

OPINION

Open Access



# Why estimands are needed to define treatment effects in clinical trials

Oliver N. Keene<sup>1\*</sup> , Helle Lynggaard<sup>2</sup>, Stefan Englert<sup>3</sup>, Vivian Lanius<sup>4</sup> and David Wright<sup>5</sup>

## Abstract

**Background** The estimand for a clinical trial is a precise definition of the treatment effect to be estimated. Traditionally, estimates of treatment effects are based on either an ITT analysis or a per-protocol analysis. However, there are important clinical questions which are not addressed by either of these analyses. For example, consider a trial where patients take a rescue medication. The ITT analysis includes data after use of rescue, while the per-protocol analysis excludes these patients altogether. Neither of these analyses addresses the important question of what the treatment effect would have been if patients did not take rescue medication.

**Main text** Trial estimands provide a broader perspective compared to the limitations of ITT and per-protocol analysis. Trial treatment effects depend on how events occurring after treatment initiation such as use of alternative medication or discontinuation of the intervention are included in the definition. These events can be accounted for in different ways, depending on the clinical question of interest.

**Conclusion** The estimand framework is an important step forward in improving the clarity and transparency of clinical trials. The centrality of estimands to clinical trials is currently not reflected in methods recommended by the Cochrane group or the CONSORT statement, the current standard for reporting clinical trials in medical journals. We encourage revisions to these guidelines.

**Keywords** Estimand, Treatment effect, Intercurrent event, ITT, Per-protocol, PICO, CONSORT

## Background

The CONSORT statement is used worldwide for the reporting of randomised controlled trials [1, 2]. While CONSORT requires specification of the outcome measure, this does not provide a precise definition of the treatment effect. Estimation of the treatment effect is typically complicated by events that occur after initiation

of treatment such as use of alternative medications or discontinuation of the assigned treatment. Such events have been termed intercurrent events, defined as “events occurring after treatment initiation that affect either the interpretation or the existence of the measurements associated with the clinical question of interest” [3].

In their recent paper, Little and Lewis [4] use a different definition of a trial estimand, referring to this simply as the “true effect of the intervention”. However, there are alternative ways in which the treatment effect can be defined as it depends on how events such as use of alternative medication or discontinuation of the intervention are included in the definition and therefore there is no single “true” treatment effect.

For example, PIONEER-1 [5] compared the effects of semaglutide with placebo on glycaemic control in patients with type 2 diabetes mellitus. Rescue medication

\*Correspondence:

Oliver N. Keene  
oliverkeene405@gmail.com

<sup>1</sup> KeeneONStatistics, Maidenhead, UK

<sup>2</sup> Biostatistics, Novo Nordisk A/S, Bagsvaerd, Denmark

<sup>3</sup> Statistical Modeling & Methodology, Janssen R&D, Janssen-Cilag GmbH, Neuss, Germany

<sup>4</sup> Statistics & Data Insights, Bayer AG, Wuppertal, Germany

<sup>5</sup> Statistical Innovation, Data Science & Artificial Intelligence, Biopharmaceuticals R&D, AstraZeneca, Cambridge, UK



© The Author(s) 2023. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

was recommended for persistent and unacceptable hyperglycaemia, and it was expected that more placebo patients would need rescue medication compared to semaglutide. Is the true effect of treatment the effect of semaglutide including the potential effects of differential use of rescue medication or is it the effect of semaglutide without the use of rescue medication? Or is it another effect? Similarly, participants could discontinue randomised treatment. Is the true effect of treatment the effect of the decision to initiate treatment with semaglutide or is it the effect of semaglutide had all patients completed their prescribed course of treatment? Unless specified, we cannot understand and interpret the estimated treatment effect.

**ICH E9 (R1) definition of estimands**

Pharmaceutical industry trials are governed by scientific guidelines produced by ICH (International Council for Harmonisation of technical requirements for pharmaceuticals for human use). Difficulties in expressing clearly the treatment effect to be estimated in clinical trials led to a new addendum to the ICH E9 guideline on statistical principles for clinical trials [3].

The ICH definition of an estimand requires a clear overview of intercurrent events and the associated strategy (see Table 1) chosen to reflect the clinical question of interest for each event. In addition, the estimand includes a complete description of the following attributes:

- *Treatment condition*, and as appropriate, the alternative treatment condition
- *Population* of patients targeted (e.g. patients with type 2 diabetes)

- *Variable* (or endpoint) obtained for each patient to address the clinical question (e.g. change from baseline to week 26 in HbA1C)
- *Population-level summary* for the variable, providing a basis for comparison between treatment conditions (e.g. difference in means in change from baseline to week 26 in HbA1C)

Strategies for intercurrent events can be incorporated in the treatment, population or variable attributes or can be specified separately. For example, “use of rescue medication as required” could be included as part of the treatment condition.

The PICO (population, intervention, control, and outcomes) format has often been used for framing clinical research questions [6]. Compared to PICO, the estimands framework in addition addresses the key issues of intercurrent events and summary measure, and without these elements, the treatment effect is not adequately defined.

In the PIONEER-1 example, employing a treatment policy strategy for rescue medication includes all data after the intercurrent event and estimates a treatment effect regardless of use of rescue medication, i.e. the treatment effect includes the impact of rescue medication on glycaemic control. A hypothetical strategy estimates the effect of semaglutide in the absence of rescue medication [7].

A composite strategy can be appropriate where the intercurrent event represents a poor (or positive) outcome of treatment. For example, the SYNAPSE trial compared mepolizumab and placebo in chronic rhinosinusitis with nasal polyps [8]. The co-primary endpoints were nasal polyps score and nasal blockage score at the end of the trial. A key intercurrent event was surgery for nasal polyps, a bad outcome for the patient that would be expected to subsequently improve nasal scores. A treatment policy strategy would include nasal scores after surgery in the comparison, but this would not reflect the negative outcome of surgery that was undertaken after treatment initiation. Instead, a composite strategy was used incorporating surgery as an adverse outcome in the endpoint definition.

**ITT and per-protocol analysis**

Historically analyses have been viewed as a dichotomy: they are either ITT analyses or “per-protocol” [9]. For example, the current Cochrane Risk of Bias tool [10] defines in Sect. 1.3 only two possible treatment effects of interest:

- (1) The effect of assignment to the interventions at baseline (regardless of whether the interventions

**Table 1** Potential strategies for intercurrent events

Strategy	Relevant values for patients with the intercurrent event
Treatment policy	Actual values of the variable regardless of whether the intercurrent event has occurred
Composite	Modified definition of the variable incorporating the intercurrent event
Hypothetical	Values the variable would have taken in the hypothetical scenario envisaged in which the intercurrent event would not occur
While on-treatment/at-risk	Values of the variable up to the time of the intercurrent event
Principal stratum	Restrict population to patients who would not experience the intercurrent event

are received during follow-up, sometimes known as the ‘intention-to-treat effect’); or.

- (2) The effect of adhering to intervention as specified in the trial protocol (sometimes known as the ‘per-protocol effect’).

When defined as above, the intention-to-treat effect corresponds to a treatment policy strategy for all intercurrent events. ITT analysis is often interpreted as referring only to including all randomised patients in the population analysed but actually requires complete follow-up of all randomised patients to the end of the planned study period [11]. When using this strategy, it is important to ensure data are still collected after the intercurrent event as these data will be included in the analysis.

It is increasingly recognised that the treatment effect estimated by the treatment policy approach may not always be of primary clinical interest and may not appropriately communicate to prescribers and patients the efficacy that is directly attributable to the treatment, i.e. what can be expected in terms of efficacy if the patient takes the medication as prescribed [12–14].

The distinguishing feature of a per-protocol analysis is that patients with major protocol violations are excluded altogether and all of their data including data collected prior to the violation is not used. The decision on whether to exclude a patient based on their adherence to the protocol can be somewhat arbitrary. When a per-protocol analysis excludes patients who discontinued due to lack of efficacy, this information on poor efficacy is lost from the analysis. The clinical question that is being addressed by a per-protocol analysis is somewhat unclear [3]. By requiring that intercurrent events are defined along with their associated strategy, the estimand framework allows the corresponding clinical questions to be precisely defined. The terms ITT and per-protocol do not accurately describe the estimand.

The PIONEER-1 trial reported an estimand for the glycated haemoglobin endpoint that used a hypothetical strategy for discontinuation of trial medication. The analysis used all data until trial medication was discontinued and predicted data that would have been observed if treatment had not been discontinued. All patients were included in the analysis, so the estimated benefit is not the effect in the subset of patients who took the medication as intended, a known bias of previous per-protocol estimates of treatment effects [15].

The estimand framework distinguishes between the target of estimation (trial estimand) and the method of estimation (estimator). In some cases, the estimand is difficult to estimate reliably or a more sophisticated analysis is required. An explicit definition of the estimand allows

a more transparent assessment of whether the estimation method appropriately addresses the clinical question of interest.

#### Intercurrent events and 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, in the PIONEER-1 study a key intercurrent event was use of rescue medication. If data are unavailable for a particular participant following use of rescue medication, this data would be missing for a treatment policy strategy but not relevant for a hypothetical strategy (see Table 1). In contrast, if data are available, these data are relevant for a treatment policy strategy, but they are not relevant for a hypothetical strategy. For the hypothetical strategy, the relevant data will need to be predicted under the hypothetical scenario envisaged, typically by multiple imputation.

#### Conclusions

The estimand framework provides an important approach to defining the treatment effect to be estimated in a clinical trial. 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.

Trials should be designed with a clearly articulated clinical question of interest. It is necessary to address intercurrent events when describing the clinical question of interest in order to precisely define the treatment effect that is to be estimated. Often, there will be multiple questions of interest, and these will lead to different estimands which result in different estimates of benefit.

A description of estimands should be included in trial publications to provide clarity on the treatment effect reported. An overview of the frequency and timing of each type of intercurrent event by treatment group is needed for proper interpretation of the estimated treatment effect. The estimand used by a trial is an important feature for meta-analyses as currently such analyses may combine estimates from different estimands when reporting treatment effects. We encourage revisions to methods recommended by the Cochrane group and to the CONSORT statement to reflect the centrality of estimands to clinical trials.

#### Abbreviations

ICH	International Council for Harmonisation of technical requirements for pharmaceuticals for human use
PICO	Population, intervention, control, and outcomes

**Acknowledgements**

Not applicable

**Authors' contributions**

ONK, HL, SE, VL and DW made substantial contributions to the conception and design of the article and drafted the work or substantively revised it. All authors read and approved the final manuscript.

**Funding**

Not applicable.

**Availability of data and materials**

Data sharing is not applicable to this article as no datasets were generated or analysed during the current study.

**Declarations****Ethics approval and consent to participate**

Not applicable.

**Consent for publication**

Not applicable.

**Competing interests**

ONK is a former employee of GlaxoSmithKline and hold shares in the company. HL is a full-time employee of Novo Nordisk and holds shares in the company. SE is a full-time employee of Janssen. VL is a full-time employee of Bayer. DW is a full-time employee of AstraZeneca.

Received: 3 April 2023 Accepted: 3 July 2023

Published online: 27 July 2023

**References**

- Schulz KF, Altman DG, Moher D. CONSORT 2010 statement: updated guidelines for reporting parallel group randomized trials. *Ann Intern Med.* 2010;152:726–32.
- Altman DG, Schulz KF, Moher D, Egger M, Davidoff F, Elbourne D, Gøtzsche PC, Lang T. The revised CONSORT statement for reporting randomized trials: explanation and elaboration. *Ann Intern Med.* 2001;134:663–94.
- International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use. ICH E9 (R1) addendum on estimands and sensitivity analysis in clinical trials to the guideline on statistical principles for clinical trials, final version, adopted Nov 2019. [https://database.ich.org/sites/default/files/E9-R1\\_Step4\\_Guideline\\_2019\\_1203.pdf](https://database.ich.org/sites/default/files/E9-R1_Step4_Guideline_2019_1203.pdf). Accessed 7 Nov 2022.
- Little RJ, Lewis RJ. Estimands, estimators, and estimates. *JAMA.* 2021;326:967–8.
- Aroda VR, Rosenstock J, Terauchi Y, Altuntas Y, Lalic NM, Villegas EC, Jeppesen OK, Christiansen E, Hertz CL, Haluzik M. PIONEER 1: randomized clinical trial of the efficacy and safety of oral semaglutide monotherapy in comparison with placebo in patients with type 2 diabetes. *Diabetes Care.* 2019;42:1724–32.
- Richardson WS, Wilson MC, Nishikawa J, Hayward RS. The well-built clinical question: a key to evidence-based decisions. *ACP J Club.* 1995;123:A12–3.
- European Medicines Agency Draft Guideline on Clinical Investigation of Medicinal Products in the Treatment or Prevention of Diabetes Mellitus. [https://www.ema.europa.eu/en/documents/scientific-guideline/draft-guideline-clinical-investigation-medicinal-products-treatment-prevention-diabetes-mellitus\\_en.pdf](https://www.ema.europa.eu/en/documents/scientific-guideline/draft-guideline-clinical-investigation-medicinal-products-treatment-prevention-diabetes-mellitus_en.pdf). Accessed 7 Nov 2022.
- Han JK, Bachert C, Fokkens W, Desrosiers M, Wagenmann M, Lee SE, Smith SG, Martin N, Mayer B, Yancey SW, Sousa AR. Mepolizumab for chronic rhinosinusitis with nasal polyps (SYNAPSE): a randomised, double-blind, placebo-controlled, phase 3 trial. *Lancet Respir Med.* 2021;9:1141–53.
- Keene ON. Intent-to-treat analysis in the presence of off-treatment or missing data. *Pharm Stat.* 2011;10:191–5.
- Sterne JA, Savović J, Page MJ, Elbers RG, Blencowe NS, Boutron I, Cates CJ, Cheng HY, Corbett MS, Eldridge SM, Emberson JR. RoB 2: a revised tool for assessing risk of bias in randomised trials. *BMJ.* 2019;28:366.
- International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use. Statistical principles for clinical trials, Final version, adopted Feb 1998. [https://database.ich.org/sites/default/files/E9\\_Guideline.pdf](https://database.ich.org/sites/default/files/E9_Guideline.pdf). Accessed 7 Nov 2022.
- Keene ON, Wright D, Phillips A, Wright M. Why ITT analysis is not always the answer for estimating treatment effects in clinical trials. *Contemp Clin Trials.* 2021;108: 106494.
- Darken P, Nyberg J, Ballal S, Wright D. The attributable estimand: a new approach to account for intercurrent events. *Pharm Stat.* 2020;19:626–35.
- Akacha M, Bretz F, Ruberg S. Estimands in clinical trials – broadening the perspective. *Statist Med.* 2017;36:5–19.
- Young JG, Vatsa R, Murray EJ, Hernán MA. Interval-cohort designs and bias in the estimation of per-protocol effects: a simulation study. *Trials.* 2019;20(1):1–9.

**Publisher's Note**

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more [biomedcentral.com/submissions](https://biomedcentral.com/submissions)

