Item 20. Trial limitations, addressing sources of potential bias, imprecision, and, if relevant, multiplicity of analyses

Example—“The preponderance of male patients (85%) is a limitation of our study … We used bare-metal stents, since drug-eluting stents were not available until late during accrual. Although the latter factor may be perceived as a limitation, published data indicate no benefit (either short term or long-term) with respect to death and myocardial infarction in patients with stable coronary artery disease who receive drug-eluting stents, as compared with those who receive bare-metal stents.”

Explanation—The discussion sections of scientific reports are often filled with rhetoric supporting the authors’ findings and provide little measured argument of the pros and cons of the study and its results. Some journals have attempted to remedy this problem by encouraging more structure to authors’ discussion of their results. For example, Annals of Internal Medicine recommends that authors structure the discussion section by presenting (1) a brief synopsis of the key findings, (2) consideration of possible mechanisms and explanations, (3) comparison with relevant findings from other published studies (whenever possible including a systematic review combining the results of the current study with the results of all previous relevant studies), (4) limitations of the present study (and methods used to minimise and compensate for those limitations), and (5) a brief section that summarises the clinical and research implications of the work, as appropriate. We recommend that authors follow these sensible suggestions, perhaps also using suitable subheadings in the discussion section. Although discussion of limitations is frequently omitted from research reports, identification and discussion of the weaknesses of a study have particular importance. For example, a surgical group reported that laparoscopic cholecystectomy, a technically difficult procedure, had significantly lower rates of complications than the more traditional open cholecystectomy for management of acute cholecystitis. However, the authors failed to discuss an obvious bias in their results. The study investigators had completed all the laparoscopic cholecystectomies, whereas 80% of the open cholecystectomies had been completed by trainees. Authors should also discuss any imprecision of the results. Imprecision may arise in connection with several aspects of a study, including measurement of a primary outcome (see item 6a) or diagnosis (see item 4a). Perhaps the scale used was validated on an adult population but used in a paediatric one, or the assessor was not trained in how to administer the instrument. The difference between statistical significance and clinical importance should always be borne in mind. Authors should particularly avoid the common error of interpreting a non-significant result as indicating equivalence of interventions. The confidence interval (see item 17a) provides valuable insight into whether the trial result is compatible with a clinically important effect, regardless of the P value. Authors should exercise special care when evaluating the results of trials with multiple comparisons. Such multiplicity arises from several interventions, outcome measures, time points, subgroup analyses, and other factors. In such circumstances, some statistically significant findings are likely to result from chance alone.