Writing results in a clinical trial report

Mara Therese P. Evangelista MD, DPDS and Nick Huber

During the course of a clinical trial, large amounts of data are usually collected at different time points (i.e. at baseline and during follow-up visits). This results in the possibility of multiple statistical tests and various ways that the results can be reported. Some authors would naturally hope to show 'interesting' data, and this can lead to a tendency to hype the results. This must be resisted. It is of the utmost importance to report a balanced view of the results. This is why it is essential, before a trial starts, to have a protocol that states the primary and secondary endpoints and includes a pre-specified analysis plan.

The CONSORT guidelines (Consolidated Standards of Reporting Trials) help authors improve reporting through the use of a checklist of items, which should be included when reporting a clinical trial, as well as a flow diagram to show the different stages that patients go through in a trial.

This article will focus on how to write the Results section of a paper, considering each of the seven subheadings from the CONSORT checklist (2010). The checklist has seven topics listed for the Results section, starting at item 13:

13. Participant flow
14. Recruitment
15. Baseline data
16. Numbers analyzed
17. Outcomes and estimation
18. Ancillary analyses
19. Harms

Writing About Participant Flow

The first question we need to answer in our Results section is: "What happened to the participants in the trial?" Most clinical trials do not enrol and follow up on all potential participants (e.g. some patients may not be eligible, some may not receive the intervention to which they are randomized, and some may be lost to follow-up). It is important to present clearly what happens to individuals in each group at each stage of the trial, so that the reader can assess the internal validity and the generalizability of the trial results.

If the trial is very simple with no dropouts or losses, it may be enough to describe the flow of participants through each stage of the trial in the text of the report. However, it is often better to present this information as a 'flow diagram', as recommended in the CONSORT guidelines (Figure 1).

The CONSORT flow diagram is designed to show what happens to patients throughout the course of a trial. It helps ensure that the authors present such data clearly, allowing the reader to see the flow of patients through the trial and the amount of follow-up available. The diagram gives information on the number of individ-

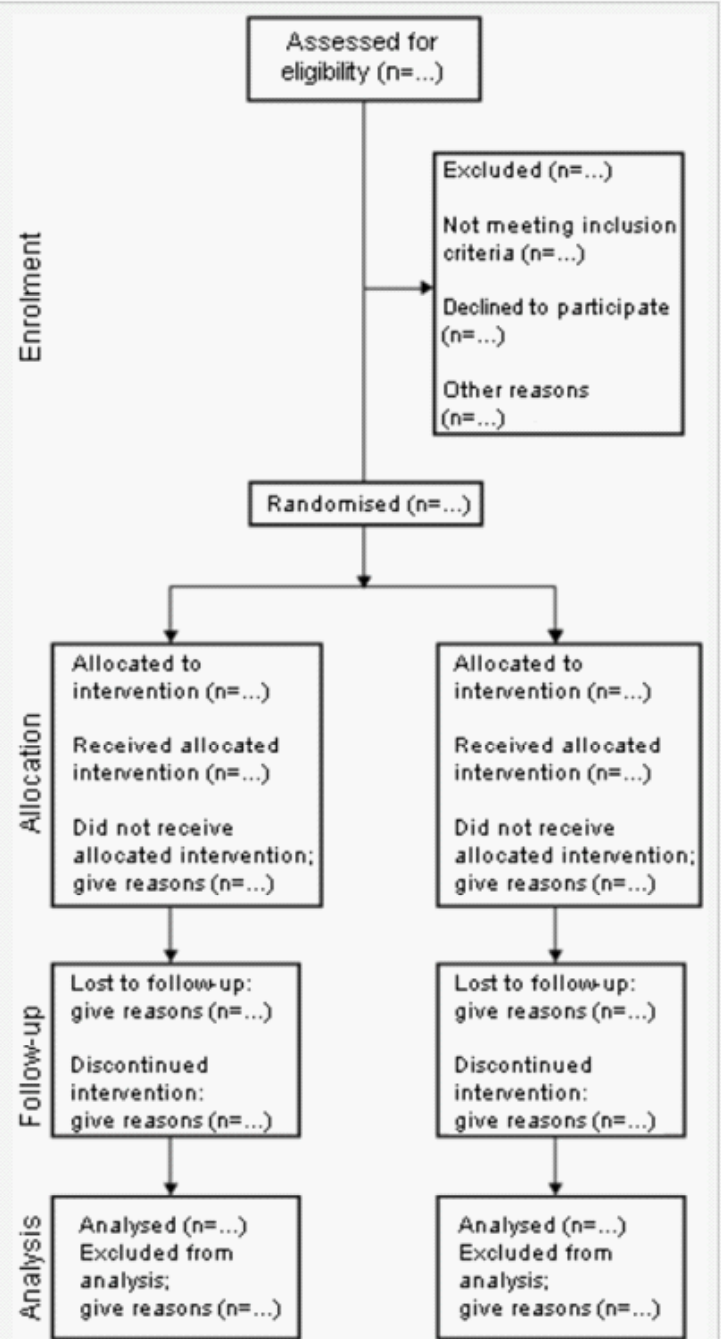

Figure 1. CONSORT Flow Diagram

uals, those who drop out and the reasons for their non-inclusion, at the five main stages of the trial:

- Enrolment
- Randomization
- Allocation of intervention
- Follow-up
- Analysis

Enrolment and Randomization

An important question in any trial is, "To whom can we apply the results?" This can be answered by considering who was included in the trial and who was excluded. Often the generalizability of a trial is limited by the inclusion criteria that restrict the trial population to only a small proportion of all patients with the disease under investigation.

Even the inclusion criteria for a trial may not give us a precise picture of the characteristics of those who were enrolled in the trial. For example, the patient profile of those who enrolled in the trial may be more specific than that outlined in the inclusion criteria such as:

- Characteristics of potential participants: only a subset of those eligible for the trial may be available in the study population. For example, the inclusion criteria may specify that all children (aged 1-16 years) are eligible, but the children with the disease of interest during the recruitment period may all happen to be children who are 10 years old and older.
- Consent: some eligible individuals may refuse to take part in the trial, and the consent rates may be lower in certain groups of individuals, for example amongst older people or among specific socio-economic groups.

All these factors play a role in the generalizability (the patient population where our results may be applicable to) of a trial. Therefore, we need to document the following in our trial report:

- How many individuals were considered for recruitment
- How many were recruited
- The reasons why individuals were not recruited (with numbers in each of these exclusion groups).

However, data may not be available on how many individuals were assessed for enrolment, and the reasons why some did not fulfill the inclusion criteria. In some trials, the time and effort put into ensuring good internal validity is done at the expense of obtaining information to help better understand generalizability. For this reason, some CONSORT flow diagrams leave out the information on enrolment and start instead with the flow of patients at randomization.

Allocation of the intervention

This stage shows the allocation of the intervention after patients have been enrolled and how many patients were randomized to each group. It also includes information on how many actually received the planned intervention and the reasons why some patients did not. This allows assessment of the internal validity of the trial.

Follow-up

The fourth stage reports the numbers in each group for those who were lost to follow-up and the numbers for those who discontinued the treatment (with reasons why). When either the number of drop-outs or the characteristics of those who dropped out differ between groups, this will likely lead to (post-randomization selection) bias, especially as the total number of patients dropping out rises. Therefore, this stage also enables the reader to assess the internal validity of a trial.

Analysis

The final stage demonstrates how many participants in each group were included in the main analysis, with reasons why anyone was excluded. This reporting allows the reader to assess if patients were analyzed in the group to which they were originally randomized, irrespective of whether they underwent the intended intervention or continued therapy according to the protocol (determining therefore, whether an intention-to-treat analysis was carried out).

If an intention-to-treat analysis is used, information on failure to follow the intended treatment should be part of the findings. It is not good practice to exclude data on those who deviate from the protocol. These deviations from the protocol should be reported in the text when possible. We will need reasons why participants did not follow the pre-specified protocol (e.g. patients used a medication not included in the specified protocol).

Review: How does an intention-to-treat analysis provide unbiased estimates?

Randomization ensures that the comparison groups are balanced at baseline for all known and unknown factors that may have an impact on the outcome. An intention-to-treat analysis preserves the equal distribution of measured and unmeasured characteristics in each group, resulting in an unbiased estimate of the treatment effect.

Review: How can an intention-to-treat analysis be applied in trials with patients who were lost to follow-up?

If there are losses to follow-up leading to missing outcome data, these patients cannot be included in the analysis even if an intention-to- treat analysis was planned. However, the missing data could be imputed using various methods (last observation carried forward, substitution of mean value, statistical models, etc.), thereby allowing the patients who were lost to follow-up to be included in the analysis.

Writing About Recruitment

It is of value that we state the start and end dates for recruiting participants and for following them up. This allows the reader to place the study in its historical context (i.e. determine the routine care the participants in the trial were likely to receive at that time).

Writing About Baseline Data

If randomization has been carried out properly, then treatment groups should be similar in their baseline demographic and clinical characteristics. Collection of baseline data allows:

- Assessment of the characteristics of the participants, and therefore the generalizability of the trial results
- Assessment of the comparability of groups in the different arms to determine whether randomization was successful in producing groups that differed only with respect to the trial intervention
- Inclusion of baseline data in pre-specified statistical analyses (sometimes done to improve the statistical precision of the effect estimate of interest, e.g. regression analysis)

Data that are collected at baseline (before randomization) from participants usually include:

- Demographic characteristics
- The nature and severity of the disease of interest
- Risk factors for the disease of interest
- Other medical conditions and treatments

These data need to be summarized and are usually presented in tables. For categorical data (such as sex), it is customary to report the numbers in each category, with percentages. For continuous data (e.g. PASI, SCORAD indices) the mean and standard deviation are usually given. It is not uncommon to see use of significance tests to assess differences between groups in baseline characteristics. This is, however, not necessary or appropriate. If randomization has been carried out correctly, any differences will have arisen due to chance.

Writing About Numbers Analyzed

As mentioned above, it is important to state the number of individuals in each group who were included in the analysis of each outcome. In addition, the number of participants with available data may vary for each outcome measure, and the denominators used for different analyses may be difficult to convey in the flow diagram. It is best to state the numbers of individuals included in each analysis in the text, irrespective of whether these numbers are also summarized in the flow diagram.

Writing About Outcomes and Estimation

For each outcome of interest in our trial, we need to report:

1. A summary of the outcome in each group; these outcomes may be binary or quantitative;
2. The effect of the intervention.

Effects can be expressed in different ways:

- For binary outcomes, we can use absolute effects (e.g. a risk difference), relative effects (e.g. a risk ratio) or relative risk reduction.
- For survival data, we can use relative effects (a hazard ratio) or differences in average survival time between groups.
- For continuous data, we usually use differences in the average (mean or median) outcome value between groups.

95% Confidence Intervals

Trial results must state how precise our effect measure is. This is because our study population is a sample of the underlying population, and so we have to consider sampling error when we interpret the effect estimates from our trial data. A 95% confidence interval (CI) reflects the uncertainty of a point estimate of an effect, and provides an indication as to where the true value is likely to lie. The main results from a trial are best presented in terms of some treatment effect together with a confidence interval and (usually) a p-value. Remember also that p-values on their own can be misleading – for example, they may reflect statistical significance but provide no measure of clinical importance.

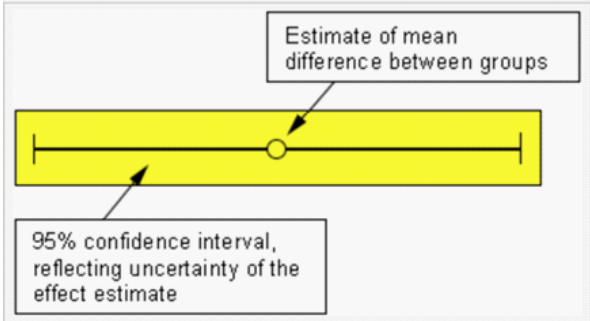

Figure 2. Confidence interval.

Writing About Ancillary Analyses

As stated in the introduction, it is possible to do many different analyses in a trial. The problem with doing multiple statistical tests on the same data is that we increase the probability of a type-1 error, i.e. obtaining a p-value of 0.05 for one of these analyses, even if there is no true effect of the intervention.

For example, the probability of having at least one significant ($p < 0.05$) result when 15 statistical tests are done on the same data is over 50%, even when no true difference exists for any of the comparisons made. Thus, it is misleading to do many analyses but only choose to present those analyses that produced statistically significant results. This results in a trial report that is likely to exaggerate the effects of the intervention. This is particularly a problem for analyses that are decided upon only after looking at the data. It is much more reliable to carry out planned analyses that were specified in the protocol before the trial started.

Writing About Harms

Adverse events are often reported poorly in trial publications. In 2003, the CONSORT members considered ways to further the reporting of adverse events in trials. As a result, ten new recommendations for the reporting of adverse events were established: one to appear in the Title/Abstract, one to appear in the Background, three to appear in the Methods, four to appear in the Results, and one to appear in the Discussion. The four recommendations for reporting adverse events in the Results section of the trial report are as follows:

- Describe participant withdrawals (or modification to the in-

tervention) in each arm of the trial that result from adverse events.
- State the denominators (total number of participants analyzed) used for analyses of adverse events.
- State the frequency of each adverse event in each arm of the trial (including information on the severity of the event, if relevant).
- Describe any subgroup analyses that are carried out to investigate adverse events.

In summary, this article focused on the CONSORT guidelines for reporting the Results section of trials. It is crucial that analyses and reporting of clinical trials adhere to high standards and are presented in an orderly manner. Doing so allows reliable conclusions on treatment efficacy and safety to be made, not only by the trial investigators but also by readers who will be interpreting the results from published reports. ■