

## CHAPTER 5

# Ambiguities and Confusions Between Reporting and Conduct

*Kenneth F. Schulz<sup>1</sup>, David Moher<sup>2</sup> and Douglas G. Altman<sup>3</sup>*

<sup>1</sup>FHI 360, Durham, and UNC School of Medicine, Chapel Hill, NC, USA

<sup>2</sup>Clinical Epidemiology Program, Ottawa Hospital Research Institute, Ottawa, ON, Canada

<sup>3</sup>Centre for Statistics in Medicine, University of Oxford, Oxford, UK

Theoretically, the reporting and conduct of research should be in harmony. Indeed, we argue that good reporting is an essential part of good conduct. In practice, however, while good reporting promotes enlightenment and clarity, poor reporting creates ambiguities and confusions. In addition, further confusion arises because a study can be excellently reported but poorly conducted or poorly reported but excellently conducted.

The primary goal of reporting guidelines is clarity, completeness, and transparency of reporting (see Chapter 1) [1, 2]. This allows readers to judge the validity of the methods and results, enabling them to develop an informed, enlightened interpretation and, for those interested researchers, to replicate the methods. Indeed, most reporting guidelines, such as the CONSORT 2010 Statement [3], do not incorporate requirements or recommendations for designing and conducting research. Reporting guidelines focus solely on a research report describing what was done and what was found [1, 2]. Obviously, reporting properly does not directly improve the design or conduct of a study, and adherence to a reporting guideline does not confer any stamp of quality on the research conduct.

Most developers of reporting guidelines embrace an intense interest in improving the design and conduct of health research. Nonetheless, they must relegate that interest to a secondary, indirect goal of their guideline. They hope that better design and conduct will materialize as a byproduct of proper reporting, which forces investigators who design and conduct deficient studies to divulge those deficiencies when they publish. Thus, good reporting in this instance provides enlightenment and clarity to readers while producing unease in authors. That unease, however, is encouraged by guideline developers as a constructive path to improve design and conduct in the future.

A further confusion between reporting and conduct emanates from the misuse of reporting guidelines. This misuse often takes the form of researchers using a guideline to develop a quality score for conduct of studies. Such a process is not advocated in most reporting guidelines, as stated in the CONSORT 2010 statement:

Moreover, the CONSORT 2010 statement does not include recommendations for designing and conducting randomized trials. The items should elicit clear pronouncements of how and what the authors did, but do not contain any judgments on how and what the authors should have done. Thus, CONSORT 2010 is not intended as an instrument to evaluate the quality of a trial. Nor is it appropriate to use the checklist to construct a "quality score." [2]

Still, some researchers have misused CONSORT as a basis for a quality score. CONSORT only states what to report but does not offer any judgment as to what is good and what is bad. Certainly, if authors report according to CONSORT, readers have the information needed to make their own judgments on the quality of design and conduct, but CONSORT does not indicate how they should make those judgments. CONSORT specifically, and reporting guidelines in general, were not conceived to serve as a springboard to a quality score. Moreover, the entire industry of quality scores is suspect [3]. This misuse of reporting guidelines creates confusion in the assessments of conduct.

## **Examples of item-specific scenarios**

Some individual items in reporting guidelines elicit clear, unambiguous information on conduct. For example, Item 1a in CONSORT 2010 asks for "identification as a randomized trial in the title."

However, other items can engender ambiguities and confusions between reporting and conduct that complicate the use of reporting guidelines. For most of us, those item-specific difficulties are hard to grasp abstractly, but examples are more easily understood.

## **Allocation concealment from the CONSORT 2010 statement**

As examples, we use conduct and reporting of an envelope method of allocation concealment in a randomization scheme, under different scenarios. As stated by some authors, an adequate envelope method would at a minimum use sequentially numbered, opaque, sealed envelopes

[4–6]. This represents a quality criterion for this item. If the description of allocation concealment reported in an article meets that quality criterion, a reader would likely judge the trial to be of good quality on the allocation concealment item in CONSORT 2010. With this as a baseline quality criterion, we present a few item-specific scenarios that serve as examples.

### **Scenario 1: Excellent reporting of excellent conduct**

In this scenario, the authors have harmonized reporting and conduct. A superb example is the following:

. . .The allocation sequence was concealed from the researcher (JR) enrolling and assessing participants in sequentially numbered, opaque, sealed and stapled envelopes. Aluminium foil inside the envelope was used to render the envelope impermeable to intense light. To prevent subversion of the allocation sequence, the name and date of birth of the participant was written on the envelope and a video tape made of the sealed envelope with participant details visible. Carbon paper inside the envelope transferred the information onto the allocation card inside the envelope and a second researcher (CC) later viewed video tapes to ensure envelopes were still sealed when participants' names were written on them. Corresponding envelopes were opened only after the enrolled participants completed all baseline assessments and it was time to allocate the intervention [7].

This description includes the quality criterion of sequentially numbered, opaque, sealed envelopes but also incorporates many other methods that greatly strengthen allocation concealment. This excellent reporting of excellent conduct leaves little room for confusion or ambiguity.

### **Scenario 2: Incomplete, ambiguous reporting of excellent conduct**

If investigators actually conducted a trial following the methods described in Scenario 1 but described their allocation approach as simply "sequentially numbered, opaque, sealed envelopes," that description would be true but incomplete. Moreover, the apparent reporting would be considered adequate reporting by many readers because it meets the sequentially numbered, opaque, sealed envelope quality criterion. Even so, the actual reporting is incomplete and ambiguous because it does not convey the true strength of this allocation concealment approach.

### **Scenario 3: Poor, ambiguous reporting of excellent conduct**

Investigators actually conducted a trial following the methods described in Scenario 1 but described their allocation approach as simply "sealed, opaque envelopes." That description would be true but distinctively incomplete. Generally, a reader would classify this description, based on the quality criterion, as ambiguous because the investigators might have only used sealed, opaque envelopes that were not sequentially numbered.

In actuality, just reporting the excellent conduct described in Scenario 1 as simply "sealed, opaque envelopes" is poor, incomplete reporting. Readers of articles face enormous ambiguity when they are confronted with that incomplete description because they have no concrete idea whether the conduct was poor or excellent.

**Scenario 4: Poor, ambiguous reporting of excellent conduct**

Investigators actually conducted a trial following the methods described in Scenario 1, but they failed to provide any details of their approach to allocation concealment, that is, no mention of envelopes, let alone adjectives. Generally, readers would consider this apparent reporting as unclear. This unclear, obscure reporting could represent anything from poor to excellent conduct. In this case, poor reporting camouflages excellent conduct.

**Scenario 5: Excellent reporting of inadequate conduct**

Assume that investigators used sealed, opaque envelopes that were not sequentially numbered, which would be considered inadequate conduct. If authors reported that they used "sealed, opaque envelopes" and explicitly stated that "the envelopes were not sequentially numbered," we would classify this description as excellent reporting of inadequate conduct. Owing to adequate, clear, and transparent reporting, no confusion or ambiguity would exist, even though the conduct was inadequate.

**Scenario 6: Ambiguous reporting of inadequate conduct**

Again assume that investigators used sealed, opaque envelopes that were not sequentially numbered. If authors reported that they simply used "sealed, opaque envelopes," this reporting is unclear because the authors may have used envelopes that were also sequentially numbered but just failed to report that fact. Scenarios 5 and 6 represent similar situations except that the authors in Scenario 5 have made it clear that sequentially numbered envelopes were not used. That led to the reporting being excellent in Scenario 5, whereas in Scenario 6 the reporting is unclear. Admittedly, these judgments are nuanced, but this illustrates the ambiguities between reporting and conduct. Whenever authors report simply "sealed, opaque envelopes" without clarifying information, ambiguity and confusion are created. Scenario 6 is, to a reader, indistinguishable from Scenario 3 above. Readers are confused whether this represents good reporting of inadequate conduct or poor reporting of adequate conduct.

**Scenario 7: Inadequate, fallacious reporting of inadequate conduct**

Authors reported that they used sequentially numbered, opaque, sealed envelopes, but they actually only used opaque, sealed envelopes that were not sequentially numbered. This form of fabricated, misleading reporting to give the appearance of having met some quality criterion (which in fact was not met) worries many of us who develop reporting guidelines. A quality criterion should be a goal by which investigators adjust conduct and not by which they adjust reporting. This scenario is particularly pernicious because it falsely describes inadequate conduct as adequate.

## Sequence generation from the CONSORT 2010 statement

As another example of less obvious ambiguities, we shall use conduct and reporting of a sequence generation method in a randomization scheme under different scenarios. As stated in some publications, an adequate sequence generation method would, at a minimum, use a specified method of randomization, such as a random number table or a reasonable computer random number generator [4–6]. This represents the quality criterion for this item. If that is reported in an article, a reader might judge the trial to be of good quality on the sequence generation item in CONSORT 2010. However, ambiguities and confusions can occur between reporting and conduct.

Assume that investigators used a computer random number generator to select random permuted blocks of size 6. In some instances, this might be considered adequate sequence generation, such as in trials where everyone is totally blinded [8, 9]. However, the trial in this example is unblinded (open-label), meaning everyone knows the treatment assigned to a participant after randomization. Thus, fixed block sizes may be discerned by the pattern of past assignments, with the obvious consequence being that some future assignments can be accurately anticipated [8, 9]. Scenarios 8 and 9 depict scenarios of reporting this form of inadequate conduct of sequence generation.

### Scenario 8: Poor, ambiguous reporting of inadequate conduct

Authors used the approach just discussed but reported their sequence generation approach as using a computer random number generator. In the absence of clarifying information, this description would be interpreted by many readers as implying a simple randomization scheme, which is adequate randomization to prevent bias under virtually all circumstances. By the quality criterion for sequence generation, many readers would judge this as apparently adequate reporting of adequate conduct. However, more astute readers, knowing this was an unblinded trial and that the authors did not explicitly state that simple randomization was used, might consider the reporting to be unclear. In any case, however, ambiguities exist between reporting and conduct. We would consider this actually poor (incomplete) reporting of inadequate conduct.

### Scenario 9: Excellent reporting of inadequate conduct

In this scenario, the actual conduct was the same as that in Scenario 8. However, in Scenario 9 the authors completely and transparently reported that they used a computer random number generator to select random permuted blocks of size 6. We view this as excellent reporting of inadequate conduct in this instance. Under this scenario, ambiguities and confusions dissipate.

## Discussion

Owing to the lack of complete and transparent reporting, ambiguities and confusions abound in the medical research literature. It is no wonder that peer reviewers and readers become uneasy and perplexed over manuscripts, whether submitted or already published. A medical epidemiologist friend terms this unease “dumbfounding.” Without adequate reporting, an ensuing knowledge abyss overwhelms readers, peer reviewers, and systematic reviewers. Indeed, for systematic reviewers, incomplete reports of studies enormously complicate the entire review process.

We hope our use of scenarios gives some clue as to the difficulties created by inadequate reporting. We rely on scenarios involving the CONSORT 2010 statement [2, 10] because it is the most familiar reporting guideline to readers and presents, arguably, the most complicated reporting issues. Nevertheless, analogous misuses and ambiguities arise from other reporting guidelines, such as STROBE [11].

What also should be clear from the scenarios presented is that ambiguities and confusions dissipate with adequate reporting. Granted, reporting guidelines may breed some unease in authors of poorly conducted studies, but this is constructive unease that encourages better study conduct in the longer term. Most importantly, complete, clear and transparent reporting eliminates ambiguities and confusions.

Reporting guidelines form part of the solution to inadequate reporting. Nevertheless, guidelines by themselves offer meager assistance if they are not used by authors, peer reviewers, and editors [12]. Additional emphasis on knowledge translation should unleash more of the benefits from existing reporting guidelines.

## References

- 1 Moher, D., Schulz, K.F., Simera, I. & Altman, D.G. (2010) Guidance for developers of health research reporting guidelines. *PLoS Medicine*, **7**, e1000217.
- 2 Schulz, K.F., Altman, D.G. & Moher, D. (2010) CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. *BMJ*, **340**, c332.
- 3 Jüni, P., Witschi, A., Bloch, R. & Egger, M. (1999) The hazards of scoring the quality of clinical trials for meta-analysis. *JAMA*, **282**, 1054–1060.
- 4 Altman, D.G. & Doré, C.J. (1990) Randomisation and baseline comparisons in clinical trials. *Lancet*, **335**, 149–153.
- 5 Schulz, K.F., Chalmers, I., Hayes, R.J. & Altman, D.G., (1995) Empirical evidence of bias. Dimensions of methodological quality associated with estimates of treatment effects in controlled trials. *JAMA*, **273**, 408–412.
- 6 Schulz, K.F., Chalmers, I., Grimes, D.A. & Altman, D.G. (1994) Assessing the quality of randomization from reports of controlled trials published in obstetrics and gynecology journals. *JAMA*, **272**, 125–128.

- 7 Radford, J.A., Landorf, K.B., Buchbinder, R. & Cook, C. (2006) Effectiveness of low-dye taping for the short-term treatment of plantar heel pain: a randomised trial. *BMC Musculoskeletal Disorders*, **7**, 64.
- 8 Schulz, K.F. & Grimes, D.A. (2002) Unequal group sizes in randomised trials: guarding against guessing. *Lancet*, **359**, 966–970.
- 9 Schulz, K.F. & Grimes, D.A. (2002) Generation of allocation sequences in randomised trials: chance, not choice. *Lancet*, **359**, 515–519.
- 10 Moher, D., Hopewell, S., Schulz, K.F. *et al.* (2010) CONSORT 2010 Explanation and Elaboration: updated guidelines for reporting parallel group randomised trials. *Journal of Clinical Epidemiology*, **63**, e1–e37.
- 11 da Costa, B.R., Cevallos, M., Altman, D.G., *et al.* (2011) Uses and misuses of the STROBE statement: bibliographic study. *BMJ Open*, **1**, e000048.
- 12 Simera, I., Moher, D., Hirst, A., *et al.* (2010) Transparent and accurate reporting increases reliability, utility, and impact of your research: reporting guidelines and the EQUATOR Network. *BMC Medicine*, **8**, 24.