NIA IMPACT COLLABORATORY
TRANSFORMING DEMENTIA CARE

Outcomes truncated by death in randomized trials

Michael Harhay, PhD, MPH
Assistant Professor of Epidemiology and Medicine (Pulmonary and Critical Care)
University of Pennsylvania

Housekeeping

- All participants will be muted
- Enter **all questions** in the Zoom **Q&A/chat box** and send to Everyone
- Moderator will review questions from chat box and ask them at the end
- Want to continue the discussion? Associated podcast released about 2 weeks after Grand Rounds
- Visit impactcollaboratory.org
- Follow us on Twitter & LinkedIN:

@IMPACTcollab1 https://www.linkedin.com/company/65346172

Funding & Disclosures

- Funding (views are entirely my own)
  - PCORI ME-2020C1-19220
  - NIH/NHLBI K99/R00-HL141678, R01-HL168202

- Personal fees
  - Unlearn.AI
  - DSMB for several academic (e.g., NIH and DOD sponsored) trials.
  - American Thoracic Society
    - Deputy Editor, American Journal of Respiratory and Critical Care Medicine
  - Elsevier
    - Statistical review, The Lancet Family Journals

Learning Objectives

Upon completion of this presentation, you should be able to:

- Articulate the risks of bias associated with having outcomes truncated by death in randomized trials.
- Assess options to address outcomes truncated by death using the estimand framework.
- Understand the tradeoffs between approaches such as statistical models and composite outcomes to address outcomes truncated by death.

Today, in a slide

- Randomization balances trial arms at the *onset* of a study.
- Post-randomization (i.e., intercurrent) events occur after the start of the trial and affect the presence and/or interpretability of observed values, and thus the results from a trial analysis.
- Death is a particularly challenging post-randomization event.
  - Death itself is an important patient-centered outcome.
  - Non-mortality outcomes become informatively missing/undefined or truncated/censored.
- There are a range of (*suboptimal*) solutions. I will present these, concluding with my (*current*) preference → *hierarchical composite endpoints*.
- Pragmatically dealing with this issue in pragmatic trials.

Intervention X for QoL for patients who had in-hospital delirium.

- 18% of participants either died or were lost to follow-up. Those on intervention X **had a higher QoL** (EQ-5D) by **0.08** (95% CI 0.02 to 0.14) at 6 months.

- …. so if we give participants intervention X, it will improve their QoL on average by 0.08?

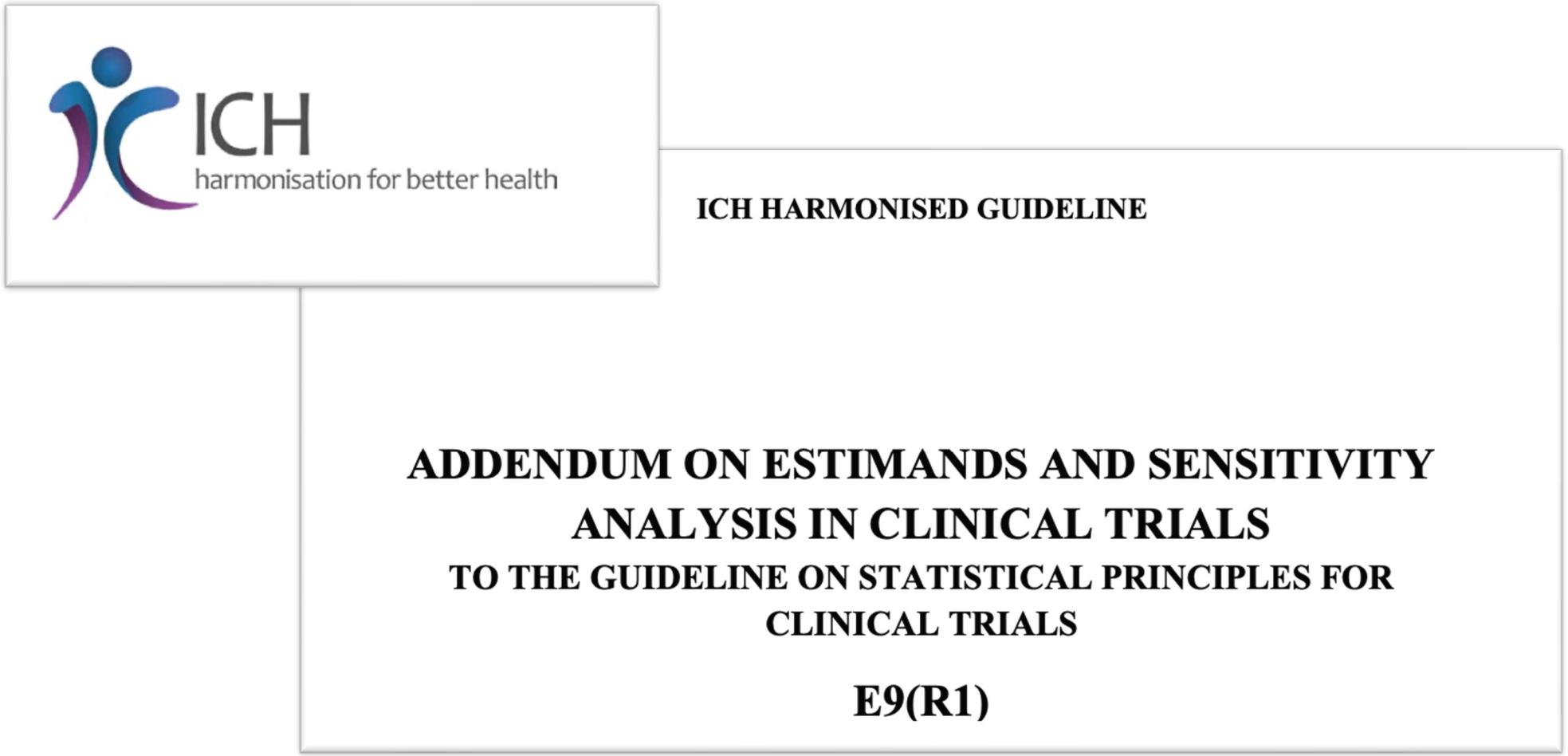
- **We don't know what that estimate actually means.**

Seemingly subtle analytic choices can produce big differences in the meaning of trial results

Statistical methods	Treatment effect
Mixed-model for repeated-measures	**0.08** is an estimate of what the treatment effect ***would be in the hypothetical setting where those in the trial never experience disease progression, drop-out, or death (missing at random).***
Among those with complete data	**0.08** is the effect in a sub-sample that fully complied with the trials' design.

How do we learn what we want to learn from a trial?

Estimands – ICH E9 (R1) Addendum (2019)

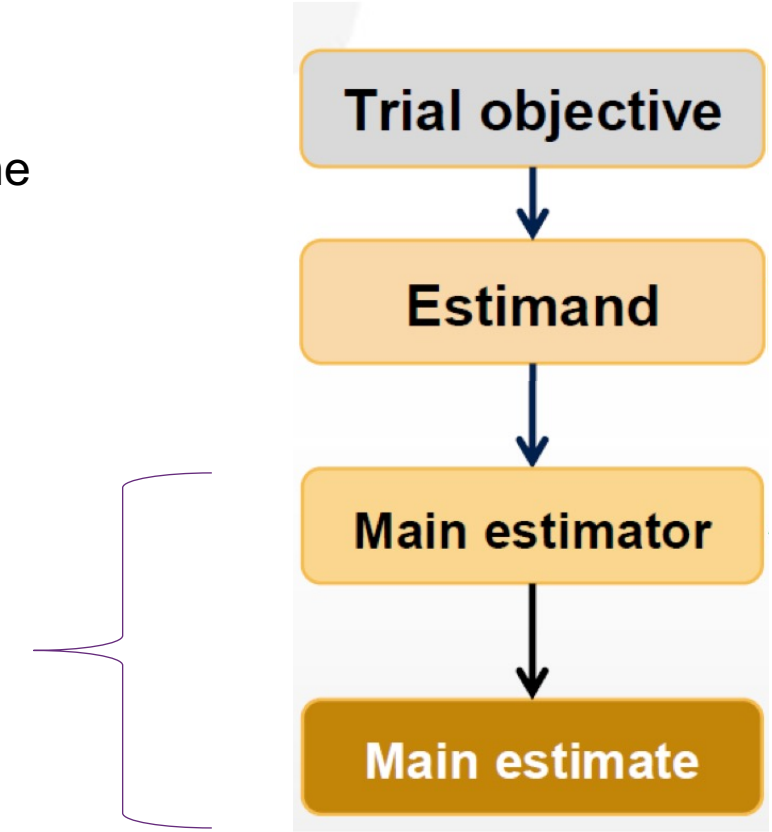


International Council for Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use. Addendum to ICH E9 – Statistical Principles for Clinical Trials.

Estimand: the quantity that is to be estimated in a statistical analysis.

Estimator: the method used to obtain an approximation of this target.

Estimate: the value produced by the selected estimator.



Target of trial and estimation =

WHAT

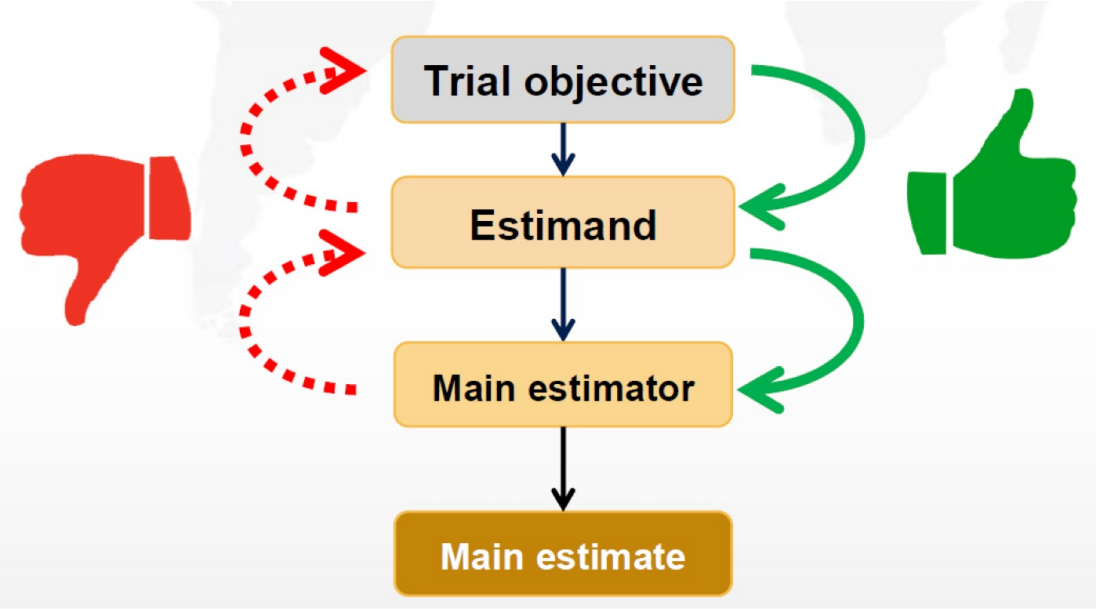
Method of estimation

=

HOW

Estimand

- By clarifying the question we can:
  - Ensure everyone understands what's being estimated.
  - Ensure what's being estimated is relevant.
  - Ensure study design/data collection/analysis are aligned with the question.



International Council for Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use. Addendum to ICH E9 – Statistical Principles for Clinical Trials.

Components of an estimand*

PICO	**Population**	**The population of patients we want to estimate the treatment effect for.**
	Treatment conditions	**The treatment conditions being compared for the treatment effect.**
	Endpoint	**The outcome measure collected for each patient that the treatment effect is based on.**
	Summary measure	**The summary measure used to compare endpoints between treatment conditions for the treatment effect (e.g., mean difference, risk ratio, odds ratio, etc.).**
	Handling of intercurrent events	**How postbaseline events, which affect the interpretation or occurrence of the endpoint (e.g., treatment discontinuation, treatment switching, use of rescue medication, or death, if not defined as part of the outcome), are handled in the definition of the treatment effect.**

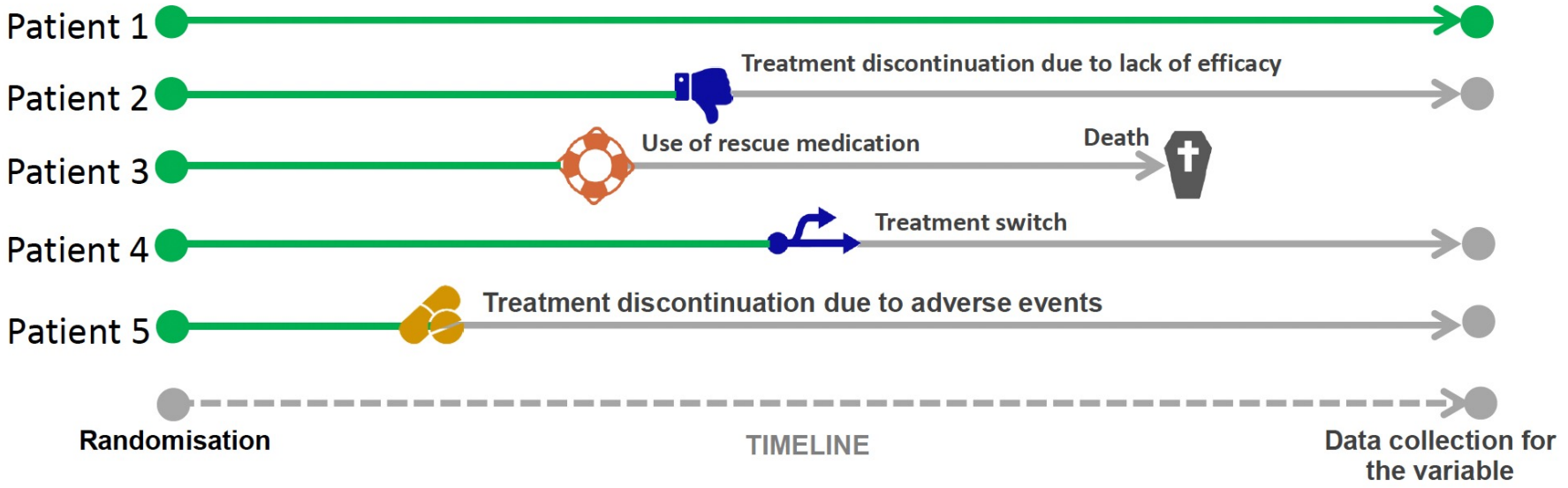
*Extending the PICO (population, intervention, control, and outcomes) format

Intercurrent events

- Events occurring after treatment initiation that affect either the interpretation or the existence of the measurements associated with the clinical question of interest.

- Examples
  - Treatment discontinuation
  - Failure to initiate treatment
  - Treatment switching / use of rescue medication
  - Wrong dose of treatment
  - Death

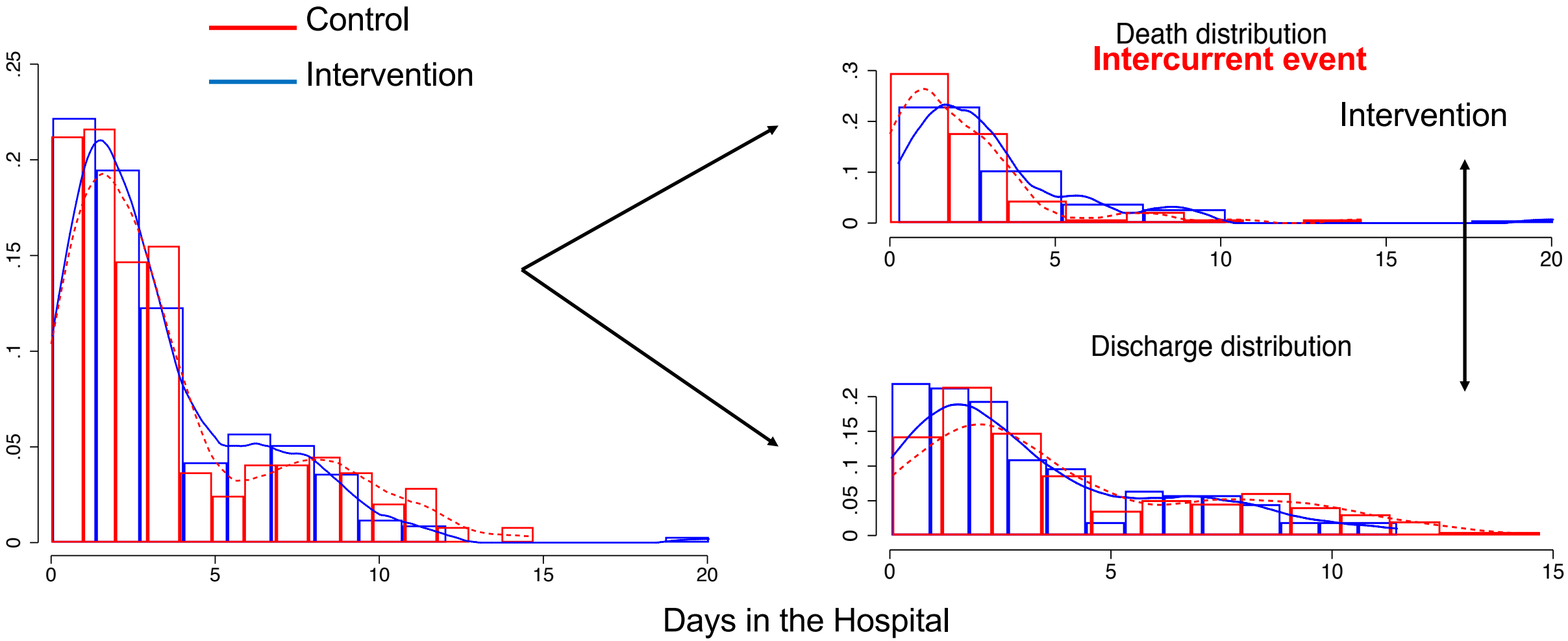
Patient journeys and intercurrent events



*An **intercurrent event** is an event that occurs after randomization/treatment initiation and either precludes observation of the variable or affects its interpretation.*	***Missing data** is data that would be meaningful for the analysis of a given estimand but were not collected.*

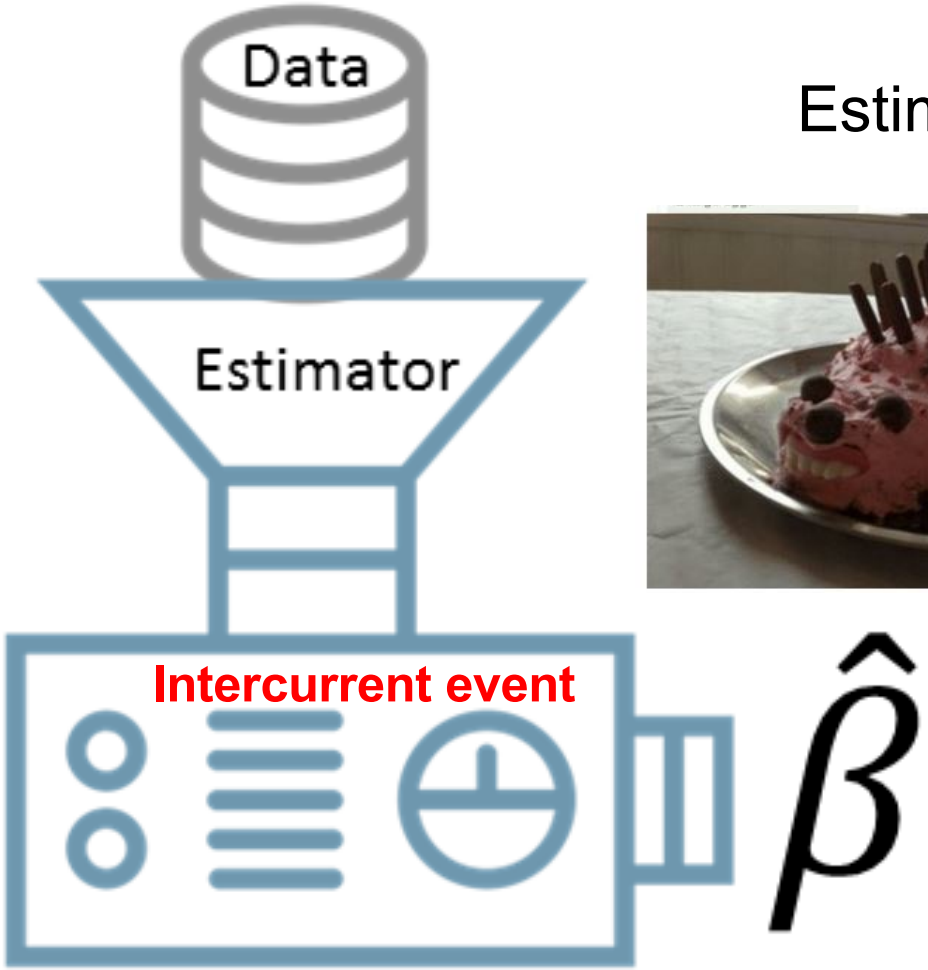
Slide from Frank Bretz, Jiawei Wei

Non-mortality outcomes represent ≥ 2 processes

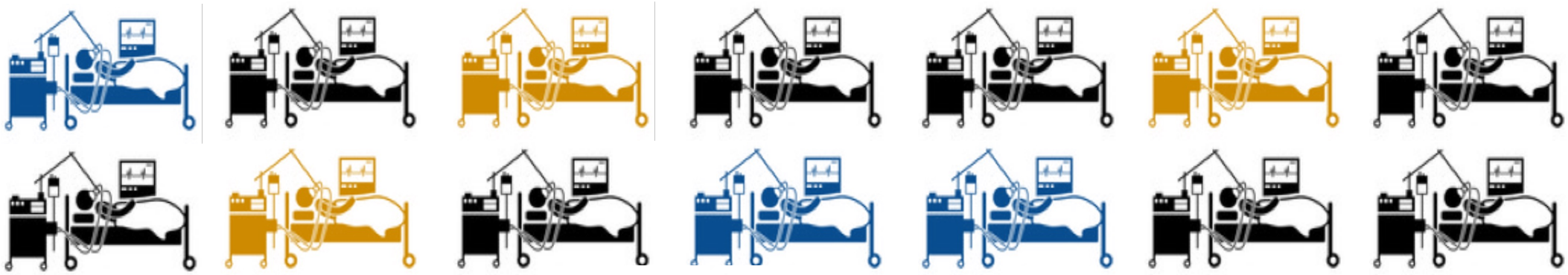


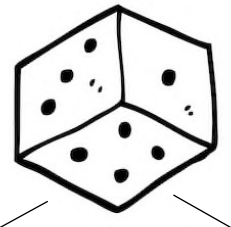
Estimand

β



Estimate





Baseline risk in each study group is the same as the target population due to randomization.

Study group 1

Study group 2

Death is a post-randomization event

- The goal of randomization is to make study groups exchangeable at baseline.
- That is, to make ‘initial conditions’ as close to identical as possible between trial arms.
- Once a trial starts, the benefits of randomization can start to break down due to intercurrent events.
    - Balance of initial conditions starts to be lost.

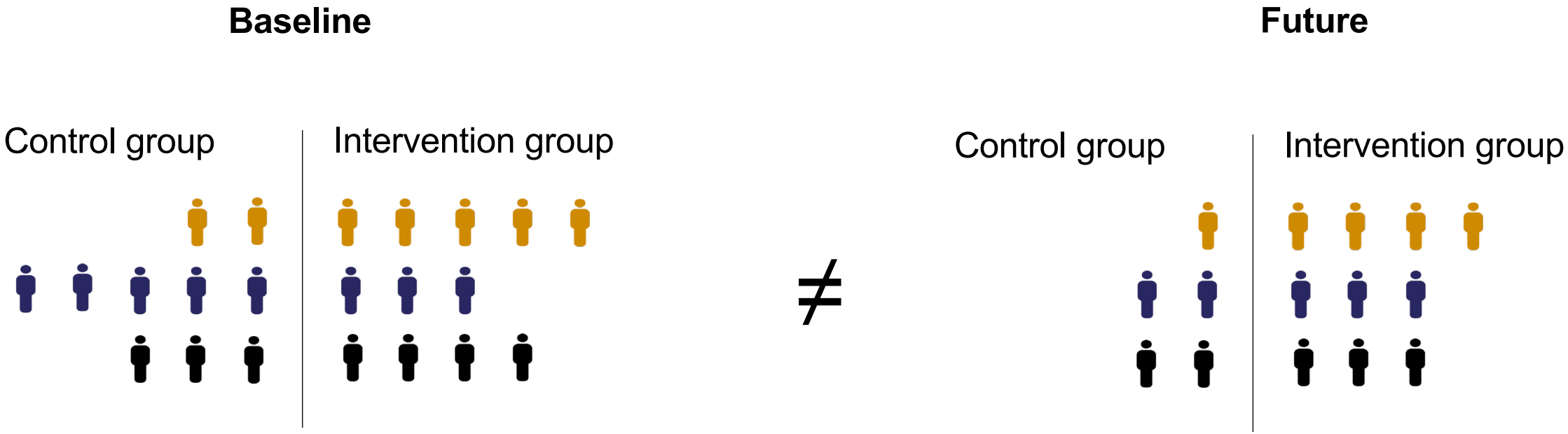
Randomization protection is not permanent

If you restrict your analysis based on post-randomization events, you throw away part of the 'initial conditions' support for comparing study groups without bias.

And this causes a new selection bias.

- Informative dropout/censoring
- Missing not at random

Baseline ≠ Future with intercurrent events



Differences, *at baseline*, are due to chance, not selection biases (e.g., confounding by indication).

Thus, imbalance can impact precision around the effect estimate if the imbalance is in prognostic baseline characteristics but won't generally bias effect estimates.

Strategies to address intercurrent events

Treatment policy

While on treatment/while alive

Hypothetical

Principal stratum

Composite

Strategies to address intercurrent events

Treatment policy

While on treatment/while alive

Hypothetical

Principal stratum

Composite

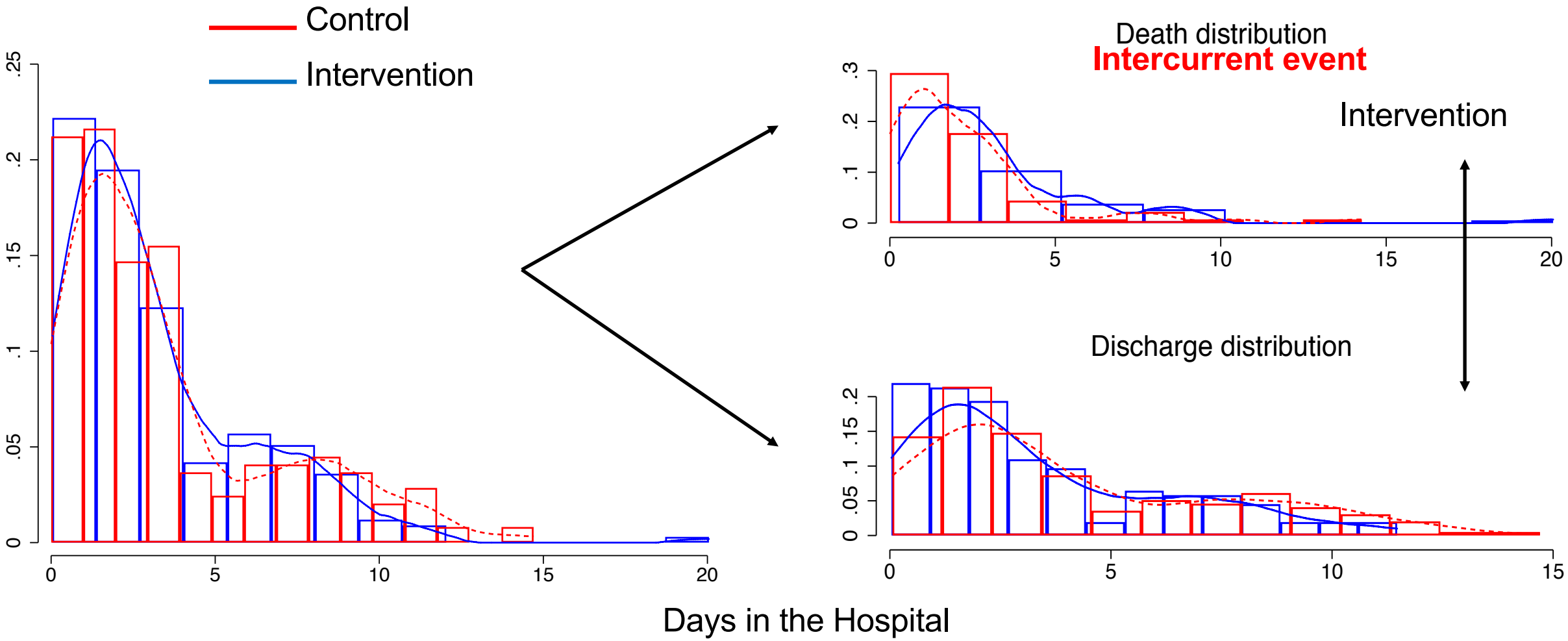
Treatment policy strategy

- This approach considers the patient's outcome regardless of whether they had the intercurrent event (i.e., it is ignored).

- Effect of intervention, regardless of treatment discontinuation.

- My view: cannot be used for terminal events like death. At best, the result produced from this approach is ambiguous.

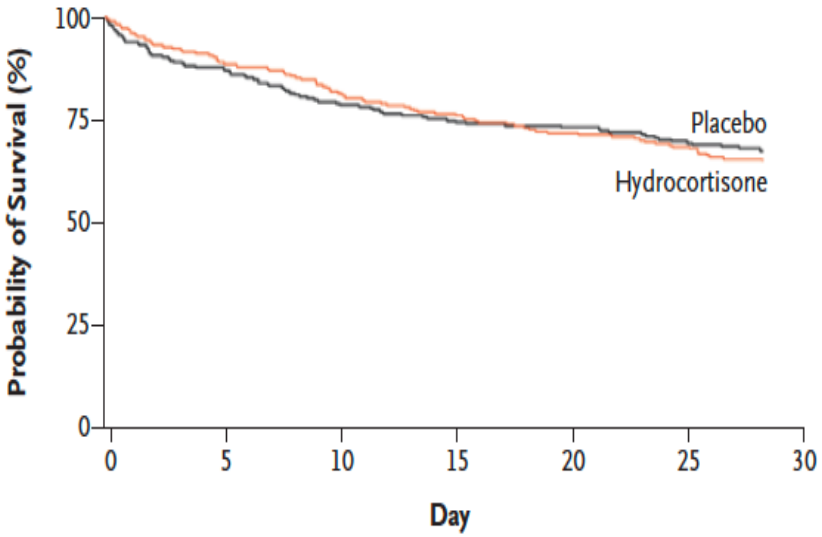
Non-mortality outcomes represent ≥ 2 processes



The NEW ENGLAND JOURNAL of MEDICINE
ESTABLISHED IN 1812 JANUARY 10, 2008 VOL. 358 NO. 2

Hydrocortisone Therapy for Patients with Septic Shock

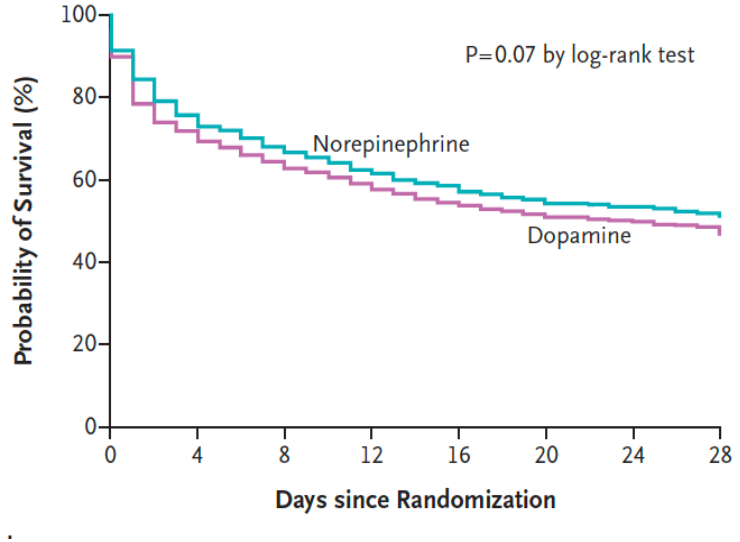
Charles L. Sprung, M.D., Djillali Annane, M.D., Ph.D., Didier Keh, M.D., Rui Moreno, M.D., Ph.D., Mervyn Singer, M.D., F.R.C.P., Klaus Freivogel, Ph.D., Yoram G. Weiss, M.D., Julie Benbenishty, R.N., Armin Kalenka, M.D., Helmuth Forst, M.D., Ph.D., Pierre-Francois Laterre, M.D., Konrad Reinhart, M.D., Brian H. Cuthbertson, M.D., Didier Payen, M.D., Ph.D., and Josef Briegel, M.D., Ph.D., for the CORTICUS Study Group*



The NEW ENGLAND JOURNAL of MEDICINE
ESTABLISHED IN 1812 MARCH 4, 2010 VOL. 362 NO. 9

Comparison of Dopamine and Norepinephrine in the Treatment of Shock

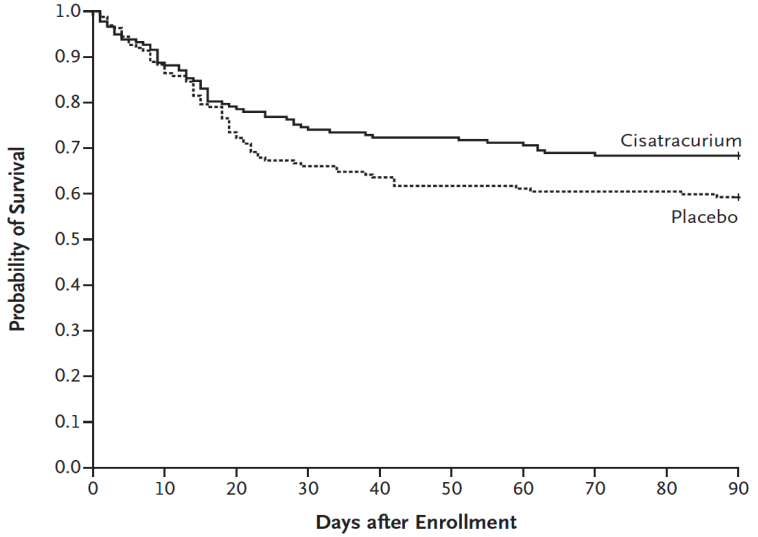
Daniel De Backer, M.D., Ph.D., Patrick Biston, M.D., Jacques Devriendt, M.D., Christian Madl, M.D., Didier Chochrad, M.D., Cesar Aldecoa, M.D., Alexandre Brasseur, M.D., Pierre Defrance, M.D., Philippe Gottignies, M.D., and Jean-Louis Vincent, M.D., Ph.D., for the SOAP II Investigators*



The NEW ENGLAND JOURNAL of MEDICINE
ESTABLISHED IN 1812 SEPTEMBER 16, 2010 VOL. 363 NO. 12

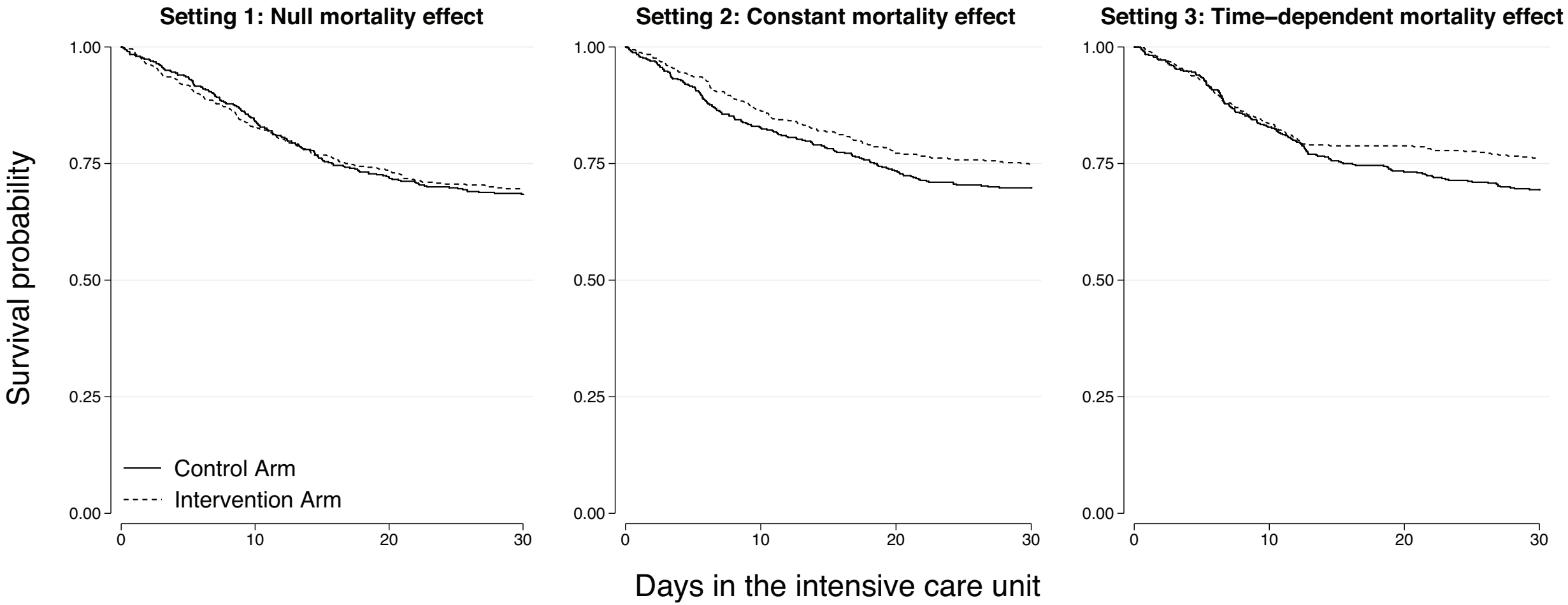
Neuromuscular Blockers in Early Acute Respiratory Distress Syndrome

Laurent Papazian, M.D., Ph.D., Jean-Marie Forel, M.D., Arnaud Gacouin, M.D., Christine Penot-Ragon, Pharm.D., Gilles Perrin, M.D., Anderson Loundou, Ph.D., Samir Jaber, M.D., Ph.D., Jean-Michel Arnal, M.D., Didier Perez, M.D., Jean-Marie Seghboyan, M.D., Jean-Michel Constantin, M.D., Ph.D., Pierre Courant, M.D., Jean-Yves Lefrant, M.D., Ph.D., Claude Guérin, M.D., Ph.D., Gwenaël Prat, M.D., Sophie Morange, M.D., and Antoine Roch, M.D., Ph.D., for the ACURASYS Study Investigators*



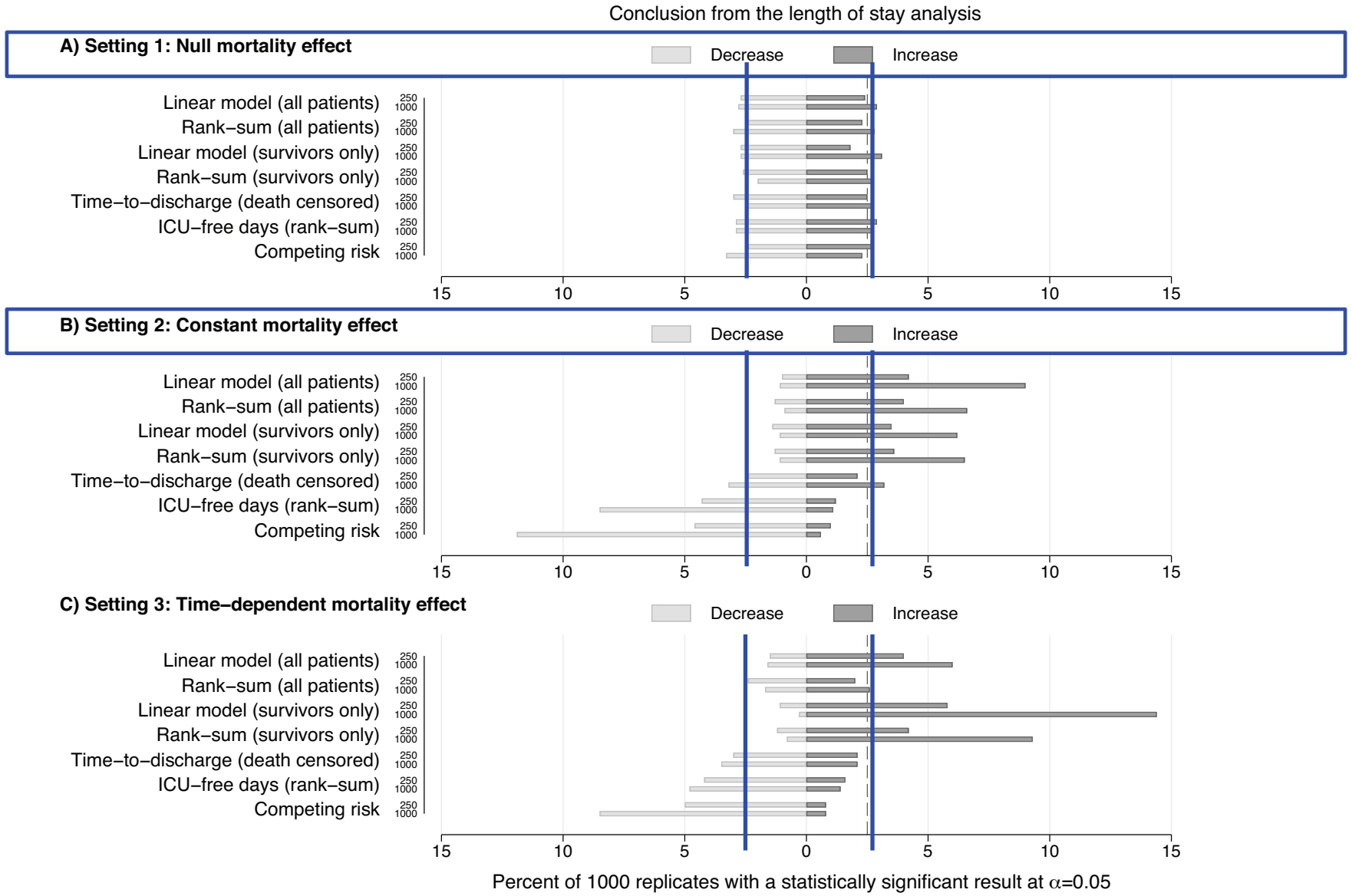
NIA IMPACT COLLABORATORY
TRANSFORMING DEMENTIA CARE

Harhay MO, Ratcliffe SJ, Small DS, Suttner LH, Crowther MJ, Halpern SD. Measuring and Analyzing Length of Stay in Critical Care Trials. Med Care. 2019 Sep;57(9):e53-e59.



NIA IMPACT COLLABORATORY
TRANSFORMING DEMENTIA CARE

Harhay MO, Ratcliffe SJ, Small DS, Suttner LH, Crowther MJ, Halpern SD. Measuring and Analyzing Length of Stay in Critical Care Trials. Med Care. 2019 Sep;57(9):e53-e59.



NIA IMPACT COLLABORATORY
TRANSFORMING DEMENTIA CARE

Harhay MO, Ratcliffe SJ, Small DS, Suttner LH, Crowther MJ, Halpern SD. Measuring and Analyzing Length of Stay in Critical Care Trials. Med Care. 2019 Sep;57(9):e53-e59.

Binary outcomes of interest are not observed

- Intervention is effective – reduces mortality.

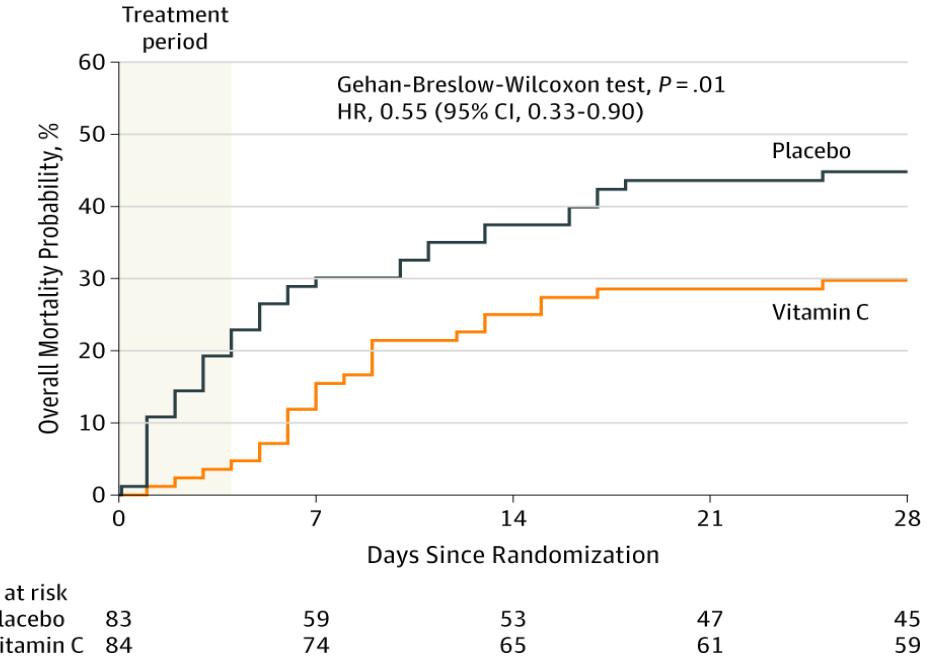
- We observe a higher incidence of AKI in intervention arm.

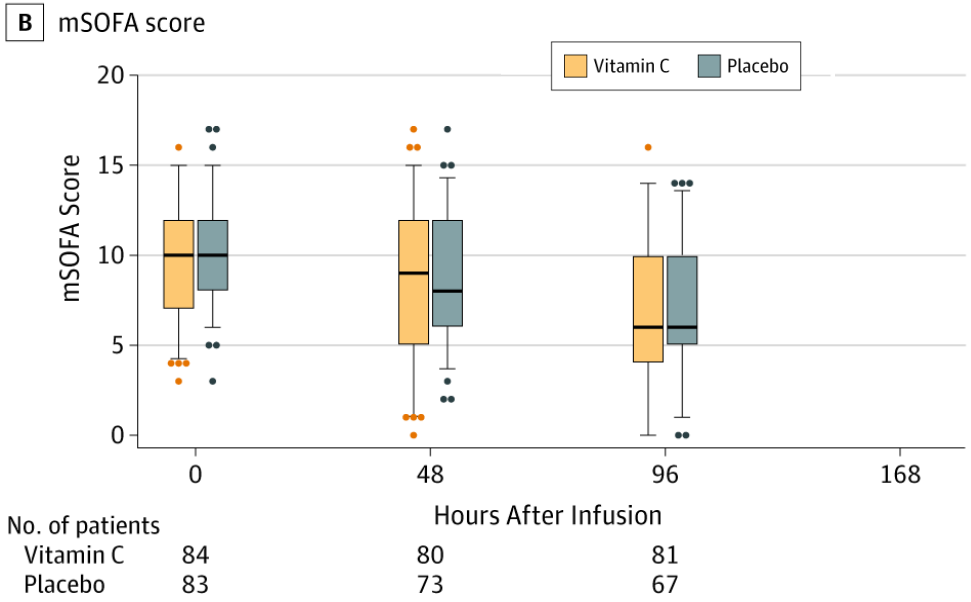
- Why?
  - Is this because of the intervention (i.e., adverse effect)?
  - Is it because more patients died in the control arm, and we didn't observe them long enough to see them develop a similar rate of AKI?

JAMA | Preliminary Communication | CARING FOR THE CRITICALLY ILL PATIENT

Effect of Vitamin C Infusion on Organ Failure and Biomarkers of Inflammation and Vascular Injury in Patients With Sepsis and Severe Acute Respiratory Failure

The CITRIS-ALI Randomized Clinical Trial





NIA IMPACT COLLABORATORY
TRANSFORMING DEMENTIA CARE

Strategies to address intercurrent events

Treatment policy

While on treatment/while alive

Hypothetical

Principal stratum

Composite

While on treatment/while alive strategy

- The endpoint prior to the intercurrent event is of interest.
  - E.g., The effect of intervention on recurrence up to 6 months or death.

- Modifies endpoint aspect of estimand.
  - From “disease up to 12 weeks” to “disease recurrence up to 12 weeks or death, whichever is first.”

- My view: cannot be used for terminal events like death. At best, the result produced from this approach is ambiguous.

Strategies to address intercurrent events

Treatment policy

While on treatment/while alive

Hypothetical

Principal stratum

Composite

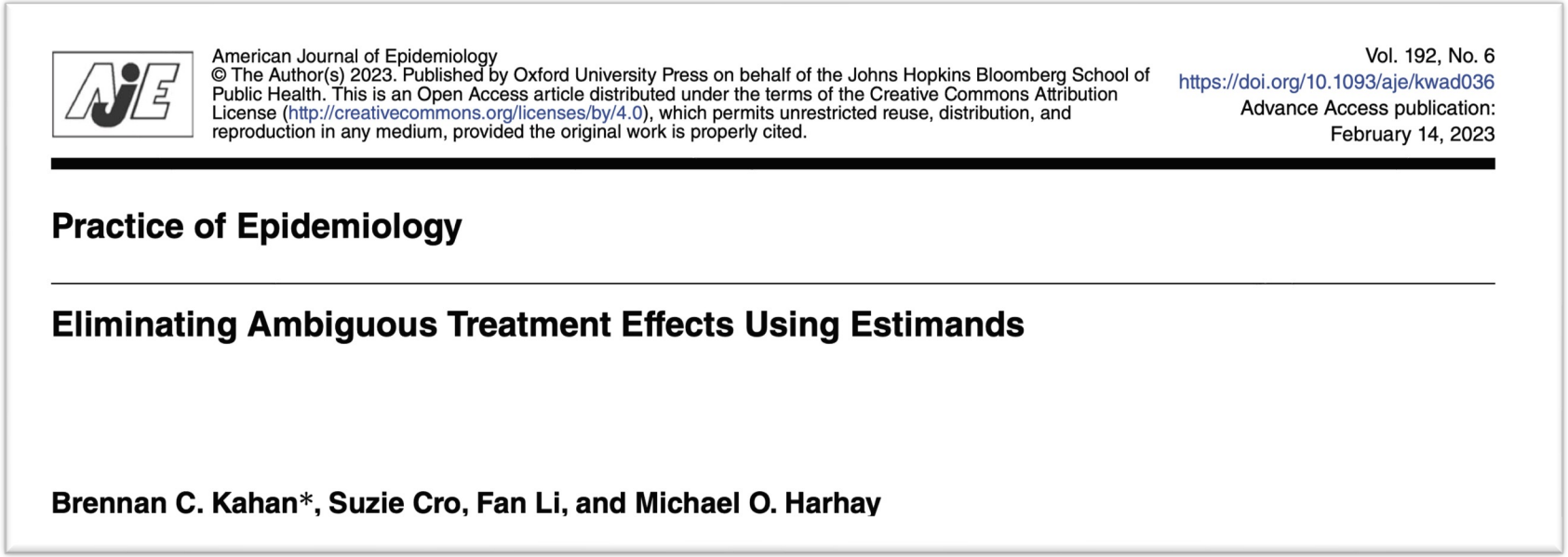
Hypothetical strategy

- This approach considers a hypothetical setting where the intercurrent event(s) *would not occur*.

- Here, that is the effect of an intervention in a hypothetical setting where participants don’t die, like the mixed-effect example we saw earlier.

The hypothetical strategy is quite common

- Cox model
- Inverse probability (of censoring) weighting (IPW)
- Competing risks
- Joint model



American Journal of Epidemiology

Vol. 192, No. 6
https://doi.org/10.1093/aje/kwad036
Advance Access publication:
February 14, 2023

Practice of Epidemiology

Eliminating Ambiguous Treatment Effects Using Estimands

Brennan C. Kahan*, Suzie Cro, Fan Li, and Michael O. Harhay

Kahan BC, Cro S, Li F, Harhay MO. Eliminating Ambiguous Treatment Effects Using Estimands. Am J Epidemiol. 2023 Jun 2;192(6):987-994.

The hypothetical strategy is quite common

Cox model

The Cox model estimates a hypothetical effect in this setting because it assumes:

1. that censored participants are still alive and at risk of the outcome (e.g., hospital discharge), and
2. there is no unmeasured confounding between the occurrence of the intercurrent event and outcomes.

The hypothetical strategy is quite common

Inverse probability (of censoring) weighting (IPW)

- Seeks to remove the selection bias that we noted can occur due to death by changing how much information trial participants contribute to an estimate of difference between arms.

- Specifically, it gives more weight to individuals with similar characteristics who have their outcomes other than death fully observed, seeking to provide an estimate of an interventions effect if mortality did not occur.

The hypothetical strategy is quite common

- Competing risks modeling

- Joint longitudinal and time-to-event model

Hypothetical strategy

- Need to consider the plausibility of the hypothetical scenario… (*alternative reality, say what?*)

- If participants discontinue due to toxicity – in what setting would they not discontinue?
  - Has the treatment been made less toxic?
  - Or are they taking treatment despite the toxicity?

- Death: in what setting would they not die?

- The questions this approach answers often are not possible in reality.

Strategies to address intercurrent events

Treatment policy

While on treatment/while alive

Hypothetical

Principal stratum

Composite

Principal stratum strategy

- Here, we are interested in the treatment effect in the principal strata in which the intercurrent event would not occur.

- That is, the effect of an intervention in a set of participants who would not discontinue treatment (irrespective of treatment arm).

Principal stratum strategy

- Modifies population aspect of estimand.
- Needs to use the same set of patients under both treatment conditions.
- Can define different principal strata sub-populations.
  - Patients who would not discontinue either treatment.
  - Patients who would not discontinue intervention.
- *These populations are often unknown in practice so statistical modelling is required.*

Principal stratification
Survivor average causal effect (SACE)

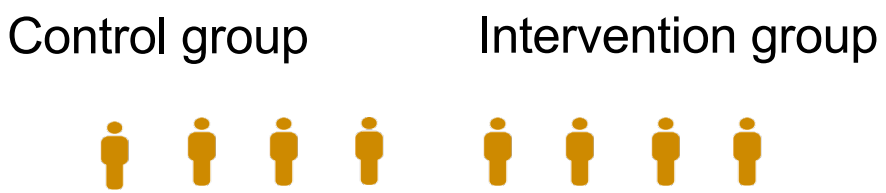
Table 1 | Patient groups based on potential survival experiences

	Survive to 12 months		Functional outcome at 12 months	
Patient stratum	**Intervention**	**Control**	**Intervention**	**Control**
Always survivors	Yes	Yes	A	C
Mortality benefiters	Yes	No	B	Undefined
Always diers	No	No	Undefined	Undefined
Specials	No	Yes	Undefined	D

RESEARCH METHODS AND REPORTING

Statistical methods to compare functional outcomes in randomized controlled trials with high mortality

Elizabeth Colantuoni,[1,2] Daniel O Scharfstein,[1,2] Chenguang Wang,[3] Mohamed D Hashem,[1,4] Andrew Leroux,[2] Dale M Needham,[1,4,5] Timothy D Girard[6]



Strategies to address intercurrent events

Treatment policy

While on treatment/while alive

Hypothetical

Principal stratum

Composite

Composite strategy

- The intercurrent event is incorporated into the endpoint definition (e.g., the endpoint is changed from “recurrence” to “recurrence or death”).

- Death is assigned a particular value of the endpoint.

- Different composite estimands could be defined based on the particular value assigned to the endpoint.

Commentary

January 20, 2010

Composite End Points in Randomized Trials
There Is No Free Lunch

George Tomlinson, PhD; Allan S. Detsky, MD, PhD

» Author Affiliations

JAMA. 2010;303(3):267-268. doi:10.1001/jama.2009.2017

Non-hierarchical versus hierarchical

- Non-hierarchical: death assigned a value possible among survivors.
  - Generation 1

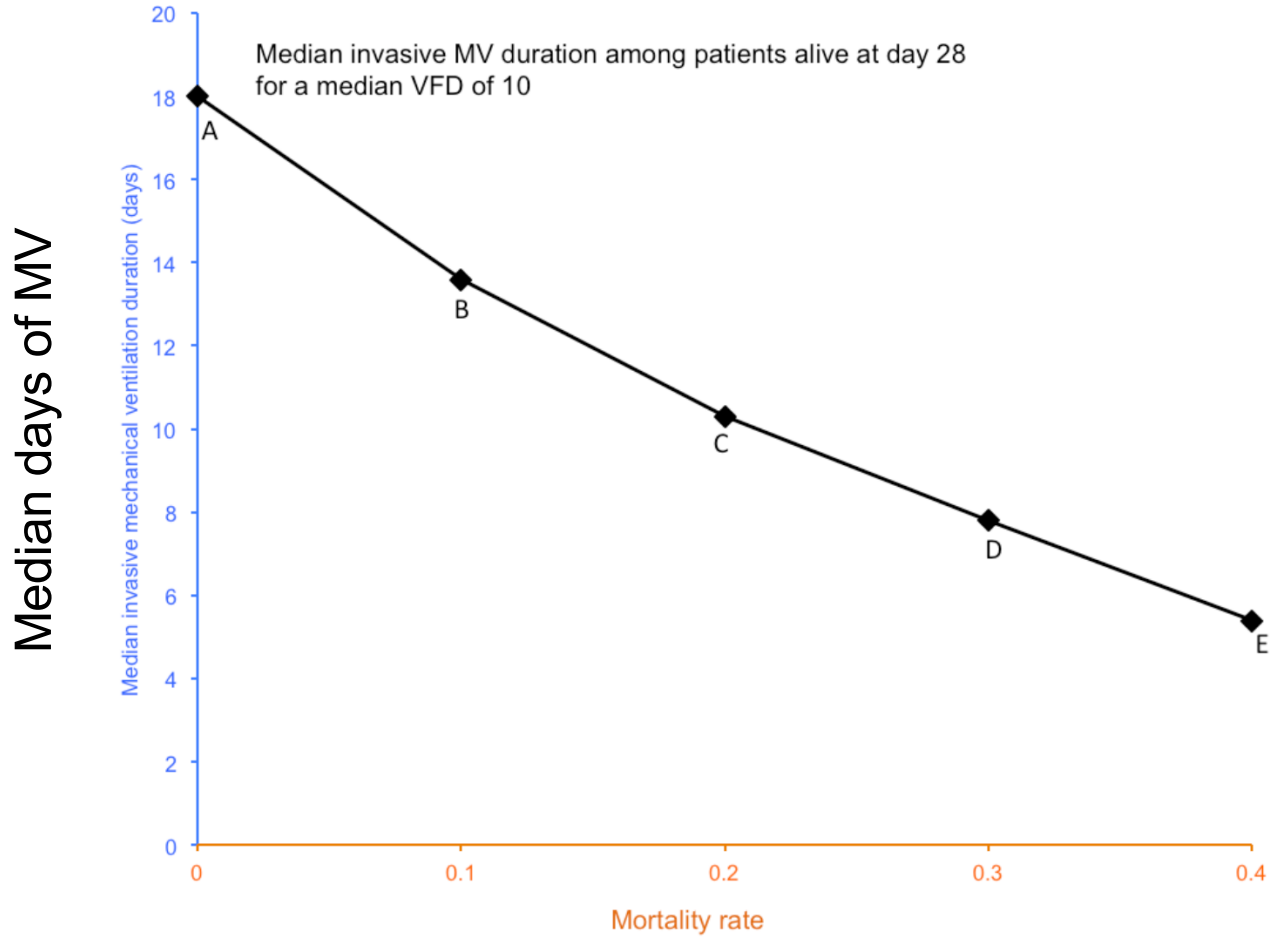
- Hierarchical: death assigned a value (or valued) different than the possible values among survivors.
  - Generation 2

Ventilator-Free Day Outcomes Can Be Misleading

The same number of VFDs in different trials does not equal:

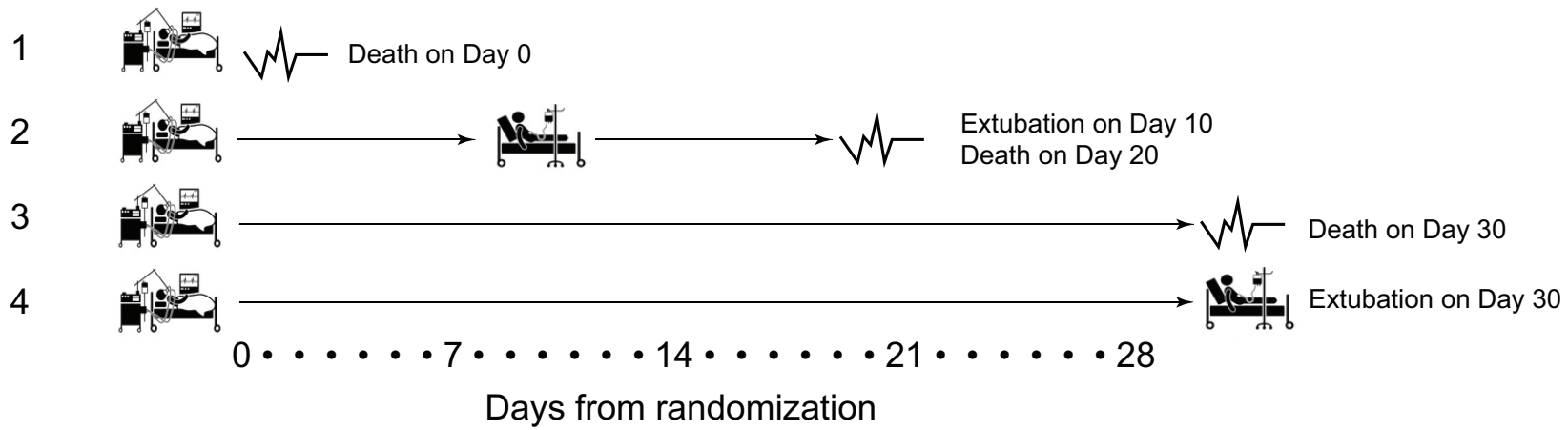
1. Same mortality,
2. Same healthcare utilization and public health impact, nor
3. Caregiver burden.

Many composite endpoints solve a statistical problem with the trade-off of a conceptual problem.

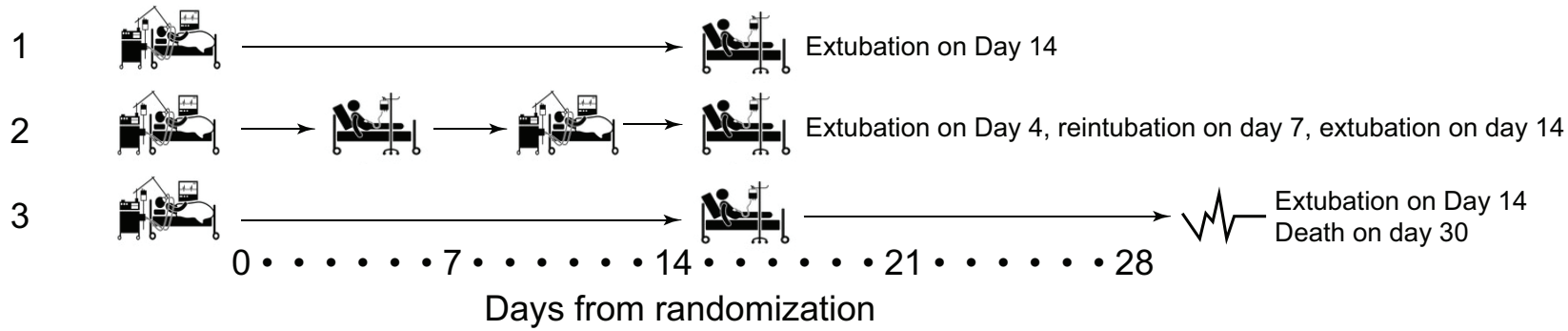


Bodet-Contentin et al, *Critical Care Medicine*, 2018

A) Ventilator free days = 0



B) Ventilator free days = 14

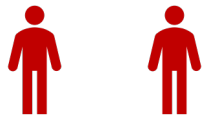


Yehya, Harhay et al. Re-appraisal of Ventilator-free Days in Critical Care Research. *AJRCCM,* 2019

Hierarchical composite endpoints

Win ratio: compares all trial participants in one arm to all trial participants in the other trial arm, and assigns each participant's outcome a win or not.

Arm A | Arm B

Neither survives hospital stay:  **= both get 0 points**

Both patients live:  **= first discharged in blue gets 1, other 0**

Only one survives:  **= surviving blue patient gets 1, other 0**

The trial arm with the most  (i.e., wins) across all possible comparisons is deemed to have responded better to the intervention.

Pulmonary arterial hypertension example

We proposed a hierarchy of five component events:

1. all-cause death
2. lung transplantation
3. PAH-related hospitalization
4. clinical improvement
5. worsening PAH

Hierarchical composite endpoints

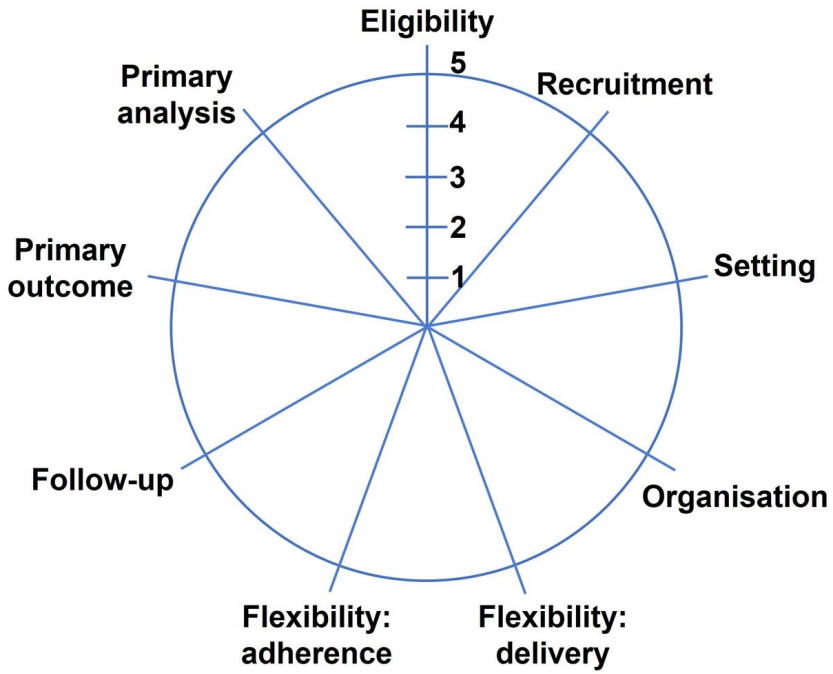
- I think this area has the most promise.

- Summary values still are not straightforward or meaningful as standalone.
  - Probability of favorable outcome

- But the net effect is captured, and outcomes can be prioritized and presented on several scales, i.e., absolute and relative differences.

Pragmatic considerations

- Ease of capture of key data elements.
- Clustering, particularly in cluster-randomized trials.



Intercurrent events versus missing data

- “There is an important distinction between an intercurrent event and missing data. Whether data are considered to be missing can depend on the choice of strategy for intercurrent events.”

- “For example, If data are unavailable for a particular participant following the use of rescue medication, this data would be missing for a treatment policy strategy but not relevant for a hypothetical strategy.”

Keene, O.N., Lynggaard, H., Englert, S. *et al.* Why estimands are needed to define treatment effects in clinical trials. *BMC Med* **21**, 276 (2023).

Strategies to address intercurrent events

Treatment policy

While on treatment/while alive

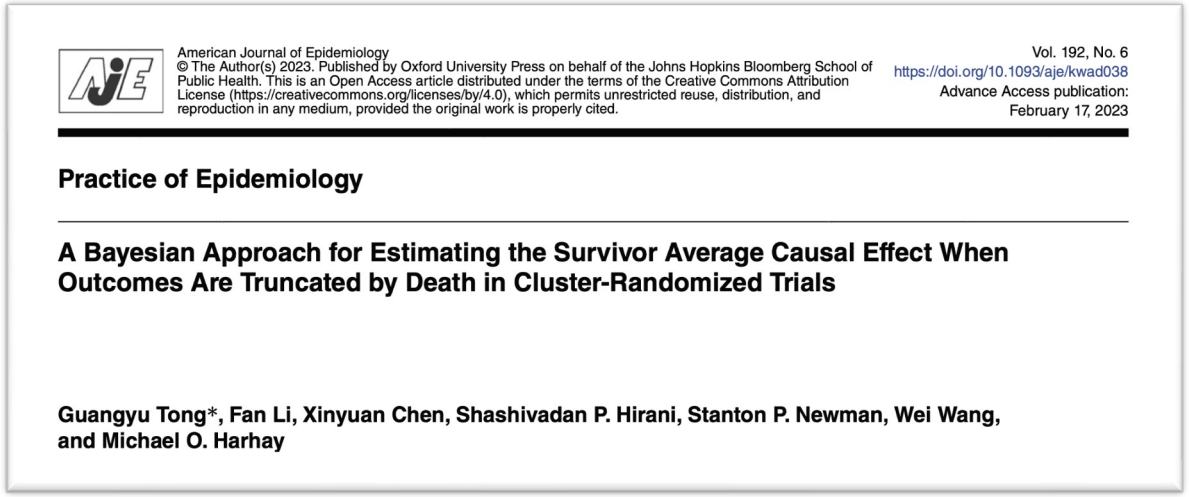
Hypothetical

Principal stratum

Composite

Survivor average causal effect (SACE)

- I like this as a planned secondary outcome approach.



American Journal of Epidemiology

Vol. 192, No. 6
https://doi.org/10.1093/aje/kwad038
Advance Access publication:
February 17, 2023

Practice of Epidemiology

A Bayesian Approach for Estimating the Survivor Average Causal Effect When Outcomes Are Truncated by Death in Cluster-Randomized Trials

Guangyu Tong*, Fan Li, Xinyuan Chen, Shashivadan P. Hirani, Stanton P. Newman, Wei Wang, and Michael O. Harhay

Received: 8 March 2022 | Revised: 5 September 2023 | Accepted: 12 October 2023
DOI: 10.1002/sim.9939

RESEARCH ARTICLE

Statistics in Medicine WILEY

A mixed model approach to estimate the survivor average causal effect in cluster-randomized trials

Wei Wang[1] | Guangyu Tong[2,3,4] | Shashivadan P. Hirani[5] | Stanton P. Newman[5,6] | Scott D. Halpern[1,7] | Dylan S. Small[8] | Fan Li[3,4] | Michael O. Harhay[1,7]

PCORI ME-2020C1-19220: https://github.com/harhay-lab

Tong G, Li F, Chen X, Hirani SP, Newman SP, Wang W, Harhay MO. A Bayesian approach for estimating the survivor average causal effect when outcomes are truncated by death in cluster-randomized trials. Am J Epidemiol. 2023 Jun 2;192(6):1006-1015.

CRT design considerations for using the SACE

- SACE may be most ideal as a pre-planned secondary analysis in trials with smaller available sample sizes due to uncertainty in the sample sizes of the always survivor strata.

- Only in larger pragmatic trials or in trials where effect sizes and always-survivors rates can be reasonably anticipated, SACE may be considered for the primary analysis.
  - Simulation studies can be undertaken to assess statistical power, but uncertainty remains.

- When there is interest in conducting primary analysis to estimate SACE, approaches for sample size re-estimation with pre-planned interim analysis may be considered.

Strategies to address intercurrent events

Treatment policy

While on treatment/while alive

Hypothetical

Principal stratum

Composite

Days alive and out of a hospital/institution

- DAOH (HFDs) for each patient:
  - Subtract the number of days since death, or spent in hospital, from a follow-up time.
  - The percentage of DAOH (%DAOH) can be calculated by dividing DAOH by the potential total follow-up time.

- Effect estimate on the absolute or relative scale.
  - Win ratio


PULMONARY PERSPECTIVE

Hospital-Free Days: A Pragmatic and Patient-centered Outcome for Trials among Critically and Seriously Ill Patients

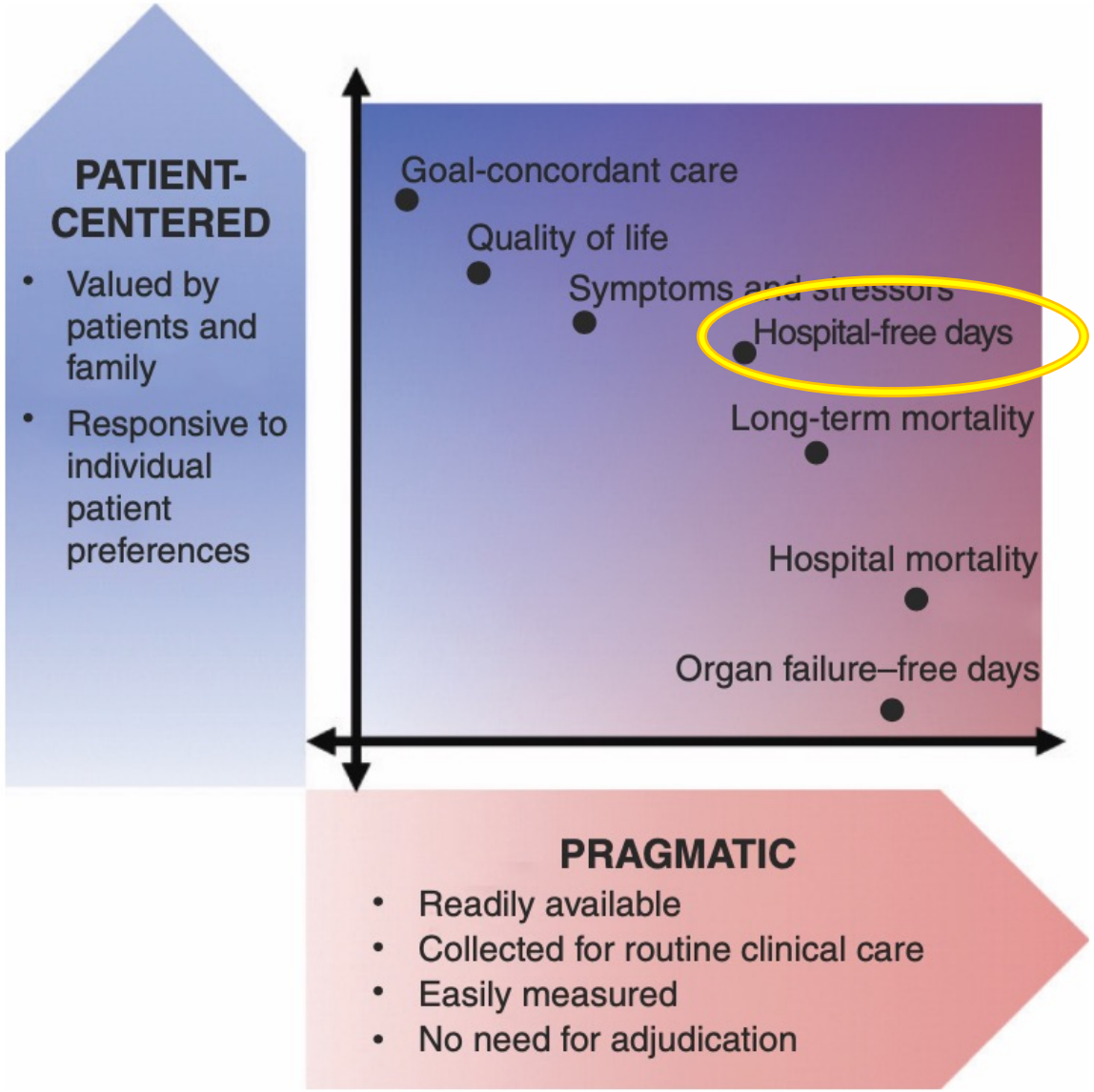
Catherine L. Auriemma[1,2,3], Stephanie P. Taylor[4], Michael O. Harhay[1,5,6], Katherine R. Courtright[1,2,3], and Scott D. Halpern[1,2,3,6,7]

NIA IMPACT COLLABORATORY
TRANSFORMING DEMENTIA CARE

Auriemma CL, Taylor SP, Harhay MO, Courtright KR, Halpern SD. Hospital-Free Days: A Pragmatic and Patient-centered Outcome for Trials among Critically and Seriously Ill Patients. Am J Respir Crit Care Med. 2021 Oct 15;204(8):902-909.

Hospital-Free Days

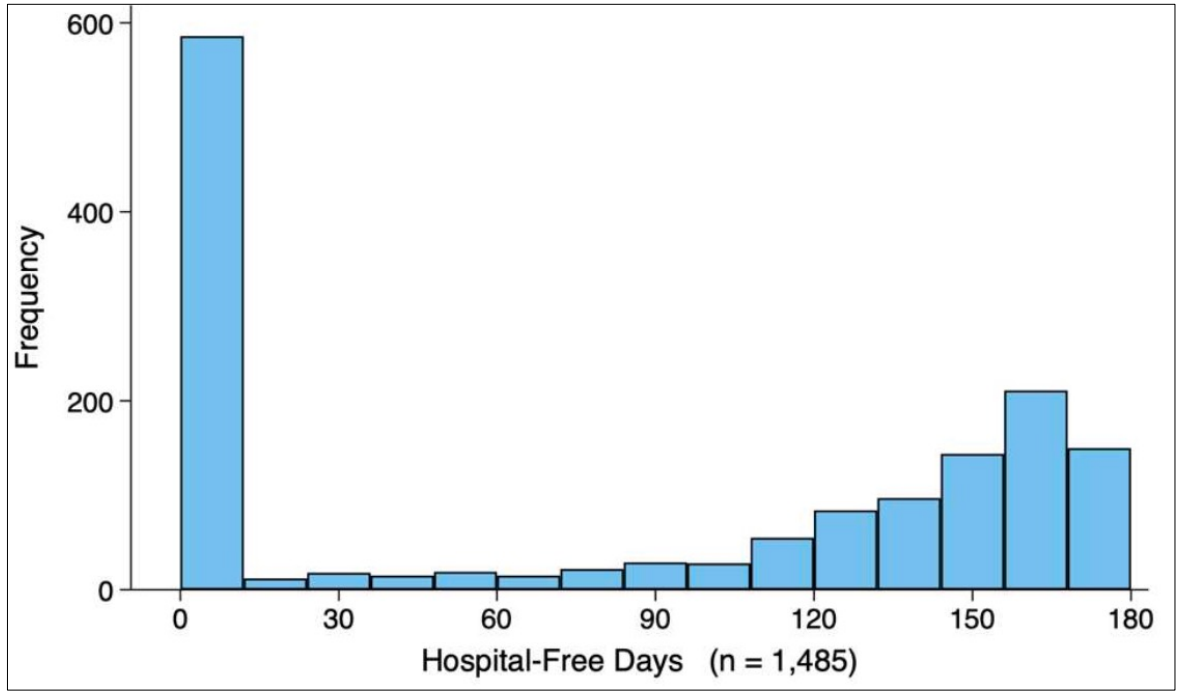
- HFDs are **patient-centered**
  - longer life >> shorter life
  - outside hospital >> inside
- Measuring HFDs is **pragmatic**
  - data available at low cost
  - data without substantial missingness
- Allows for **direct comparison** of interventions with either restorative or palliative intents



Auriemma CL, Taylor SP, Harhay MO, Courtright KR, Halpern SD. Hospital-Free Days: A Pragmatic and Patient-centered Outcome for Trials among Critically and Seriously Ill Patients. Am J Respir Crit Care Med. 2021 Oct 15;204(8):902-909.

Days alive and out of a hospital/institution

- Estimand? Analysis?
  - An outcome (death of days free) in the overall measure will be reflected in the overall difference in means.
    - Easy to fit mixed-effects and GEE models.



Brennan Kahan and M. Harhay, In progress

Summary

- We need to be careful with translating trial results into practice.
  - Not to say anyone intends to be misleading. This is hard stuff.

- “Different treatment effects can be considered depending on how intercurrent events are included in the estimand definition and therefore there is no single “true” treatment effect.” (Keene et al, 2023)

- We can use different strategies for different intercurrent events.
  - None are optimal in my view.

Summary

- But I think some are clearly more pragmatic than others, and currently, I'm most often a proponent of hierarchical composite outcomes as I think knowing the net benefit is perhaps the best of the options I have seen.

- Hierarchical composite outcomes can also be adapted to each study's pragmatic dimensions.

Final thoughts

- There are other approaches out there that I did not cover today.
- There are several disease-specific adaptations to a lot of these. I think that is great.
- Big role for stakeholder (broadly speaking) engagement that remains untapped.
- My thinking and views continue to evolve on this topic.
- Happy to answer questions! mharhay@pennmedicine.upenn.edu

Further reading (open access)

- Lawrance R, Degtyarev E, Griffiths P, Trask P, Lau H, D'Alessio D, Griebsch I, Wallenstein G, Cocks K, Rufibach K. What is an estimand & how does it relate to quantifying the effect of treatment on patient-reported quality of life outcomes in clinical trials? J Patient Rep Outcomes. 2020 Aug 24;4(1):68.

- Clark TP, Kahan BC, Phillips A, White I, Carpenter JR. Estimands: bringing clarity and focus to research questions in clinical trials. BMJ Open. 2022 Jan 3;12(1):e052953.

- Colantuoni E, Scharfstein DO, Wang C, Hashem MD, Leroux A, Needham DM, Girard TD. Statistical methods to compare functional outcomes in randomized controlled trials with high mortality. BMJ. 2018 Jan 3;360:j5748.

Without implicating

- Scott Halpern
- Fan Li
- Brennan Kahan
- Kate Courtright
- Nadir Yehya
- Anna Heath
- Elizabeth Colantuoni
- Ewan Goligher
- Katie Auriemma

NIA IMPACT COLLABORATORY
TRANSFORMING DEMENTIA CARE

Questions?