10. Study size: Explain how the study size was arrived at.

**Example 1**

“The number of cases in the area during the study period determined the sample size.”[73](http://journals.lww.com/epidem/Fulltext/2007/11000/Strengthening_the_Reporting_of_Observational.28.aspx#R73-28)

**Example 2**

“A survey of postnatal depression in the region had documented a prevalence of 19.8%. Assuming depression in mothers with normal weight children to be 20% and an odds ratio of 3 for depression in mothers with a malnourished child we needed 72 case-control sets (one case to one control) with an 80% power and 5% significance.”[74](http://journals.lww.com/epidem/Fulltext/2007/11000/Strengthening_the_Reporting_of_Observational.28.aspx#R74-28)

**Explanation**

A study should be large enough to obtain a point estimate with a sufficiently narrow confidence interval to meaningfully answer a research question. Large samples are needed to distinguish a small association from no association. Small studies often provide valuable information, but wide confidence intervals may indicate that they contribute less to current knowledge in comparison with studies providing estimates with narrower confidence intervals. Also, small studies that show ‘interesting’ or ‘statistically significant’ associations are published more frequently than small studies that do not have ‘significant’ findings. While these studies may provide an early signal in the context of discovery, readers should be informed of their potential weaknesses.

The importance of sample size determination in observational studies depends on the context. If an analysis is performed on data that were already available for other purposes, the main question is whether the analysis of the data will produce results with sufficient statistical precision to contribute substantially to the literature, and sample size considerations will be informal. Formal, a priori calculation of sample size may be useful when planning a new study. Such calculations are associated with more uncertainty than implied by the single number that is generally produced. For example, estimates of the rate of the event of interest or other assumptions central to calculations are commonly imprecise, if not guesswork. The precision obtained in the final analysis can often not be determined beforehand because it will be reduced by inclusion of confounding variables in multivariable analyses, the degree of precision with which key variables can be measured, and the exclusion of some individuals.

Few epidemiological studies explain or report deliberations about sample size. We encourage investigators to report pertinent formal sample size calculations if they were done. In other situations they should indicate the considerations that determined the study size (eg, a fixed available sample, as in the first example above). If the observational study was stopped early when statistical significance was achieved, readers should be told. Do not bother readers with post hoc justifications for study size or retrospective power calculations.[77](http://journals.lww.com/epidem/Fulltext/2007/11000/Strengthening_the_Reporting_of_Observational.28.aspx#R77-28) From the point of view of the reader, confidence intervals indicate the statistical precision that was ultimately obtained. It should be realized that confidence intervals reflect statistical uncertainty only, and not all uncertainty that may be present in a study (see STROBE item 20).