

Beware of on-treatment safety analyses

Fred Yang¹, Janet Wittes² and Bertram Pitt³

Clinical Trials

1–8

© The Author(s) 2018

Article reuse guidelines:

sagepub.com/journals-permissions

DOI: 10.1177/1740774518812774

journals.sagepub.com/home/ctj



Abstract

Introduction: Assessing safety is important to evaluating new medications. In many randomized clinical trials, assessment of safety relies on so-called on-treatment analysis, where data on adverse events are collected only while the participant is taking study medication and perhaps for a few (7, 14, or 30) days after stopping. This article discusses the consequence of such failure to use intent-to-treat analyses in assessing safety.

Methods: This article discusses two approaches to analysis of safety data: intention-to-treat and on-treatment analysis with reference to principles of the design of randomized clinical trial.

Results: On-treatment analysis violates randomization and is often not well defined. Moreover, because the typical on-treatment analysis ignores the reason participants in clinical trials stop treatment, on-treatment analyses can lead to biased estimates of risk. Examples show biases that can result from failure to count all adverse events. An example from a study of rofecoxib shows an on-treatment analysis that led to likely underestimation of harm; an example from a study of saxagliptin shows an on-treatment analysis that led to a likely overestimate of harms.

Conclusion: For major safety outcomes in long-term clinical trials, intention-to-treat analysis should be performed in the framework of benefit–risk evaluation. More generally, analyses of safety should be tailored to the specific question being asked with the specific study design under consideration. On-treatment analyses are subject to bias; however, the direction of that bias is not necessarily clear.

Keywords

Randomized clinical trial, safety analyses, cardiovascular outcomes, on-treatment analysis, intent-to-treat analysis

Introduction

When a patient taking a drug experiences an adverse event, an important question is whether the drug actually caused the event. When such events occur during a randomized clinical trial, in the absence of direct biological evidence of causation, the question that the trial formally addresses is whether assignment to the drug caused the events. The formal analysis of causation should compare the treatment arms defined by assignment to the drug rather than by directly linking the individual event to the drug in question. Such an analysis uses an intention-to-treat approach, counting every event from the time of randomization to the end of the study and assigning all participants their randomized treatment. The conclusion is that the probability of the event in the treated arm is greater, or less, than the probability of the event in the control. To simplify the discussion, we consider randomized two arm trials comparing an experimental drug to placebo.

While clinical trialists as a group accept the principle of analysis by intent-to-treat for efficacy, many

question the paradigm for safety: how can a drug have caused an event if the affected person never took the drug or is no longer taking it? The usual approach to safety employs “on-treatment” analyses, where “on-treatment” refers to what the participant is taking at a specific time. Rather than characterize the trial’s participants by their assigned drug, the typical analysis uses a “safety” population: those who received any amount of experimental drug are classified in the “drug” group; those who took only placebo are counted in the “placebo” group. To keep the argument simple, consider trials in which all participants receive the treatment to which they were randomized; no one receives any amount of product from the other arm. In such trials,

¹KBP BioSciences, Princeton, NJ, USA

²Statistics Collaborative, Inc., Washington, DC, USA

³University of Michigan, Ann Arbor, MI, USA

Corresponding author:

Janet Wittes, Statistics Collaborative, Inc., 1625 Massachusetts Avenue, NW, Suite 600, Washington, DC 20036, USA.

Email: janet.wittes@statcollab.com

some participants fail to take their entire prescribed course of study medication. Some in each treatment group prematurely stop taking study medication; some experience adverse events after stopping study medication. Many trialists argue that, for safety analysis, an event occurring after the participant is no longer taking the experimental medication should not count against it. Rather, they argue that analysis should focus on events that occur while the participant is taking the study medication.

The definition of “on-treatment” may vary from study to study. Usually, the “on-treatment” period extends for several days after the participant has stopped taking study drug. Sometimes the period is five half-lives of the drug because, as the argument goes, how could what is essentially a homeopathic dose of a drug cause harm? In this article, we warn that such on-treatment analyses can sometimes dramatically underestimate or overestimate the harms of a drug. An example of underestimation comes from studies of rofecoxib while one of overestimation comes from saxagliptin; both are presented below.

Because efficacy analyses have traditionally respected randomization, intention-to-treat analysis is standard for assessing efficacy. Intention-to-treat analysis focuses on “effectiveness” (how a population will respond) rather than “efficacy,” which is driven by pharmacologic response.¹ Researchers generally recognize that changes due to treatment can introduce a cascade of events even after participants stop study intervention. For example, in studies examining the effect of lowering low-density lipoprotein cholesterol, one counts all clinical events, even among participants who have stopped study intervention. The justification is twofold. First, this type of analysis respects the randomization. Second, short-term exposure to an intervention can lead to long-term effects. For example, in the Multiple Risk Factor Intervention Trial, the intervention arm did not show statistically significant evidence of benefit during the period of the study itself;² however, long-term follow-up showed a convincing benefit of intervention, even though the dietary habits of the intervention and control groups had converged by the end of the 20 years after the trial was officially over.³

Given our understanding of the potential role of intervention after participants stop taking experimental therapy, why ignore that lesson in safety analyses? Many short-term exposures lead to measurable long-term harm. Lung cancer in ex-smokers is an example of damage after the end of exposure; the rate of lung cancer in ex-smokers, long after the half-life of components in the smoke, is higher than in those who never smoked; we do not exonerate smoking as a causative agent because exposure has ended. An example more relevant to drugs is situations in which a drug causes a transient decrease in systemic blood pressure. This, occurring in

patients with preexisting cerebral, coronary, or renal vascular disease, could result in ischemia, cell death, or increased inflammatory cytokines—the consequences of which might not become manifest for some time. Myocardial fibrosis as a consequence of myocardial inflammation and/or cell death, for instance, could be a focus for subsequent ventricular arrhythmias and/or sudden cardiac death.⁴

These examples indicate how failure to count events after stopping study drug can underestimate harms; the Saxagliptin Assessment of Vascular Outcomes Recorded in Patients with Diabetes Mellitus-TIMI 53 (SAVOR-TIMI 53) trial below shows how undercount can overestimate the relative rate of harm in a study arm.

Problems with on-treatment analyses of safety

We describe three problems with on-treatment analyses: the first is fundamental, and the other two are secondary.

Problem #1: violation of randomization

Randomization provides the basis for unbiased comparison between groups in randomized experiments.⁵ Assignment of treatment based on chance leads to the expectation that the experimental groups will be well balanced with respect to both measured and unmeasured variables. The larger the sample size, the more similar the actual groups are likely to be. The only differences expected between the groups are those induced by the experimental treatment. Randomization allows comparison of the risks and benefits of the intervention and control treatment over the period of the study.

Randomization, though necessary, is not sufficient for unbiased comparison. Randomization must go hand-in-hand with methods of analysis that are based on the randomized groups. The methods used should not themselves introduce bias. Ignoring data from randomized participants, when the ignored data are based on post-randomization observations, violates randomization inviting bias due to post-randomization confounding. On-treatment analysis is a prime example of ignoring these types of data because it removes follow-up time from the randomized groups as a function of post-randomization experience. Some well-known examples of this problem come from the Anturane Reinfarction Study,^{6,7} which showed the danger of excluding some events; a classic paper on the Coronary Drug Project,⁸ which demonstrated bias induced by analysis of treatment adherers; and the paper by Peduzzi et al.⁹ on coronary artery bypass surgery, which describes bias caused by on-treatment and as-treated approaches.

Consider an adverse event experienced by a participant randomized to active therapy who never received it. Of course, that active therapy could not have “caused” the event if “cause” is defined in an Aristotelian sense of efficient cause. Clinical trials, however, rarely can identify such efficient causes when the events, like myocardial infarctions and cancers, are not surprising in the population being studied; the best a trial can usually do is ask whether the event rate of interest differs between the treated and control groups. Such imbalance indicates causality. Distorting the randomized groups violates the fundamental principle that allows unbiased comparison of treatment to control. Such distortion can occur by removing some participants after randomization because they failed to comply with study procedures or by not counting some events during the protocol-defined follow-up time. Should the trialists decide to break randomization, they must recognize that they have converted their randomized trial into an observational study, which they should analyze under the principles governing observational analyses.¹⁰

Problem #2: being “on” or “off” treatment is not well defined

In the experimental group of studies comparing drug to placebo, it is clear when participants stop study drug. The definition of stopping placebo, however, is slippery because placebo introduces no additional pharmacological intervention. (Stopping placebo could, however, cause increased anxiety, leading to sympathetic activation which might have an adverse effect, including sudden cardiac death.)

In blinded studies where the control is placebo on top of standard of care, the active group can go off-treatment, but the control group, by definition, cannot. This asymmetry can lead to longer follow-up time in the control group and, since safety is assessed by comparing the two groups, the comparison can lead to underestimation of the extent of harms in the active group. Correcting this problem by calculating rates of events (i.e. events divided by time) makes the often unwarranted assumption of constant hazard.

From the perspective of benefit–risk of an intervention, the evaluation of both benefits and risk should be performed within the same window of time.

Problem #3: it ignores the reasons for stopping study drug and it fails to count events that occurred after stopping

An important reason for intent-to-treat analysis of safety stems from the common observation that many participants of clinical trials who stop taking their prescribed study drug do so because of illness. Some

experience a seemingly minor event that presages a major one. For example, they may have experienced fluid retention before they stopped blinded study medication; 3 weeks later, they are hospitalized for congestive heart failure.¹¹ Analyses that count all serious adverse events, regardless of their timing, would capture these episodes of heart failure; however, analyses that include only the on-treatment period would ignore these episodes, thus introducing post-randomization confounding and potential bias.

The Adenomatous Polyp Prevention on Vioxx (APPROVe) trial, which studied whether rofecoxib could prevent adenomatous polyps, provides an example. The report of the original analysis only counted cardiovascular events that occurred within 14 days of cessation of study medication.¹² Further follow-up that counted later events increased the estimated relative risk of rofecoxib to placebo. Presumably, limiting the period of interest to 14 days failed to account for the delayed effect of increased cardiovascular risk caused by the drug. In this trial, the proportion of participants who stopped rofecoxib was higher than the proportion who stopped placebo; thus, the on-treatment analysis understated the adverse effect of rofecoxib.¹³

Example of on-treatment analysis overestimating harm: the SAVOR-TIMI 53 trial in type 2 diabetes mellitus

In response to the Food and Drug Administration’s guidance describing requirements for new drugs for type 2 diabetes mellitus,¹⁴ many companies have conducted double-blind placebo controlled trials assessing the cardiovascular safety of drugs for type 2 diabetes mellitus.¹⁵ The endpoints are cardiovascular events. Most of these studies are event-driven with such long durations that many participants experience worsening of their type 2 diabetes mellitus during the trial. Consequently, to adhere to diabetes treatment guidelines, these participants modify their diabetes regimen. These cardiovascular outcome trials compare the new drug against placebo, both on the background of “standard of care,” with protocolized “rescue” procedures specifying that, regardless of whether participants are on or off their randomized medication, those in both the new drug and placebo arms should receive needed additional medications to treat their diabetes. The protocols state that investigators should follow participants who withdraw from active treatment for collection of data on cardiovascular outcomes. In general, the studies have had low proportions (typically <3%) of loss to follow up.

These studies aim to show non-inferiority or, better yet, superiority to placebo with respect to cardiovascular outcomes. Some researchers have suggested using an on-treatment approach for the non-inferiority analysis and an intent-to-treat approach for testing superiority.

Table 1. SAVOR-TIMI 53 study: all-cause mortality using various definitions of “on-treatment.”

	Saxagliptin N = 8240	Placebo N = 8173	Hazard ratio (95.1% CI) ^a
All “on study” deaths	416 (5.1%)	376 (4.6%)	1.10 (0.96–1.27)
Deaths on-treatment + 30 days	297 (3.6%)	248 (3.0%)	1.18 (0.99–1.39)
Deaths on-treatment + 7 days	256 (3.1%)	204 (2.5%)	1.23 (1.02–1.48)

CI: confidence interval.

^aA 95.1% CI was used to address the fact that the study had interim analyses.

Source: Adapted from Food and Drug Administration meeting material for SAVOR-TIMI 53 advisory committee.²⁰

The literature is inconsistent regarding the proper population for analysis for non-inferiority trials.^{16,17} Geiger et al.¹⁸ discuss the choice in the context of cardiovascular outcome trials. We define two approaches:

An “intent-to-treat” approach, which counts all events until a participant is lost to follow up, withdraws consent for follow-up or the trial ends. For participants with no event, time on the trial is censored at the time of loss to follow up, time to withdrawal of consent, or when the trial ends. Such censoring is assumed non-informative (although loss to follow up and withdrawal of consent may not be truly non-informative).

An “on-treatment plus x days” approach counts events that occur while participants are taking study medication and a subsequent x-day period after discontinuation of medication. For participants with no event, time in the trial is censored x days after discontinuation of study medication.

Consider, for example, mortality assessed with an “on treatment plus 7 day” analysis. This excludes all deaths occurring 8 or more days after last dose of study drug. Many of these excluded events, if evaluated carefully, would turn out to have been the consequence of an adverse event that began while the participants were taking the randomized treatment. In a clinical trial, once a patient has experienced a life-threatening adverse event, the treating physician would likely stop all blinded study medication; however, an on-treatment analysis would inappropriately exclude those deaths, thus leading to an underestimated mortality rate. If the two arms had equal rates of stopping study medication for impending death, the mortality rates in both arms would be underestimated, but the relative rates might remain unchanged. Depending on the nature of the events excluded, the proportion of events excluded, and the imbalance of exclusions between study arms, an on-treatment approach may introduce unacceptable bias, either absolute or relative, in the evaluation of safety. What is worse, the direction of the bias is often not predictable which makes interpretation difficult.

In trials of type 2 diabetes mellitus, given the nature of the study design, in both the active treatment and placebo arms, patients are to receive “rescue” treatment for hyperglycemia while staying on treatment. If the experimental therapy in fact leads to glycemic control

superior to that of the placebo, more participants will be “rescued” in the placebo arm. Our example, the SAVOR-TIMI 53 trial, compared saxagliptin to placebo with all participants treated also with standard of care in 16,500 type 2 diabetes mellitus patients aged 40 or above. Participants were randomized 1:1 to saxagliptin or placebo and followed for an average of 3 years. Eligibility required an HbA1c between 6.5% and 12%, and high cardiovascular risk.¹⁹

The Food and Drug Administration’s review of the safety of saxagliptin cited a hazard ratio of 1.18 for the on-treated plus 30-day analysis and even higher hazard ratios for on-treated plus shorter periods (see Table 1), implying that saxagliptin led to increased mortality rates.

For each participant in the study, the years off-treatment are based on the duration of time while the participant was off-treatment; that time does not include any time before going off study drug. Looking only at the mortality rate for the placebo group classified by time on- and off-treatment, the data show a remarkably higher mortality rate for placebo participants who went off their placebo medication than for those who remained on it (see Figure 1(a)). Absent a belief that placebo is biologically protective, this graph is a demonstration, reminiscent of similar findings in the Coronary Drug Project, that adherers to placebo tend to have better outcomes than do non-adherers.⁸ One likely explanation in this case is that non-adherence to placebo was a marker of having experienced a life-threatening adverse event; once such an event occurred, the participant was likely to stop randomized treatment and would have a higher mortality rate than those remaining on their randomized treatment during the same time period.

The figure shows that the off- and on-treatment curves for the two treatment groups are almost superimposable: in both groups, the mortality rate while participants are off their treatment is roughly five times higher than when they are on-treatment (Figure 1(b)). Going off placebo is not logically associated with a fivefold increase in the chance of death. The figure is indicative of the danger of estimating the event rate purely on the basis of on-treatment rates—in this case, the on-treatment rates underestimate the event rates experienced over the course of the trial.

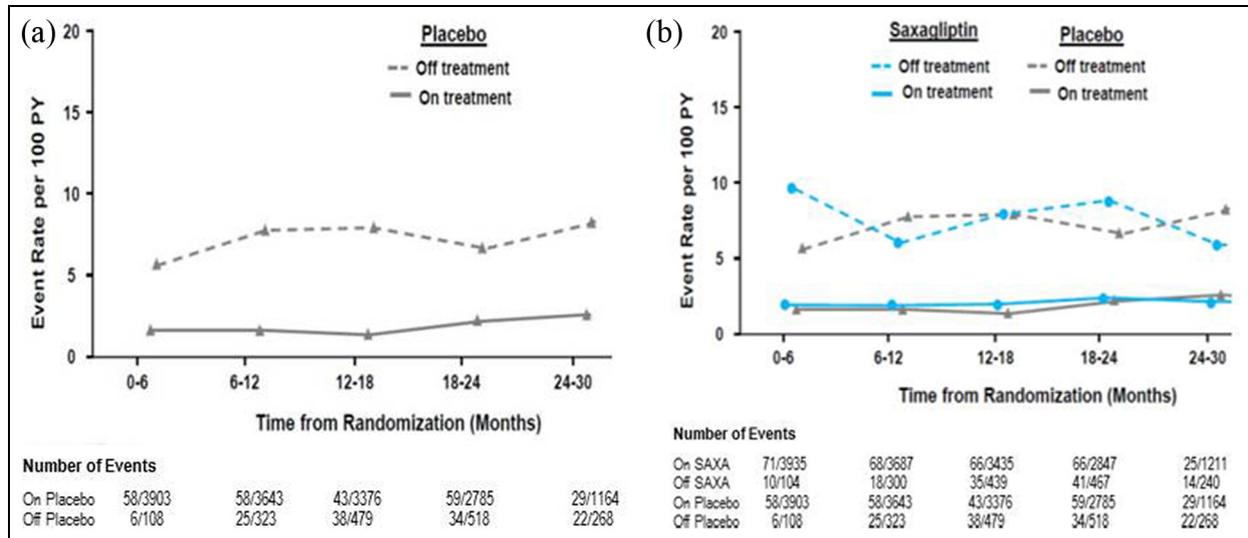


Figure 1. Mortality rate in participants on- and off-assigned randomized treatment. (a) Mortality rate for participants randomized to placebo. (b) Mortality rate for participants randomized to saxagliptin superimposed on the rates in part (a). Each tick in the graphs represents the mortality rate per 100 patient-years corresponding to the midpoint of the period represented by the x-axis. The ticks above the “0–6” on the x-axis depict the mortality rate for the participants on- and off-treatment throughout 0 to 6 months. For example, a person who was on-treatment for 2 of those 6 months and off for 4 would contribute 2 months to the “on-treatment” and 4 months to the “off-treatment” graphs. In that 0-6 month period, a total of six events occurred to participants “off” placebo, 58 to those “on” placebo, 10 to those “off” saxagliptin, and 71 to those “on saxagliptin.” For each participant in the study, the years off-treatment are based on the duration of time while the participant was off-treatment; the times do not include any time before going off study drug. Source: Adapted from Slide ST-26 and ST-27 used by AstraZeneca in Food and Drug Administration advisory meeting on the SAVOR-TIMI 53 study²⁰ (AstraZeneca provided the authors with the denominators).

Table 2. Example of bias caused by on-treatment analysis.

Participants	Treatment N = 100 n (%)	Placebo N = 100 n (%)	Relative risk (treatment/placebo)
Who completed treatment	60	40	
Who discontinued treatment	40	60	
Whose first event occurred while on-treatment	15	10	1.5
Whose first event occurred while off-treatment	5	10	0.5
Overall proportion having at least one event	20/100 (20%)	20/100 (20%)	1.0
Proportion having at least one event while on-treatment	15/100 (15%)	10/100 (10%)	1.5

The data underlying these curves show a second type of bias, one that occurs in estimating the relative risk when the proportion of events excluded differs by treatment group. When one treatment is more effective than the other, more participants in the less effective group are likely to stop study medication. In the SAVOR TIMI 53 study, the group randomized to saxagliptin experienced glycemic control superior to that of those randomized to placebo. Therefore, more placebo than saxagliptin participants stopped study medication and started rescue treatment. Thus, the on-treated analyses excluded more deaths from the placebo group than from the saxagliptin group, leading to an overestimate of relative risk. An intention-to-treat approach for safety analysis is similar to what happens in the “real world.” A cardiovascular outcome study that compares randomized regimens

serves as a bridge between data from a randomized clinical trial and the “real world.”

At the SAVOR TIMI 53 Food and Drug Administration advisory meeting, the authors of this article presented earlier versions of the figure. The members of the advisory panel discussed the issues related to mortality and agreed that the data led to a moderate concern about all-cause mortality but, they said, “the statistically significant findings were based on sensitivity analyses that censored deaths near treatment exposure and [they] were uncertain about the strength and validity of these sensitivity analyses.”²² The Food and Drug Administration included the intent-to-treat analysis results for all-cause mortality in the label for saxagliptin.²³

Table 2, a hypothetical example, illustrates the potential bias, both for absolute and relative risk, from an on-treatment type of analysis.

Conclusion

We have argued thus far that on-treatment analyses are problematic because of their inherent biases; however, pure intent-to-treat analyses that capture all events from randomization to the end of the trial have limitations as well. Long trials with mortality or commonly occurring events (e.g. infections) as measures of harm may find no difference in the proportion of participants experiencing the event even when they find differences in time to event. Similarly, because some events do in fact occur in proximity to exposure, counting all events over time will tend to attenuate the adverse effects of treatment.

Evaluation of safety data should be tailored to the study design as well as to the objectives of the trial. The ultimate goal in a randomized trial should be to evaluate benefits and risks. “Short term” small randomized controlled studies are suitable for evaluating short-term effects of the product on both benefit and risk. The estimands (the statistics to be estimated²⁴) assess tolerability and short-term pharmacologically induced efficacy; data on adverse events unrelated to tolerability, however, are typically sparse. One evaluates benefit and tolerability in similar ways by using approaches that include all randomized subjects and employing data up to certain defined protocol-defined times. This framework of evaluating short-term benefits and risks is balanced. Short-term trials, however, rarely can identify the scope or frequency of adverse events that patients who remain on a drug long term would be likely to experience.

In long-term outcome studies, patients are randomized to one of two (or more) treatment regimens instead of to a specific rigidly define treatment. Such a study becomes vulnerable to post-randomization confounding and selection bias in the evaluation of both efficacy and safety. Evaluation of benefit focuses on effectiveness, for it is based on intention-to-treat analysis, which includes all efficacy data regardless of whether the patient is on- or off-treatment. Including in the risk evaluation only “on-treatment” events (similar to what is done in a short-term study with full compliance) would lead to a mismatch of the estimands for benefit and risk. Importantly, excluding different periods of time for the treatment groups introduces selection bias into the evaluation of risk. The net result is an imbalance in the assessment of benefit and risk.

While we argue against routine on-treatment analysis, analyses should consider mechanism (even though often the assumed mechanism is incorrect or incomplete). For safety and tolerability events with a known mechanism of action associated with the drug, choice of analysis is relatively straightforward. For example, an on-treatment analysis can evaluate the frequency of

nausea and vomiting associated with glucagon-like peptide-1 (GLP-1) receptor agonists.²¹ On the other hand, chronically used medication may have unknown harms caused by unknown, or poorly understood, mechanisms of action. Ignoring adverse events that occur post-treatment, especially within the framework of evaluating benefits and risks, may cause a mismatch between estimands and biased conclusions. As mentioned earlier, often the direction of the bias is unclear.

Because intention-to-treat analyses that include post-treatment data may introduce noise, thus attenuating signals of harms, we suggest methods as outlined in Table 3 along with a twofold strategy:

- An evaluation of safety in the context of an assessment of benefit and risk should use an intention-to-treat approach to match the approach used to evaluate efficacy.
- An intention-to-treat analysis that shows a signal of harm should prompt exploratory analyses to facilitate better understanding of the likely mechanism of the causal relationship of the product to the event.

Analyses of adverse events often are reported in long lists of proportions of participants experiencing a specific event. In exploring whether a drug causes a specific adverse experience, proportions are often insensitive to true effects, and biased removal of participants from denominators or biased removal of time of follow-up may overestimate or underestimate the product’s true risks. Even large clinical trials are often too small to produce precise estimates of rates of uncommon adverse events. Analyses of both safety and efficacy should be tailored to the specific question being asked; methods that address time-to-event rather than proportions are often useful so long as time is censored in an unbiased manner. Note that even for the events in Table 3 labeled “tolerability” and “known drug-induced adverse events” where we do recommend on-treatment types of analysis, one may need to include analyses of time off-treatment to explore whether the drug has lingering effects.

Declaration of conflicting interests

Dr Yang was an employee of AstraZeneca and is a current employee of KBP BioSciences. Dr Wittes is President of Statistics Collaborative which has contracts with many pharmaceutical companies, including AstraZeneca. Dr Pitt has served as a consultant for Bayer, Sanofi, AstraZeneca, Relypsa/Vifor, scPharmaceuticals, Stealth Peptides, Sarfez, Tricida, and KBP Pharmaceuticals. He has stock option in Relypsa, scPharmaceuticals, Tricida, Sarfez, and KBP Pharmaceuticals. He has a patent pending for site-specific delivery of eplerenone to the myocardium.

Table 3. Suggestions for approaches to analysis of safety in randomized clinical trials.

Type of event	Example	How to assess causality	Consequences to study design	Analysis
Tolerability	Nausea and vomiting with GLP-1 antagonists; Hand-foot syndrome with capecitabine	Dechallenge–rechallenge or understanding of mechanism of action	No specific consequence	Capture data on these events carefully so as to be able to calculate the proportion of participants affected and, if relevant, the duration of event. Use on-treatment analysis (but consider including analyses of some time off-treatment to explore whether the drug has lingering effects) Capture data on these events carefully so as to be able to calculate the proportion of participants affected and, if relevant, the duration of event. Use intent-to-treat analysis
Laboratory value	Increased potassium in use of spironolactone	Dechallenge–rechallenge or understanding of mechanism of action	Likely increase in patient discontinuations and reduced dose of study drug. May be increase in sudden death where cause of death is not identified in trial, consequently need for increased sample size No specific consequence	Capture data on these events carefully so as to be able to calculate the proportion of participants affected and, if relevant, the duration of event. Use on-treatment analysis (but consider including analyses of some time off-treatment to explore whether the drug has lingering effects) Consider altering the study. Use on-treatment analysis
Known drug-induced adverse events	Anaphylaxis; Stevens–Johnson syndrome; acute liver toxicity	Assume the drug caused the event	No specific consequence	Capture data on these events carefully so as to be able to calculate the proportion of participants affected and, if relevant, the duration of event. Use on-treatment analysis (but consider including analyses of some time off-treatment to explore whether the drug has lingering effects) Consider altering the study. Use on-treatment analysis
Rare event in population under study	Heart attacks in 20-year-old healthy women on high-dose experimental contraceptive	Assume the drug caused the event	Consider reducing dose	Capture data on these events carefully so as to be able to calculate the proportion of participants affected and, if relevant, the duration of event. Use on-treatment analysis (but consider including analyses of some time off-treatment to explore whether the drug has lingering effects) Consider altering the study. Use on-treatment analysis
Event anticipated in the population under study	Heart attack or stroke in middle-aged or elderly participant, leukemia, infections	If the rate of event in the treatment group under study is convincingly higher than in the control, the treatment can be said to be causal; however, a specific event cannot usually be attributed to the drug	Protocol should specify the need for follow-up of each participant until the end of the trial. This will increase the cost of clinical trials but will allow more accurate assessment of risks	Use intention-to-treat analyses as described in the body of the text

GLP-1 : glucagon-like peptide-1.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

References

1. Feinstein A. An additional basic science for clinical medicine: II. *Ann Intern Med* 1983; 99(4): 544–550.
2. Burch P. The Multiple Risk Factor Intervention Trial. *JAMA* 1983; 249: 1435–1437.
3. Stamler J and Neaton JD. The Multiple Risk Factor Intervention Trial (MRFIT)—importance then and now. *JAMA* 2008; 300: 1343–1345.
4. Mak G, Murphy NF, Ali A, et al. Multiple neurohumoral modulating agents in systolic dysfunction heart failure: are we lowering blood pressure too much? *J Cardiac Failure* 2008; 14: 555–560.
5. Fisher RA. *The design of experiments*. New York: Hafner, 1935.
6. Sherry S. The Anturane Reinfarction Trial. *Circulation* 1980; 62(Suppl. V): 73–78.
7. Temple R and Pledger GW. The FDA's critique of the Anturane Reinfarction Trial. *N Engl J Med* 1980; 303(25): 1488–1492. doi:10.1056/NEJM198012183032534.
8. The Coronary Drug Project Research Group. Influence of adherence to treatment and response of cholesterol on mortality in the Coronary Drug Project. *N Engl J Med* 1980; 303(18): 1038–1041. doi:10.1056/NEJM198010303031804.
9. Peduzzi P, Detre K, Wittes J, et al. Intent-to-treat analysis and the problem of crossovers. *J Thorac Cardiovasc Surg* 1991; 101(3): 481–487.
10. Hernan MA, Hernandez-Diaz S and Robins JM. Randomized trials analyzed as observational studies. *Ann Intern Med* 2013; 159(8): 560–562. doi:10.7326/0003-4819-159-8-201310150-00709.
11. Chin MP, Wrolstad D, Bakris GL, et al. Risk factors for heart failure in patients with type 2 diabetes mellitus and stage 4 chronic kidney disease treated with bardoxolone methyl. *J Card Fail* 2014; 20(12): 953–958. doi:10.1016/j.cardfail.2014.10.001.
12. Bresalier RS, Sandler RS, Quan H, et al. Cardiovascular events associated with rofecoxib in a colorectal adenoma chemoprevention trial. *N Engl J Med* 2005; 352: 1092–1102.
13. Baron JA, Sandler RS, Bresalier RS, et al. Cardiovascular events associated with rofecoxib: final analysis of the APPROVe trial. *Lancet* 2008; 372(9651): 1756–1764. doi:10.1016/S0140-6736(08)61490-7.
14. U.S. Department of Health and Human Services, Food and Drug Administration, Center for Drug Evaluation and Research. *Guidance for industry: diabetes mellitus—evaluating cardiovascular risk in new antidiabetic therapies to treat type 2 diabetes*, December 2008.
15. Yang F, Stewart M, Ye J, et al. Type 2 diabetes mellitus development programs in the new regulatory environment with cardiovascular safety requirements. *Diabetes Metab Syndr Obes* 2015; 8: 315–325. doi:10.2147/DMSO.S84005.
16. Wiens BL and Iglewicz B. Testing noninferiority of response rates for regulatory filings using transformations. *Drug Information Journal* 2001; 35: 1165–1171.
17. Schumi J and Wittes J. Through the looking glass: understanding non-inferiority. *Trials* 2011; 12: 106. doi:10.1186/1745-6215-12-106.
18. Geiger MJ, Mehta C, Turner JR, et al. Clinical development approaches and statistical methodologies to prospectively assess the cardiovascular risk of new antidiabetic therapies for type 2 diabetes. *Ther Innov Regul Sci* 2015; 49(1): 50–64. doi:10.1177/2168479014549860.
19. Spinar J and Smahelova A. Saxagliptin and cardiovascular outcomes in patients with type 2 diabetes mellitus. *Vnitř Lek* 2013; 59(11): 1003–1007.
20. AstraZeneca. SAVOR saxagliptin assessment of vascular outcomes recorded in patients with diabetes mellitus: sNDAs for Onglyza (22-350/S-014) and Kombiglyze XR (200-678/S-013) briefing document for endocrinologic and metabolic drugs advisory committee meeting. https://www.pharmamedtechbi.com/~media/Supporting%20Documents/The%20Pink%20Sheet%20DAILY/2015/April/Saxagliptin_AC_AZ_brfg.pdf (2015, accessed November 8, 2018).
21. Prasad-Reddy L and Isaacs D. A clinical review of GLP-1 receptor agonists: efficacy and safety in diabetes and beyond. *Drugs Context* 2015; 4: 212283. doi:10.7573/dic.212283.
22. Food and Drug Administration Center for Drug Evaluation and Research. Summary minutes of the Endocrinologic and Metabolic Drugs Advisory Committee Meeting. <http://wayback.archive-it.org/7993/20161023221510/http://www.fda.gov/downloads/AdvisoryCommittees/CommitteesMeetingMaterials/Drugs/EndocrinologicandMetabolicDrugsAdvisoryCommittee/UCM453900.pdf> (2015, accessed 30 August 2018).
23. Food and Drug Administration. Full prescribing information ONGLYZA (saxagliptin). https://www.accessdata.fda.gov/drugsatfda_docs/label/2016/022350s014lbl.pdf (2016, accessed 30 August 2018).
24. Mehrotra DV, Hemmings RJ and Russek-Cohen E. Seeking harmony: estimands and sensitivity analyses for confirmatory clinical trials. *Clin Trials* 2016; 13(4): 456–458. doi:10.1177/1740774516633115.