European Journal of Clinical Nutrition (2018) 72:1291–1295
https://doi.org/10.1038/s41430-018-0207-x

PERSPECTIVE

Estimands: improving inference in randomized controlled trials in clinical nutrition in the presence of missing values

Christian Ritz[1] · Birgitte Rønn[2]

Received: 16 April 2018 / Accepted: 22 April 2018

Abstract
For randomized controlled trials, the impact of the amount and handling of missing data on the interpretation of the treatment effect has been unclear. The current use of intention to treat, per protocol, and complete-case analysis has shortcomings. The use of estimands may lead to improved estimation of treatment effects through more precise characterizations of the fate of treatments after dropout or other post-randomization events. A perspective on current and future developments with a view toward clinical nutrition is provided.

Basics in clinical nutrition

Over the last decade there have been a number of initiatives toward improving inference in randomized controlled trials (RCTs) with respect to missing data. The impact of the amount and handling of missing data on the interpretation of the treatment effect has been unclear. These initiatives were mainly driven by regulatory requirements for medical research in the pharmaceutical industry. Recently it has also become a very active area of research.

The concepts, definition, and methods are equally applicable to RCTs conducted in clinical nutrition, and they are useful for improving how data should be analyzed and results should be interpreted, improving our understanding of the findings from RCTs.

What is a treatment effect?

This seems like a straightforward question: What is the effect of treatment? In a clinical trial protocol, this component is rarely subject to any discussion or argumentation. A relevant outcome that will serve as measure of the effect is simply presented and sometimes also the time point at which the evaluation of the outcome should take place. But what is the treatment that we want to evaluate? Is it when taken exactly as described in the protocol or are the questions of interest rather what the treatment effect is when taken/administered as the trial population would actually do in practice?

In drug development, the treatment effect is how the outcome of treatment compares to what would have happened to the same subjects under different treatment conditions, e.g., had they not received the treatment or had they received a different treatment [1]. This is a causal treatment effect as the unit and time at which the treatment is compared with no/another treatment is the same. This is also a counterfactual treatment effect, since the same subject cannot both receive and being denied a treatment at the same time. This obstacle has traditionally been dealt with by randomization.

Randomization and blinding

The causal treatment effect is often evaluated using an RCT where the study population is randomly divided into two (or more) groups before initiation of any investigational treatment. At baseline, the two groups will both be random samples of the same study population and it can be expected that they on average (if the experiment is repeated) will have the same effect if both received treatment. If the trial is also blinded, the subjects in the control group can represent the subjects in the treatment group when being denied the treatment (a counterfactual situation). The fact that we have causality by design is the well-known strength of the

✉ Christian Ritz
ritz@nexs.ku.dk

[1] Department of Nutrition, Exercise and Sports, Faculty of Science, University of Copenhagen, Copenhagen, Denmark

[2] Biometrics, Leo Pharma a/s, Copenhagen, Denmark

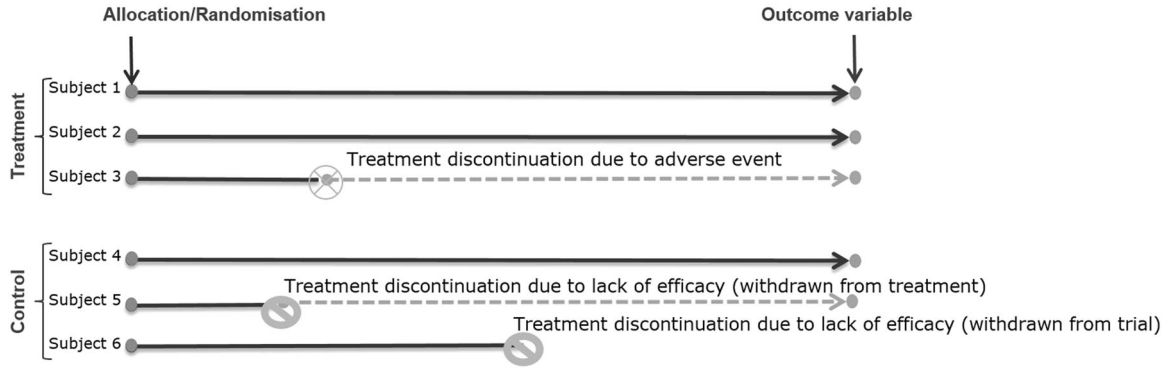

Fig. 1 Schematic display of types of dropouts that can occur in the control and treatment arms

parallel, double-blinded randomized clinical trials and with measurements at baseline and at the time point of the measure of effect for all randomized subjects (the ideal RCT scenario) the counterfactual treatment effect can be evaluated by comparing outcomes from the two groups.

Departures from the ideal RCT scenario

Post-randomization events may lead to various departures from the ideal RCT scenario. We discuss two important and frequently occurring departures: missing (outcome) data, compliance/non-adherence.

Missing data

An important deviation from the ideal RCT that occur in real life is missing data: many randomized controlled trials (on humans) conducted in nutrition but also medical research are facing missing outcome values, essentially reflecting the free will of study participants, that may result in some degree of attrition. Missing values may result from dropouts, which are subjects that withdraw from the trial before it is terminated, e.g., due to lack of efficacy or adverse events or for reasons unknown to the study investigators. There are also intermittent missing values due to subjects not showing up for all scheduled visits, but still participating until study termination, but we will not spend time on these as they have less impact in the conclusion concerning efficacy to be made at a specific time point (most likely causing a slight loss in efficiency). Despite of missing observations, the causal treatment effect can still be estimated if events that make subjects drop out are independent of the trajectory of the outcome values after the event took place. This can rarely be assumed to be the case, e.g., often more drop-outs due to lack of efficacy would be expected in a control arm treated with placebo than in the treatment arm (see Fig. 1). For subjects dropping out due to adverse events the opposite pattern can be anticipated in many settings. Therefore, the primary outcome in real life often is available for groups of subjects that are not random samples of the study population at baseline and the methods chosen to guess or impute the impact of the missing data become important.

Compliance and non-adherence

Another deviation from the ideal RCT scenario happens if some subjects during the course of the trial only partly adhere to the treatment. Such non-missing data may also jeopardize the causality by design. Lack of compliance or non-adherence due to adverse events, change in treatment, or other kinds of events that violates the intention of the protocol, occurring after randomization, may also influence the outcome used for evaluation for efficacy. For example measurements of plasma glucose taken one week after a subject stopped adhering to a planned diet will no longer reflect the effect of that diet, but rather whatever diet the subject had chosen to follow instead during the last week. Therefore, even if the plasma glucose is measured at the planned time, the evaluation will not be relevant to find the causal effect of the planned diet. However, it may be relevant to evaluate a treatment effect that is defined as the effect of "being allocated to" the planned diet.

Present research activities

As pointed out earlier a treatment effect is not just a treatment effect. If there are intercurrent, post-randomization events leading to departures from the ideal RCT scenario the treatment effect may be estimated in several ways depending on the assumptions made.

Definition of estimands

Estimands formalize what is being estimated by providing precise characterizations of the study populations to be used for estimating treatment effects [2]. Specifically, defining an

estimand may require to specify the study outcome and the time point or time period of interest, the target population, handling of post-randomization events (how subjects are assumed to behave after withdrawal; fate of treatments after discontinuation/dropout), and the model parameter corresponding to the treatment effect of interest. It is important to realize that the definition of an estimand does not imply that a certain statistical model has to be used. However, estimands imply certain model restrictions.

In principle, the choice of estimand will impact the study design and execution and hence the estimand should ideally be defined simultaneously with the formulation of the scientific question of interest [3].

There are estimands for estimating efficacy, which may be defined as the effect of a diet if taken, ideally as specified in the protocol although adherence may be difficult to ensure in reality, and for estimating effectiveness, which may be viewed as the effect of a diet when assigned [2]. However, efficacy and effectiveness may be understood in several ways [4]. Therefore, the preferred terminology is instead *de jure* and *de facto* estimands: de jure estimands estimate the treatment effect as it would be if control and treatment were taken as specified in the protocol. On the other side, de facto estimands estimate the treatment effect as it would be in practice based on the assignment of control and treatment to subjects [5].

Here are three examples of strategies for estimands [1]:

Hypothetical

The hypothetical strategy aims to estimate the treatment effect if no post-randomization events occur. The strategy leads to de jure estimands as it makes a comparison of the two randomized group of subjects, so it is not the same as a per protocol analysis or a complete-case analysis as have been common practice previously. The strategy does not imply how the corresponding estimands should be estimated: the statistical model and the assumptions behind it are not given as a part of the definition of the estimands. The estimands are hypothetical since clinical trials without post randomization events are hypothetical and the estimated treatment effect is counterfactual since some subjects will drop out due to adverse events or lack of adherence with treatment.

Treatment policy

The treatment policy strategy aims to estimate the treatment effect regardless of the post-randomization events that occur. The strategy leads to de facto estimands as it addresses the effect of the treatment as it has been manifested in the current trial. In practice, the effect is estimated based on observations collected regardless of whether the subjects have stopped treatment due to adverse events or change to other treatments. With this estimation strategy, it will be required to try to keep the subjects in the trial no matter how the subjects adhere/do not adhere to the treatment. The strategy leads to estimands that estimate the effect of being randomized to the control/treatment in the trial—but it may also be heavily influenced by the design of the trial and it may in the end not mimic the treatment policy as it will be in real life.

While on treatment

The while on treatment strategy aims at estimating the effect of treatment until the post-randomization events occur. This strategy also leads to de jure estimands, but instead of evaluating the effect at a specific time point it considers the time frame until the treatment stopped. In practice this strategy can correspond to last observation carried forward and, in trials evaluating weight loss, this may result in low placebo response, since placebo patients may withdraw early from the trial, due to lack of efficacy, and hence contribute to the estimated effect by short observation time and very low weight losses.

Other examples and more detailed explanations are provided in refs. [6, 7]. Estimands for area under the curve and for studies with run-in periods could also be relevant to consider. There is a recent application of estimands in a clinical nutrition context [8].

Intention to treat and per protocol analysis

Intention to treat (ITT) has been interpreted to imply quite different analyses [9]. In the past, it was often taken to mean that an available-case analysis was carried out [10, 11].

We will assume that the intention-to-treat principle means that statistical analysis is based on the dataset encompassing all randomized subjects and respecting their allocation to control or treatment groups [12]. Both de facto and de jure estimands may be defined while respecting the intention-to-treat principle. We refer to these estimands as intention-to-treat estimands. In fact, it has been argued that meaningful estimands have to be ITT estimands [9, 12]. An available-case analysis will only correspond to an intention-to-treat estimand under strong and usually rather unrealistic assumptions.

It is common to supplement the main analysis based on the intention-to-treat principle with one or more so-called per protocol analyses, which are based on study populations that satisfy certain criteria defining a high degree of compliance or adherence, possibly defined using additional collected data on self-reported intake or measured intake (tracers in urine samples, amount of leftovers). However, such analyses lacks a sound interpretation as they are based

on comparing two non-randomized groups. A complete-case analysis is one example of a per protocol analysis. These analyses will usually be based on fewer subjects than is the case for analyses derived from intention-to-treat estimands, leading to a risk of loss of information [2, 13]. Moreover, there is a risk of selection bias where selections may differ between groups, depending on the missing data mechanism.

Sensitivity analysis

Typically, a single estimand is chosen for the main statistical analysis of data on the primary outcome. It may, however, be necessary to supplement this analysis by means of additional analyses (on the same outcome) as to provide a more nuanced picture of the treatment effect of interest. We refer to such analyses as sensitivity analyses.

There are two types of sensitivity analyses [3]. Internal validity is investigated by fitting different statistical models (e.g., different correlation structures for longitudinal data, different covariate adjustments, inclusion of more or less intermittent repeated measurements) for the same estimand. This type of sensitivity analysis explores the robustness of the findings of the main analysis (under different model assumptions) while retaining the exact same study population and way of handling missing data. External validity is investigated by fitting similar models for different estimands, and it provides a means for exploring how treatment effects would possibly look like in a different study population (to which extent are results generalizable). Thus, strictly speaking, these analyses do not directly explore the robustness of the findings in the main analysis based.

Statistical analysis

We just briefly mention a few general considerations for the statistical analysis. The choice of estimand will have some implications for the construction of a suitable statistical model. Often the chosen estimand will necessitate certain assumptions about the fate of treatments after dropout; these model assumptions will most of the time be unverifiable from the observed data, but could be sustained by knowledge about the observed biological or physiological mechanisms or processes. For instance, hypothetical estimands will often be estimated by a linear-mixed model for repeated measurements where all subjects are included with their observations obtained until they are dropping out or completing the study. There might often also be a need for using deterministic or multiple imputation to be able to evaluate the effects of assumptions about the missing data mechanism (missing at random or missing not at random) [14–16].

Need of future research

Estimands provide a unified approach toward defining treatment effects in terms of treatment contrasts derived under varying assumptions regarding the study population and handling post-randomization events such as dropouts and non-adherers.

We tentatively suggest that researchers and trialists in (clinical) nutrition research define one or more estimands when planning RCTs or, as a minimum, when analyzing and interpreting results from RCTs. For instance, compliance and adherence could be quantified by means of appropriate estimands, abandoning the misused chi-square tests.

At the same time, we acknowledge that RCTs in nutrition research may not need (at present) the same rigor as is the case for medical trials in the pharmaceutical industry. However, it would be helpful to have a more realistic understanding of what is indeed being estimated. In particular, a better appreciation of ITT estimands and, especially, what are not ITT estimands.

Recently, it has been argued that it is difficult to obtain reasonable treatment effects as they would be if a treatment was provided to patients in real life (de facto estimands) where subjects are not monitored by a study infrastructure, which may promote adherence directly or indirectly [17]. Therefore, it seems that the common advice on making all possible efforts to keep subjects enrolled in RCTs is counter-productive if the purpose of the study is mainly defined in terms of a de facto estimand. Therefore, in our view, many RCTs in clinical nutrition at present essentially only allow estimation of de jure estimands.

Within clinical nutrition it is admittedly challenging to define and characterize the types of estimands that are most suitable when evaluating efficacy, effectiveness, compliance, and adherence. Also, establishing and applying suitable statistical models is another challenge.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

References

1. International Conference on Harmonisation (ICH). E9(R1): Statistical principles for clinical trials: Addendum Estimands and Sensitivity Analyses in Clinical Trials; 2017.
2. Mallinckrodt CH. Preventing and treating missing in longitudinal clinical trials. Cambridge: Cambridge University Press; 2013.
3. Leuchs A-K, Zinserling J, Brandt A, Wirtz D, Benda N. Choosing appropriate estimands in clinicals. Ther Innov Regul Sci. 2015;49:584–92.

4. Helms RW. Precise definitions of some terminology for longitudinal clinical trials: subjects, patient populations, analysis sets, intention to treat, and related terms. Pharm Stat. 2016;15:471–85.
5. Carpenter JR, Roger JH, Kenward MG. Analysis of longitudinal trials with protocol deviation: a framework for relevant, accessible assumptions, and inference via multiple imputation. J Biopharm Stat. 2013;23:1352–71.
6. Permutt T. A taxonomy of estimands for regulatory clinical trials with discontinuations. Stat Med. 2016;35:2865–75.
7. Mallinckrodt CH, Molenberghs G, Rathmann S. Choosing estimands in clinical trials with missing data. Pharm Stat. 2017;16:29–36.
8. Geiker NRW, Ritz C, Pedersen SD, Larsen TM, Hill JO, Astrup A. A weight-loss program adapted to the menstrual cycle increases weight loss in healthy, overweight, premenopausal women: a 6-mo randomized controlled trial. Am J Clin Nutr. 2016;104:15–20.
9. White IR, Carpenter J, Horton NJ. Including all individuals is not enough: lessons for intention-to-treat analysis. Clin Trials. 2012;9:396–407.
10. Olsen MF, Abdissa A, Kæstel P, Tesfaye M, Yilma D, Girma T, et al. Nutrient supplements improve weight gain, lean mass, grip strength and immune recovery in HIV-patients initiating antiretroviral treatment: a randomised controlled trial in Ethiopia. BMJ. 2014;348:g3187.
11. Perkin MR, Logan K, Tseng A, Raji B, Ayis S, Peacock J, et al. Randomized trial of introduction of allergenic foods in breast-fed infants. New Engl J Med. 2016;374:1733–43.
12. Leuchs A-K, Brandt A, Zinserling J, Benda N. Disentangling estimands and the intention-to-treat principle. Pharm Stat. 2017;16:12–19.
13. Johnston BC, Guyatt GH. Best (but oft-forgotten) practices: intention-to-treat, treatment adherence, and missing participant outcome data in the nutrition literature. Am J Clin Nutr. 2016;104:1197–201.
14. Lu K. An analytic method for the placebo-based pattern-mixture model. Stat Med. 2014;33:1134–45.
15. Holzhauer B, Akacha M, Bermann G. Choice of estimand and analysis methods in diabetes trials with rescue medication. Pharm Stat. 2015;14:433–47.
16. Little R, Kang S. Intention-to-treat analysis with treatment discontinuation and missing data in clinical trials. Stat Med. 2015;34:2381–90.
17. Akacha M, Bretz F, Ruberg S. Estimands in clinical trials – broadening the perspective. Stat Med. 2017;36:5–19.