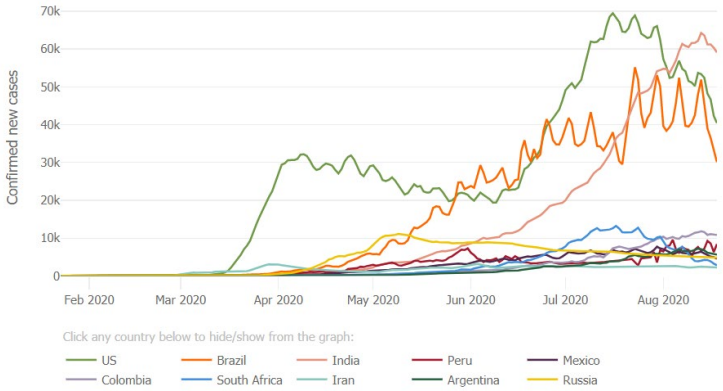
Revisiting ICH E9 (R1) During the COVID-19 Pandemic

Ilya Lipkovich, Yongming Qu
Eli Lilly and Company, Indianapolis, IN 46285, USA

Biopharmaceutical Applied Statistics Symposium (BASS) XXVII
Sept 9, 2020

The COVID-19 pandemic and the impacts on clinical trials

Daily confirmed new cases (5-day moving average) for 10 most affected countries

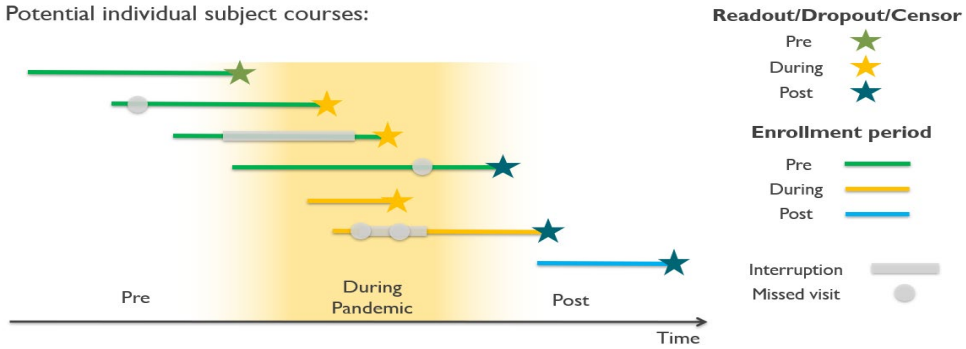


https://coronavirus.jhu.edu/data/new-cases
Meyer, et al. (2020)

COVID-19-related factors that are likely to impact ongoing trials

- Quarantines, travel limitations, site closures, or reduced availability of site staff
- Interruptions to supply chain of experimental drug and/or other medications
- Temporarily stopping drug due to safety concerns
- Alternative administration of drug
- Alternative collection of specimens
- Alternative data collection
- COVID-19 infection/treatment

How COVID-19 impacts ongoing clinical trials

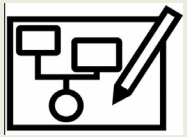


Key Point: Introduces challenges by impact to analyses and PRSS

A systematic approach to addressing challenges

ASSESS IMPACT

- Impact on data quality
- Impact on recruitment and retention
- Impact on treatment effects and study power
- Blinded/Unblinded Review



DEFINE RISK

- Lack of interpretability
- Confounding or inconclusive results
- Loss of power

MITIGATION

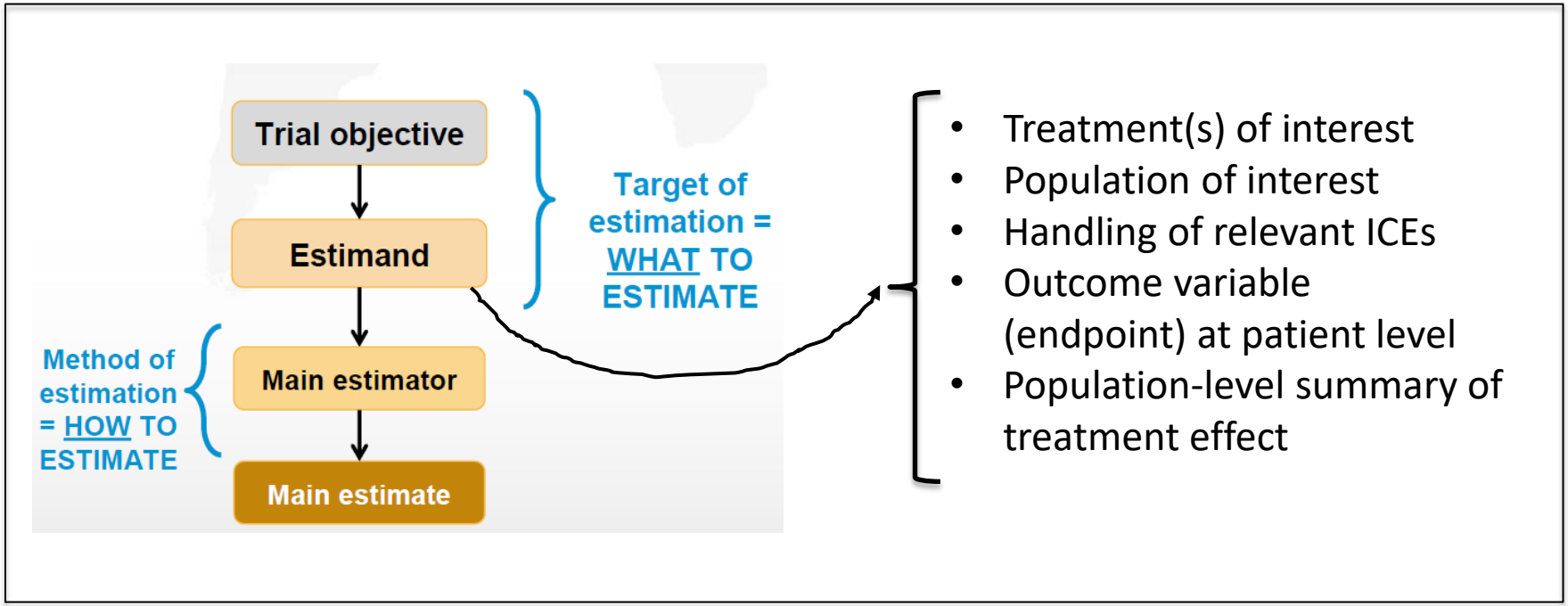
- Different ways of collecting data
- Trial modifications – sample size, analysis methods, missing data, sensitivity analyses
- Documentation in Protocol, SAP, and CSR
- Consult with regulatory agencies

Meyer, et al. (2020)

Outline

- Estimands and intercurrent events (ICEs)
- Causal-inference and potential outcome (PO) framework to define estimand
- Handling ICEs differently according to the underlying reasons of ICEs
- Handling missing values
- Overview of the estimand/estimation framework based on the nature of ICEs and missing values

Estimand framework [ICH E9 (R1)]



ICE, intercurrent events

Potential ICEs related to the COVID-19 pandemic

- Prolonged treatment interruptions due to COVID-19 illness
  - The definition of “prolonged” should depend on the disease state, study objectives, and mechanism of action of the study medications and should be the same for treatment interruptions due to COVID-19 or other reasons
- Prolonged treatment interruptions due to COVID-19 controlled measures.
- Study treatment discontinuations due to COVID-19 illness (an adverse events [AE])
- Study treatment discontinuations due to COVID-19 controlled measures.
- Death as a result of COVID-19 illness
- Use of protocol prohibited medications to treat COVID-19 illness

Regulatory guidance and scientific publications for estimands/estimation related to the COVID-19 pandemic

- FDA Guidance on Conduct of Clinical Trials of Medical Products during COVID-19 Public Health Emergency, Guidance for Industry, Investigators, and Institutional Review Boards (March 2020, Updated on April 16, 2020)
- EMA Guidance on the Management of Clinical Trials during the COVID 19 (Coronavirus) pandemic. Version 2 (27/03/2020)
- EMA Points to consider on implications of Coronavirus disease (COVID-19) on methodological aspects of ongoing clinical trials. https://www.ema.europa.eu/en/documents/scientific-guideline/points-consider-implications-coronavirus-disease-covid-19-methodological-aspects-ongoing-clinical_en-0.pdf
- Considerations on impact of COVID-19 on estimands in oncology clinical trials compiled by industry working group "Estimands in Oncology", most recent version available on http://tinyurl.com/oncoestimand
- Statistical Issues and Recommendations for Clinical Trials Conducted During the COVID-19 Pandemic. COVID-19 Pharmaceutical Industry Working Group. https://www.tandfonline.com/doi/full/10.1080/19466315.2020.1779122
- Comment on: Statistical Issues and Recommendations for Clinical Trials Conducted During the COVID-19 Pandemic. Sylva H. Collins & Mark S. Levenson. https://www.tandfonline.com/doi/full/10.1080/19466315.2020.1779123
- Statistical Considerations for Clinical Trials During the COVID-19 Public Health Emergency. FDA https://www.fda.gov/media/139145/download

What can we learn from the pandemic in terms of estimands?

ICEs

ICH E9 (R1) defines ICEs as “events occurring after treatment initiation that affect either the interpretation or the existence of the measurements associated with the clinical question of interest.”

- Interplay between ICE and treatment of interest (especially for dynamic treatment regimen).
- An example: For a study evaluating an insulin treatment, the insulin dose is adjusted weekly after the initial dose. Should we consider each dose adjustment as an ICE?
- In practice, we should only consider the events that are NOT part of treatment of interest as ICEs.

ICE, intercurrent events

Outline

- Estimands and intercurrent events (ICEs)
- Causal-inference and potential outcome (PO) framework to define estimand
- Handling ICEs differently according to the underlying reasons of ICEs
- Handling missing values
- Overview of the estimand/estimation framework based on the nature of ICEs and missing values

Review of potential outcome framework (Neyman 1923; Rubin 1978)

- Y_i: outcome of interest for *i*-th patient
- S: stratum (subset) of the population, and n is the sample size for S
- A_i: treatment assigned to *i*-th patient (0 = control; 1 = experimental treatment)
- $Y_i(a)$: the potential outcome of Y for a randomly selected (i-th) patient IF assigned to treatment a ($a = 0, 1$)
- Connection between PO with observed outcome: $Y_i = Y_i(1)A_i + Y_i(0)(1 - A_i)$ [consistency]
- The ***causal*** estimand for a subset S is the average treatment effect (ATE)

$$ATE(S) = \frac{1}{n}\sum_{i=1}^{n} E[Y_i(1) - Y_i(0)|S]$$

  S can be all patients, a subset defined by baseline covariates, or a principal stratum. Some special cases are
  - $S = \{A_i = 1\}$ represents *all on experimental treatment*
  - $S = \{A_i = 0\}$ represents *all in control treatment*
- For the whole population (all randomized patients), we may remove S

$$ATE = \frac{1}{n}\sum_{i=1}^{n} E[Y_i(1) - Y_i(0)]$$

- Because $Y_i(a)$ are iid's with the same expected value, this is often simplified in causal literature as (dropping subject index),

$$ATE = E[Y(1) - Y(0)] = E\big(Y(1)\big) - E\big(Y(0)\big)$$

Randomization (without ICEs) can guarantee a causal estimand

- Patients were randomly assigned to two treatments
- The expected treatment difference for the mean of the PO between two treatment groups is

$$E\left\{\frac{1}{n_1}\sum_{i=n_0+1}^{n} Y_i - \frac{1}{n_0}\sum_{i=1}^{n_0} Y_i\right\}$$

Observable outcomes

$$= E\left\{\frac{1}{n_1}\sum_{i=n_0+1}^{n} Y_i(1) - \frac{1}{n_0}\sum_{i=1}^{n_0} Y_i(0)\right\}$$

Consistency of PO

$$= E\left\{\frac{1}{n_1}\sum_{i=1}^{n} Y_i(1)\cdot I(A_i=1) - \frac{1}{n_0}\sum_{i=1}^{n} Y_i(0)\cdot I(A_i=0)\right\}$$

$$= \frac{1}{n_1}\sum_{i=1}^{n} E[Y_i(1)]\cdot\frac{n_1}{n} - \frac{1}{n_0}\sum_{i=1}^{n} E[Y_i(0)]\cdot\frac{n_0}{n}$$

$Y_i(a) \perp A_i$

$$= \frac{1}{n}\sum_{i=1}^{n} E[Y_i(1) - Y_i(0)]$$

- This is a causal estimand

ICE, intercurrent events

Defining estimands based on potential outcomes in causal-inference framework (Lipkovich, et al., 2020)

- Y: outcome of interest
- S: stratum (subset) of the population, and n is the sample size for S
- A: treatment (0 = control; 1 = experimental treatment)
- $Y(a,b)$: the PO of Y assigned to treatment a but actually taking b
  - *As we will see, actual treatment is a PO on its own and can depend on intermediate outcomes of initial treatment, $Z(a)$*
- When $a = b$, we write $Y(a,a) = Y(a)$
- The ***causal*** estimand for a subset S if patient would adhere to their assigned treatment is the average treatment effect (ATE)

$$\frac{1}{n}\sum_{i=1}^{n} E[Y_i(1,1) - Y_i(0,0)|S] = \frac{1}{n}\sum_{i=1}^{n} E[Y_i(1) - Y_i(0)|S]$$

Subscript i may be omitted to simplify the notation

- For the whole population (all randomized patients), we may remove S

$$\frac{1}{n}\sum_{i=1}^{n} E[Y_i(1,1) - Y_i(0,0)] = \frac{1}{n}\sum_{i=1}^{n} E[Y_i(1) - Y_i(0)]$$

PO, potential outcome

Strategies to handle ICEs

- Treatment policy
- Hypothetical
- Composite variable
- While on treatment (WOT)
- Principal stratum (PS)

- ICH E9 (R1) provides a framework for defining estimand
- Key components to be considered
  - Treatment(s) of interest
  - Population of interest
  - Handling of relevant intercurrent events (ICEs)
  - Outcome variable (endpoint) at patient level
  - Population-level summary of treatment effect

PS is to define a population, not a strategy to handle ICEs (although ICEs can be used to define PS)

ICE, intercurrent events

Outline

- Estimands and intercurrent events (ICEs)
- Causal-inference and potential outcome (PO) framework to define estimand
- Handling ICEs differently according to the underlying reasons of ICEs
- Handling missing values
- Overview of the estimand/estimation framework based on the nature of ICEs and missing values

Use a mix of strategies in handling ICEs in a study (Qu et al., 2020)

- One common drawback in most current clinical studies is that only ONE strategy is used to handle all ICEs
- Strategies in handling ICEs should be based on the underlying reasons
  - ICEs due to AE
    - AE at “normal time”
    - AE of COVID-19 illness
  - ICEs due to lack of efficacy (LoE)
    - Treatment discontinuation due to LoE
    - Use of rescue medication due to LoE
  - ICEs due to administrative reasons
    - Relocation, family situation changed, COVID-19 controlled measures, etc.

ICE, intercurrent events

WOT and composite strategies

- WOT strategy: there are very few cases in which WOT is useful
- Composite strategy: Using ICEs to define the endpoint. It may be more appropriate to define the composite endpoint explicitly. For example,
  - In rheumatoid arthritis (RA), the binary variable of ACR20 is often used
  - Composite strategy may treat a patient with an ICE of using rescue medication as a non-responder
  - It is more appropriate to define the endpoint as a composite endpoint "achieving ACR20 at the end of study without using rescue medications"

ICE, intercurrent events; WOT, while on treatment

Treatment policy strategy

- ICH E9 (R1) describes the treatment policy strategy as "the occurrence of the intercurrent event is considered irrelevant in defining the treatment effect of interest: the value for the variable of interest is used regardless of whether or not the intercurrent event occurs."
- Let $A_i^* = \{A_i, g_i(Z_i(A_i))\}$ be the treatment regimen (policy) patient i takes
  - g_i maps intermediate outcomes Z_i to a treatment regimen (i.e., stopping study meds when having AE)
  - g_i generally is not precisely defined in the protocol (certain things may be left to physician's discretion)
- The estimand using this treatment policy strategy is defined by

$$E\left\{Y_i\left(1, g_i(Z_i(1))\right) - Y_i\left(0, g_i(Z_i(0))\right)\right\}$$ → g() with subscript i

- Estimand for the dynamic treatment regimen (DTR) (Murphy et al., 2001; Moodie et al., 2007)

$$E\left\{Y_i\left(1, g(Z_i(1))\right) - Y_i\left(0, g(Z_i(0))\right)\right\}$$ → g() without subscript i

The time-varying treatment regimen function g is defined clearly and in a same way for all patients

Treatment policy strategy (continued)

- One key argument for using treatment policy strategy is that it reflects the real world clinical practice
- The visit schedules, inclusion criteria, allowed rescue medication use, etc. make a clinical trial setting drastically different from that in real clinical practice
- For example, in a basal insulin study comparing basal insulin peglispro with insulin glargine (Bergenstal et al., 2016)
  - Patients had clinical visits every week for the first 12 weeks
  - Patients were not allowed to take “rescue” medication unless they discontinue the study medication
- To use treatment policy strategy, it is recommended to clearly define the treatment regimen. For example,
  - The treatment of interest is the randomized study medication with addition to any additional rescue concomitant medications based on protocol defined rescue criteria

Hypothetical strategies

- As estimands should be defined in terms of the potential outcome, most strategies in handling ICEs should be "hypothetical")
- We introduce 4 different hypothetical strategies
  - *Controlled direct hypothetical* (CDH) strategy
  - *No treatment hypothetical* (NTH) strategy
  - *Partial treatment hypothetical* (PTH) strategy
  - *Null hypothesis hypothetical* (NHH) strategy

Controlled direct hypothetical (CDH) strategy

- The PO of interest is the outcome if patients could complete the treatment even in the presence of ICEs
- The estimand is

$$E\{Y_i(1,1) - Y_i(0,0)\}$$

- "Controlled direct" was borrowed from *controlled direct effect* (Pearl, 2009)
- This approach may be appropriate for
  - ICEs due to administrative reasons (e.g., ICEs related to COVID-19 controlled measures)
  - ICEs that do not represent the "normal" time (e.g., COVID-19 illness)
  - ICEs due to LoE

LoE, lack of efficacy; PO, potential outcome

No treatment hypothetical (NTH) strategy

- The interested PO is the outcome assuming patients with ICEs would have no benefit from the treatment (as if the patients were left untreated starting from randomization):

$$E\left[\left\{Y_i(1,-1)\Delta_i(1)+Y_i(1)\left(1-\Delta_i(1)\right)\right\}-\left\{Y_i(0,-1)\Delta_i(0)+Y_i(0)\left(1-\Delta_i(0)\right)\right\}\right]$$

  where "-1" in the second parameter $Y_i(\cdot,\cdot)$ indicates no treatment received and $\Delta_i(a)$ is the ICE indicator (0 for no ICE and 1 for ICE occurring).

- This approach may be appropriate for ICEs due to AE (occurring at "normal time")

AE, adverse event; ICE, intercurrent events; PO, potential outcome

Partial treatment hypothetical (PTH) strategy

- The interested PO is the outcome if the patient benefit from (or be harmed by) the study medication until the ICE and then stops taking the medication.
- The estimand is defined as

$$E\left[\left\{Y_i(1, g_i(T_i(1)))\Delta_i(1) + Y_i(1)\left(1 - \Delta_i(1)\right)\right\}\right.$$
$$\left. - \left\{Y_i(0, g_i(T_i(0)))\Delta_i(0) + Y_i(0)\left(1 - \Delta_i(0)\right)\right\}\right]$$

  where $T_i(a)$ is the time to the ICE under treatment a and $g_i\left(T_i(\mathrm{a})\right)$ is the treatment regimen: taking treatment a until the occurrence of the ICE and then having no access to treatment until a specified assessment time.

- This strategy may be suitable for handling ICEs due to AE at a “normal circumstances” (not for AE related to the COVID-19 pandemic), especially for treatment with potential long-term or disease-modification effect.

ICE, intercurrent events; PO, potential outcome

Null hypothesis hypothetical (NHH) strategy

- The interested PO is the outcome (for a patient with an ICE) if the experimental treatment could have "null" effect from randomization compared to the control treatment (Lipkovich et al., 2020; Qu et al., 2020c).
  - The PO for the control and experimental treatments follow the relationship under the null hypothesis

$$Y_i(1, \text{No Treatment}) = Y_i(0, 0) + \delta,$$

  where δ is the average treatment difference under the null hypothesis.
- This leads to an estimand

$$E\left[\left\{(Y_i(0) + \delta)\Delta_i(1) + Y_i(1)\left(1 - \Delta_i(1)\right)\right\} - \left\{Y_i(0)\Delta_i(0) + Y_i(0)\left(1 - \Delta_i(0)\right)\right\}\right]$$

- For superiority studies, $\delta = 0$, and for non-inferiority studies, δ is the non-inferiority margin (assuming the smaller the outcome, the better).
- This approach can only be applied to estimands with a hypothesis, which is most often the case in clinical trials.

PO, potential outcome

Outline

- Estimands and intercurrent events (ICEs)
- Causal-inference and potential outcome (PO) framework to define estimand
- Handling ICEs differently according to the underlying reasons of ICEs
- Handling missing values
- Overview of the estimand/estimation framework based on the nature of ICEs and missing values

Classification of Missingness

- In the context of a longitudinal clinical trial, missingness can be classified into four categories (Rubin, 1987; Little, 1995):
  - Missing not at random (MNAR). Conditional on the observed values, the probability of missingness is dependent of unobserved (missing) outcomes.
  - Missing at random (MAR). Conditional on the observed values, the probability of missingness is independent of any unobserved outcomes.
  - Covariate dependent MAR (Cov-MAR). Conditional on the baseline covariates, the probability of missingness is independent of any observed or unobserved outcomes (including treatment assignment).
  - Missing completely at random (MCAR). The probability of missingness is independent of any observed and unobserved variables.
- MCAR and Cov-MAR are special cases of MAR

Missing values

- Missing values
  - As a result of handling ICEs with hypothetical strategies
  - *True* missing values due to data not being collected
- Assumptions for missingness and methods to handle missing values should be based on the underlying reasons of ICEs or missingness
  - ICEs due to AE
    - AE at "normal circumstances"
    - AE of COVID-19 illness
  - ICEs due to LoE
  - ICEs due to administrative reasons

AE, adverse event; ICE, intercurrent events; LoE, lack of efficacy

Imputation or direct likelihood-based method in handling missing values

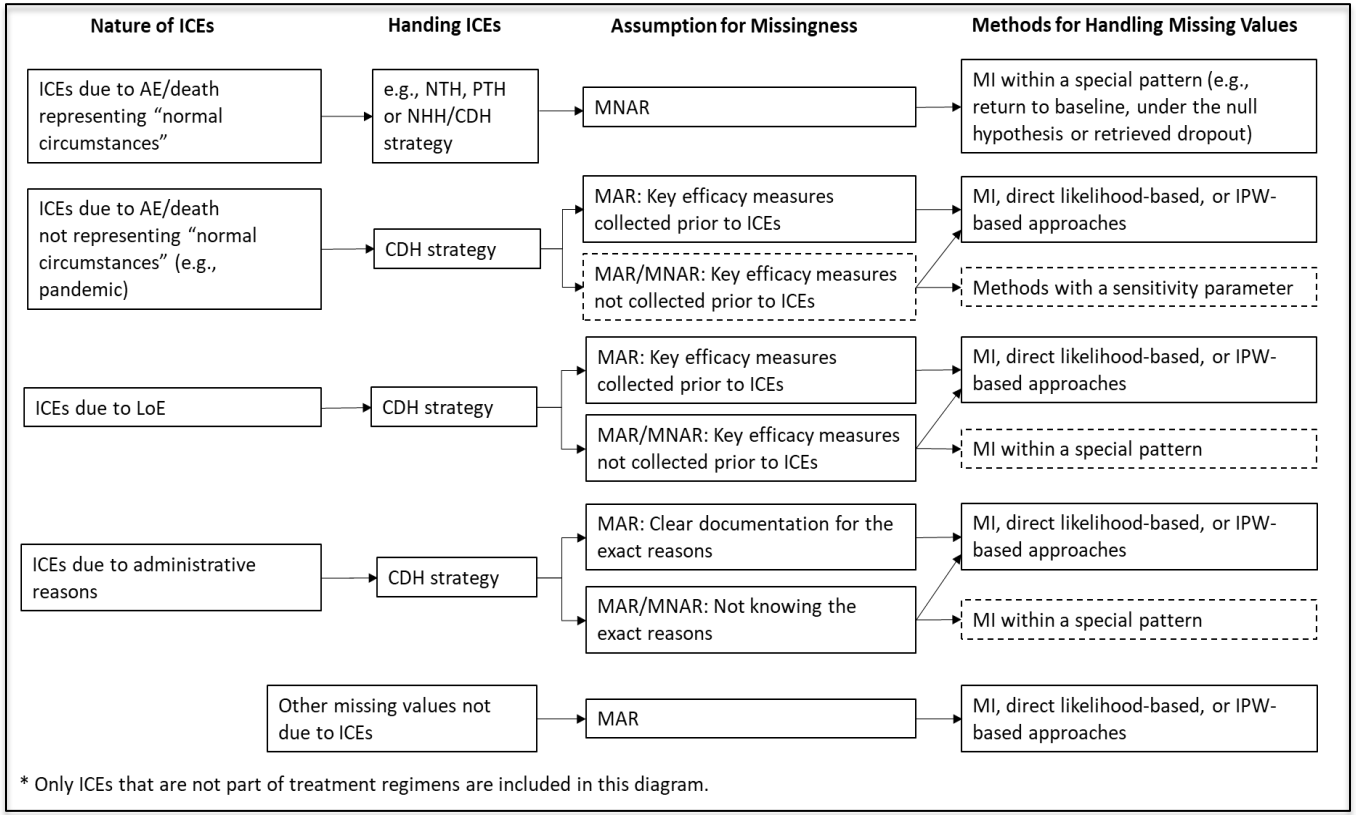
- MAR
  - Multiple imputation using patients in the same treatment group
  - Direct likelihood-based method, e.g., mixed model for repeated measures (MMRM)
  - Inverse probability weighing (IPW) based on the probability of missingness
- MNAR
  - Multiple imputation under a special pattern, e.g., reference-based imputation
  - Direct likelihood-based methods
  - Specifying sensitivity parameters during multiple imputations or direct likelihood-based methods

MAR, missing at random; MNAR, missing not at random

Outline

- Estimands and intercurrent events (ICEs)
- Causal-inference and potential outcome (PO) framework to define estimand
- Handling ICEs differently according to the underlying reasons of ICEs
- Handling missing values
- Overview of the estimand/estimation framework based on the nature of ICEs and missing values

Handling ICEs and missing values according to the nature of ICE/missingness



AE, adverse event
CDH, controlled direct hypothetical
ICE, intercurrent events
IPW, inverse probability weighting
LoE, lack of efficacy;
MAR, missing at random
MI, multiple imputation
MNAR, missing not at random
NHH, null hypothesis hypothetical
NTH, no treatment hypothetical
PTH, partial treatment hypothetical

An example – a study for heart failure indication

- The primary endpoint is the 6-minute walk distance (6MWD)
- The interest is the average of
  - Worst benefit for those who dies
  - No additional benefit for patients who discontinue treatment due to AE (PTH)
  - Hypothetical treatment effect otherwise (CDH)

Intercurrent events	Estimand	Missing value	Estimation (handling the resulting missing values)
Death	Worst outcome (hypothetical)	Yes	Impute as 0
Treatment discontinuation due to AE	No benefit (hypothetical)	If no measurement collected at study end	MI under a special pattern - “Retrieved dropout” imputation
Treatment discontinuation due to other reasons	Hypothetical	Yes (as a result of censoring)	MI using patients in the same treatment group (MAR)
		Missing measurements due to being unable to perform the 6MWD test	Impute as 0
		Other missing measurements (e.g., due to COVID-19 controlled measures)	MI using patients in the same treatment group (MAR)

MAR, missing at random; MI, multiple imputation

Conclusions and discussion

- Pandemic revealed gaps in ICH E9 (R1) leading to confusion and causing study teams to spend considerable effort redefining estimands in the protocols to accommodate COVID-19
- We see several sources of confusion across teams
  - PO language for causal inference is not explicitly used in ICH E9 (R1)
  - Circularity on whether ICEs or treatment regimens are defined first
  - Use of “treatment policy” as an approach to bypass the need to clearly define the treatment regimen of interest
  - “While on treatment” strategy is often used as a disguise for “old good” LOCF
  - In most cases when “composite strategy” is proposed, it can be avoided by explicitly defining a composite endpoint
  - The “principal stratum” is not a method for handling ICEs but a method for defining subpopulations

Recommendations

- **Describing estimands**
  - Using PO language may help define and communicate estimands more succinctly. It also helps evaluate the plausibility of certain strategies for handling ICEs.
- **Defining ICEs**
  - Prior to discussing ICEs, treatment regimens of interest need to be defined precisely.
  - To be considered an ICE, this event should be a deviation from the treatment regimens of interest.
- **Handling ICEs**
  - Hypothetical strategies should be predominately used to define causal estimands.
  - Treatment policy strategies should generally be avoided except for two situations:
    - ICEs are explicitly included in treatments of interest under a rigorous treatment regimen
    - For selected ICEs in a pragmatic study in which the study setting is similar to the real-world setting.
  - A mix strategies handling ICEs are often clinically relevant.
- **Estimation**
  - Multiple imputation is a flexible tool allowing for implementing a mix of strategies in handling ICEs.

References

Bergenstal, R. M., Lunt, H., Franek, E., Travert, F., Mou, J., Qu, Y., ... & Bastyr III, E. J. (2016). Randomized, double-blind clinical trial comparing basal insulin peglispro and insulin glargine, in combination with prandial insulin lispro, in patients with type 1 diabetes: IMAGINE 3. Diabetes, Obesity and Metabolism, 18(11), 1081-1088.

ICH E9 (R1) (2019). "ICH Harmonised Guideline Addendum on Estimands and Sensitivity Analysis in Clinical Trials to the Guideline on Statisical Principles for Clinical Trials E9(R1) ," International Conference on Harmonisation [online], Available at https://database.ich.org/sites/default/files/E9-R1_Step4_Guideline_2019_1203.pdf.

Lipkovich, I., Ratitch, B., & Mallinckrodt, C. H. (2020). Causal inference and estimands in clinical trials. *Statistics in Biopharmaceutical Research*, *12*(1), 54-67.

Little, R. J. A. (1995), "Modeling the Drop-Out Mechanism in Repeated Measures Studies," Journal of the American Statistical Association, 90, 1112–1121.

Meyer, R. D., Ratitch, B., Wolbers, M., Marchenko, O., Quan, H., Li, D., ... & Englert, S. (2020). Statistical Issues and Recommendations for Clinical Trials Conducted During the COVID-19 Pandemic. Statistics in Biopharmaceutical Research, (just-accepted), 1-22.

Neyman, J. (1923), "On the Application of Probability Theory to Agricultural Experiments. Essay on Principles. Section 9," Statistical Science, 5, 465–480.

Pearl, J. (2009), "Causal Inference in Statistics: An Overview," Statistics Surveys, 3, 96–146.

Qu, Y., Shurzinske, L., & Sethuraman, S. (2020). Defining Estimands Using a Mix of Strategies to Handle Intercurrent Events in Clinical Trials. arXiv preprint arXiv:2006.03105.

Rubin, D. B. (1978), "Bayesian Inference for Causal Effects: The Role of Randomization," Annals of Statistics, 6, 34–58.

Rubin, D. B. (1987), "Multiple Imputation for Nonresponse in Surveys", New York: Wiley.

Thank you!

Q & A