiveness was calculated after primary immunization with BNT162b2 (Pfizer–BioNTech), ChAdOx1 nCoV-19 (AstraZeneca), or mRNA-1273 (Moderna) vaccine and after a booster dose of BNT162b2, ChAdOx1 nCoV-1A, and mRNA-1273.
Test negative designs

Compare individuals with symptoms who test positive (cases) with those who test negative (controls)
Conditioning on common effects induces associations

Selection bias (‘collider bias’)
Academic ability and sporting ability

• In the general population, academic ability and sporting ability are unrelated.
• However, expensive private schools in England recruit on the basis of both academic and sporting ability:
• Among children at expensive private schools, the two characteristics are inversely associated.
Using selected populations for VE research

- Suppose that perceived risk of infection and attitudes to vaccination each influence use of a health monitoring app
- Associations between causes of vaccination and risk factors for the target infection may be distorted among app users
- To correct for this, we would need to measure and adjust for influences on use of the app
Case-control studies

- Useful to think of a cohort study in relation to its target randomized trial.
- Useful to think of a case-control study in relation to its target cohort.
  - For example, in the target cohort, potential confounders are measured at the start of follow up, not when the outcome occurs.
- Do we need to sample controls?
  - Modern computers can handle analyses based on many millions of individuals.
  - If the whole population is defined, we can sample based on any characteristic, and use inverse sampling probability weighting to recreate the result from the whole cohort.
  - For example, if vaccination is rare we could include all outcome events and all vaccinated individuals, together with a random sample of other individuals.
- Main justification for case-control studies may be when the population is not well-defined, and we sample controls on the basis of geographical or social proximity to cases.
Include “negative control” outcomes

*and not clinically vulnerable; **all dates are in 2021
Thank you for your attention