BOX 3

Evaluating the quality of epidemiological studies

Epidemiological findings are only as good as the studies that produce them. To assess the quality of epidemiological studies, readers should ask many critical questions about their design and execution, including the following:

Intervention trials

- Were the subjects assigned randomly to the treatment and control groups?
- Was randomization successful (that is, were the treatment and control groups truly comparable with respect to important variables)?
- Were efforts made to determine whether subjects complied with the treatment protocol?
- Was compliance good?
- Were subjects kept "blind" to their treatment assignment through the use of a placebo or other means?
- Was blinding successful (that is, were subjects unable to guess their group assignment correctly)?
- Were the researchers kept "blind" as to the subjects' treatment assignment?
- Was the assessment of outcome blinded?
- Was the number of subjects large enough to yield statistically reliable results?

Cohort studies and nested case-control studies[1]

- Was the number of subjects large enough to allow effects to be detected?
- Was the follow-up period long enough to yield meaningful results?
- Did the investigators clearly define the reference population from which the cohort was drawn?
- Were exposures and outcomes assessed by accurate and appropriate methods?
- Were exposure assessments repeated during the study to account for possible changes over time?
- Were many subjects lost to follow-up? If so, could this have affected the study outcome?
- Were data collected on potential confounding variables? Were these variables taken into account in the data analysis?

BOX 3 – continued

- Did sufficient time elapse between the initial examination and the diagnosis of disease to ensure that metabolic effects of disease could not have influenced biological variables measured in the study?
- How were the diagnoses confirmed in the subjects who developed the disease under investigation?
- If biological samples were collected at the outset of the study and stored for later analysis, were suitable storage conditions used? Were all samples handled in the same manner?

Case-control studies[1]

- Was the number of subjects large enough to allow effects to be detected?
- How was the diagnosis confirmed in the cases?
- Were exposures and outcomes assessed by accurate and appropriate methods?
- Were the rates of participation in the study (by both potential cases and potential controls) high? If not, could this have influenced the study outcome?
- Were controls selected from the same population that yielded the cases?
- Were controls comparable to cases in all respects other than those under investigation?
- Was information collected in the same way from cases and controls?
- Did the subjects themselves provide the exposure data, or was it necessary to resort to proxy respondents?
- Did the assessment cover an appropriate time frame?
- Were interviewers blinded to the subjects' case or control status?
- Did the researchers consider the possibility that metabolic effects of disease could have influenced measurements such as blood pressure or blood nutrient levels?
- Was the purpose of the study defined beforehand rather than after the data were collected?

Studies (of any design) that showed a statistical association between an exposure and an outcome

- Could the association have been due to chance?
- Could the association have been due to bias?
- Could the association have been due to confounding?
- Is the association biologically plausible?
- To whom does the association apply?
- Is the association likely to represent a cause-and-effect relationship? (See Table 2 for a further discussion of criteria for assessing the likelihood of causality.)

Studies (of any design) that did not show a statistical association between an exposure and an outcome[2]

- Did the study have the power to detect a clinically or biologically significant effect if it were present?

BOX 3 – continued

Reviews and meta-analyses[3]

- Were the question(s) and methods clearly stated?
- Were comprehensive search methods used to locate relevant studies?
- Were explicit and appropriate methods used to determine which papers to include in the review?
- Was the methodologic quality of the primary studies assessed?
- Were the selection and assessment of the primary studies reproducible and free from bias?
- Were differences in individual study results adequately explained?
- Were the results of the primary studies combined appropriately?
- Were the reviewers' conclusions supported by the data cited?

1. Recently, a group of epidemiologists developed a quantitative scoring system to judge the scientific quality of case-control and cohort studies of nutrition and disease. In this system, case-control studies are scored in three areas (dietary assessment, recruitment of subjects and analysis), and cohort studies are scored in four areas (dietary assessment, definition of cohort, ascertainment and analysis). For more information on this scoring system and its application see Margetts BM, Thompson RL, Key T, et al, Development of a scoring system to judge the scientific quality of information from case-control and cohort studies of nutrition and disease. Nutrition and Cancer 24:231–239, 1995.
2. For more information see Friedman GD, Primer of Epidemiology, 4th ed. (© New York:McGraw-Hill, 1994), p321.
3. For more information see Sackett DL, Haynes RB, Guyatt GH, Tugwell P, Clinical Epidemiology: A Basic Science for Clinical Medicine, 2nd ed. (© Boston: Little, Brown & Company, 1991), p380.