Beyond the intention-to treat effect:
Per-protocol effects in randomized trials

Miguel Hernán

DEPARTMENTS OF EPIDEMIOLOGY AND BIOSTATISTICS

HARVARD T.H. CHAN
SCHOOL OF PUBLIC HEALTH

Intention-to-treat **analysis** (estimator) estimates intention-to-treat **effect** (estimand)

- Intention-to-treat effect
  - The effect of being assigned to a treatment strategy, regardless of treatment received, in a particular setting

- Intention-to-treat effects are agnostic about post-randomization decisions
  - Changes in studied treatment: discontinuation, switching…
  - Use of concomitant therapies prohibited by the study protocol
  - etc.

Demystifying intention-to-treat effects: Not necessarily preserve the null

- Consider a non-blinded trial
- The ITT effect may not be null even if treatment has a null effect on the outcome
  - Patients and doctors may just alter their behavior in ways that affect their outcome

- *Most pragmatic trials are not blinded*

Demystifying intention-to-treat effects: Not necessarily biased towards the null

- When the treatment effect is not monotonic
  - not in the same direction for all individuals
- Trial of active treatment vs placebo
  - 30% of the individuals assigned to treatment did not adhere to treatment
  - direction of the effect in adherers opposite to that in non-adherers

- An ITT analysis may misleadingly indicate a beneficial effect of the less efficacious treatment

Demystifying intention-to-treat effects: Not necessarily biased towards the null

- Even if the treatment effect is monotonic
- Trial of 2 active treatments with differential adherence
  - due to a mild, easily palliated side effect
- An ITT analysis may misleadingly indicate a beneficial effect of the less efficacious treatment

- *Many pragmatic trials are head-to-head trials*

Demystifying intention-to-treat effects: Bias towards the null is often undesirable

- Safety trials
- Non-inferiority trials

- In these trials, a "conservative" ITT analysis is statistical malpractice
  - A trial designed to quantify harm and whose protocol foresees only an ITT analysis could be referred to as a 'randomized cynical trial'
- *Many pragmatic trials are for safety, non-inferiority*

Demystifying intention-to-treat effects: Not necessarily a measure of effectiveness

- Degree of adherence outside the trial may change drastically after doctors and patients learn of the trial’s findings

- Actual effectiveness in the community may differ from ITT effect estimate from trial

Demystifying intention-to-treat effects: Not of primary interest for doctors and patients

- For example, a couple trying to decide whether to use a contraceptive method would want to know
  - the effectiveness of the method when used as indicated
  - not the estimated effectiveness in a population in which, say, 40% of couples failed to use the method properly
  - That is, not the ITT effect

- *Pragmatic trials are designed to guide clinical decisions by patients and doctors*

Need a complement to the ITT effect:

- An effect measure (an estimand)
  - not affected by the degree of adherence
  - usable in safety, noninferiority trials
  - clinically relevant, patient-centered

- Per-protocol effect:
  - the effect of implementing the treatment strategies as described in the protocol

A big difference between ITT effect and per-protocol effect

- We have a universally accepted way of estimating ITT effects
  - ITT analysis
  - Almost uncontroversial
- We don't have a universally accepted way of estimating per-protocol effects
  - There are many types of per-protocol analysis
  - Including the commonly used, unadjusted, naïve per-protocol analysis

Intention-to-treat effect
Analysis plan

- Simple
- Compare outcome distribution between group assigned to different strategies
  - Regardless of whether individuals actually followed the strategies
- Often overlooked problem:
  - ITT analysis cannot be conducted if there are losses to follow-up
  - Potential selection bias due to informative censoring

Intention-to-treat effect
Analysis plan

- Estimating ITT effect requires adjustment for selection bias due to loss to follow-up
  - Adjustment for baseline and post-baseline covariates
  - Little et al, NEJM 2012
- In fact, intention-to-treat effect is more precisely defined as
  - the effect of being assigned to a strategy, regardless of strategy received, *while staying under follow-up throughout the study*

Per-protocol effect
Analysis plan

- Not so simple
- Treatment decisions after baseline are not randomized
  - Potential post-randomization confounding and selection bias
- Example
  - In a statins trial, statin use after baseline may depend on post-baseline cholesterol levels; dropout may depend on side effects and prognosis

Per-protocol effect
Analysis plan

- Estimating the per-protocol effect requires adjustment for confounding
  - Adjustment for baseline and post-baseline covariates

- In addition to adjustment for selection bias
  - same as for ITT effects

Effects (estimands) vs. *analyses* (estimators)
The elephant in the room

- Typical ITT and per-protocol **analyses**
  - adjust for neither pre- nor post-randomization variables
  - Potentially biased estimates of ITT and per protocol **effects**

- This is a problem for all randomized trials
  - because treatment choices and participation decisions after baseline are not randomly assigned
- But especially for pragmatic trials
  - with lots of room for non-adherence and loss to follow-up

A pragmatic randomized trial is a follow-up study with baseline randomization

- Analysis methods to adjust for post-baseline confounding and selection bias are the same methods used for observational follow-up studies
- Adjustment for post-randomization (time-varying) variables require special techniques
  - Inverse probability (IP) weighting, g-formula, etc
    - Developed by Robins et al since 1986
  - Instrumental variable estimation

Case study

Hormone therapy and breast cancer

Question

- What is the effect of postmenopausal hormone therapy on risk of breast cancer in postmenopausal women?

Data

- A Women's Health Initiative randomized trial
  - ~16,000 postmenopausal U.S. women
  - Toh et al. *Epidemiology* 2010; 21:528-539

Effect of hormone therapy, what effect?

- Effect of assignment to hormone therapy under the study's conditions?
  - Intention-to-treat effect
- Effect of hormone therapy use as instructed by the study's protocol?
  - Per-protocol effect
- BOTH
  - They answer different questions

Methodological challenges for per protocol effect

- Time-varying treatment
  - Women may not adhere to their assigned treatment (hormone therapy or placebo)
- Time-varying confounders
  - Use of hormone therapy depends on age, BMI, symptoms…
  - may be affected by prior treatment
- Also better to estimate absolute risks
  - Appropriately adjusted survival curves
  - Not only hazard ratios

Methodological approach to estimate per protocol effect

- Estimate IP weights to adjust for time-varying confounding
  - Need data on post-randomization variables
- Estimate IP weighted hazards model to estimate
  - Hazard ratios
  - Survival (or cumulative incidence)
- Compare survival curves for continuous treatment vs. no treatment
  - Standardize curves to baseline variables

Hazard ratio of breast cancer
Hormone therapy vs. placebo

- Intention to treat effect estimate
  - 1.25 (1.01, 1.54)
- Per protocol effect estimate
  - 1.68 (1.24 to 2.28)

- Suppose you are a woman considering initiation of hormone therapy and who plans to take it as instructed by your doctor
  - Which hazard ratio do you want?

Validity of per-protocol effect estimates

- Relies on adjustment for post-randomization confounding and selection bias
- via the same analytic methods
  - and under the same untestable assumptions
- that we usually reserve for observational studies

Review: Classification of treatment strategies according to their time course

- **Point** interventions
  - Intervention occurs at a single time
  - Examples: one-dose vaccination, short-lived traumatic event, surgery...
- **Sustained** strategies
  - Interventions occur at several times
  - Examples: medical treatments, lifestyle, environmental exposures...

Choice of statistical adjustment method depends on type of strategies

- Comparison of strategies involving point interventions only
  - All methods work
  - if all confounders are measured or the instrumental variable conditions hold
- Comparison of sustained strategies
  - Generally only g-methods work
  - Developed by Robins and collaborators since 1986

Per-protocol effect is generally a contrast of sustained (dynamic) treatment strategies

- Not a comparison of continuous treatment A vs. continuous treatment B
- But a comparison of strategies of the sort
  - "start taking A, continue taking A until toxicity arises, then switch to B"

- Implications for
  - definition of per-protocol effect
  - definition of adherence
  - data collection requirements: need post-randomization data on treatment adherence and (time-varying) confounders

Conclusions (I)

- There are good reasons for ITT analyses to remain the primary analyses of many randomized trials
- Also good reasons for appropriately adjusted per-protocol analyses as an integral component of randomized trial analysis
  - especially relevant to patients and clinicians
  - can also be used by modelers and healthcare planners to estimate an upper bound of the impact of changes in recommendations

Conclusions (II)

- The validity of per-protocol effects requires
  - Explicit definition of per-protocol effect and adherence
  - A priori specification of the statistical plan for the per-protocol analysis
  - High-quality data on adherence and prognostic factors
  - Appropriate adjustment methods

- These requirements necessitate changes in the way we design and conduct trials

Thank you
(more on Twitter @_MiguelHernan)

- This work is supported by PCORI ME-1503-28119
- Additional readings
  - Hernán MA, Hernández-Díaz S. Beyond the intention to treat in comparative effectiveness research. *Clinical Trials* 2012; 9(1):48-55.
  - Hernán MA, Hernández-Díaz S, Robins JM. Randomized trials analyzed like observational studies. *Annals of Internal Medicine* 2013; 159(8): 560-562
  - Toh S, Hernán MA. Causal inference from longitudinal studies with baseline randomization. *International Journal of Biostatistics* 2008; 4(1): Article 22