17. Other analyses: Report other analyses done—e.g., analyses of subgroups and interactions, and sensitivity analyses.

Explanation

In addition to the main analysis other analyses are often done in observational studies. They may address specific subgroups, the potential interaction between risk factors, the calculation of attributable risks, or use alternative definitions of study variables in sensitivity analyses. There is debate about the dangers associated with subgroup analyses, and multiplicity of analyses in general . In our opinion, there is too great a tendency to look for evidence of subgroup-specific associations, or effect measure modification, when overall results appear to suggest little or no effect. On the other hand, there is value in exploring whether an overall association appears consistent across several, preferably pre-specified subgroups especially when a study is large enough to have sufficient data in each subgroup. A second area of debate is about interesting subgroups that arose during the data analysis. They might be important findings, but might also arise by chance. Some argue that it is neither possible nor necessary to inform the reader about all subgroup analyses done as future analyses of other data will tell to what extent the early exciting findings stand the test of time. We advise authors to report which analyses were planned, and which were not (see also items 4, 12b and 20). This will allow readers to judge the implications of multiplicity, taking into account the study’s position on the continuum from discovery to verification or refutation. A third area of debate is how joint effects and interactions between risk factors should be evaluated: on additive or multiplicative scales, or should the scale be determined by the statistical model that fits best (see also item 12b and Box 8)? A sensible approach is to report the separate effect of each exposure as well as the joint effect—if possible in a table, as in the first example above , or in the study by Martinelli et al. Such a table gives the reader sufficient information to evaluate additive as well as multiplicative interaction (how these calculations are done is shown in Box 8). Confidence intervals for separate and joint effects may help the reader to judge the strength of the data. In addition, confidence intervals around measures of interaction, such as the Relative Excess Risk from Interaction (RERI) relate to tests of interaction or homogeneity tests. One recurrent problem is that authors use comparisons of P-values across subgroups, which lead to erroneous claims about an effect modifier. For instance, a statistically significant association in one category (e.g., men), but not in the other (e.g., women) does not in itself provide evidence of effect modification. Similarly, the confidence intervals for each point estimate are sometimes inappropriately used to infer that there is no interaction when intervals overlap. A more valid inference is achieved by directly evaluating whether the magnitude of an association differs across subgroups. Sensitivity analyses are helpful to investigate the influence of choices made in the statistical analysis, or to investigate the robustness of the findings to missing data or possible biases (see also item 12b). Judgement is needed regarding the level of reporting of such analyses. If many sensitivity analyses were performed, it may be impractical to present detailed findings for them all. It may sometimes be sufficient to report that sensitivity analyses were carried out and that they were consistent with the main results presented. Detailed presentation is more appropriate if the issue investigated is of major concern, or if effect estimates vary considerably . Pocock and colleagues found that 43 out of 73 articles reporting observational studies contained subgroup analyses. The majority claimed differences across groups but only eight articles reported a formal evaluation of interaction (see item 12b)