How to assess epidemiological studies

J H Zaccai

Assessing the quality of an epidemiological study equates to assessing whether the inferences drawn from it are warranted when account is taken of the methods, the representativeness of the study sample, and the nature of the population from which it is drawn. Bias, confounding, and chance can threaten the quality of an epidemiological study at all its phases. Nevertheless, their presence does not necessarily imply that a study should be disregarded. The reader must first balance any of these threats or missing information with their potential impact on the conclusions of the report.

Epidemiology underpins good clinical research. It is any research with a defined numerator, which describes, quantifies, and postulates causal mechanisms for health phenomena. Epidemiology gives insight into the natural history and causes of disease and can provide evidence to help prevent occurrence of disease. It promotes effective treatments either to cure or to prolong the lives of those with disease. Epidemiology, also referred to as "population medicine", is used to estimate the individual risk of disease and the chances of avoiding it from group experience averages. Such information is crucial to planning interventions and allocating resources.

The epidemiological approach needs to be applied to clinical research to evaluate both its effectiveness and its importance. Hence clinicians need to gain the skills that will allow them to properly update and re-evaluate their knowledge and thus provide the best evidence based patient care. Epidemiology is an interdisciplinary field that draws its techniques and methodologies from biostatistics, social sciences, and clinical medicine as well as from a vast range of biological sciences such as genetics, toxicology, and pathology and for this reason the interpretation of epidemiological studies is not always easy.

There are several reviews and books available that provide advice on how best to assess epidemiological studies. The favoured outline for these is by listing types of common errors. This review provides an alternative approach that is hoped will be helpful. After briefly characterising the main threats to the quality of epidemiological studies, a map is provided to assess studies based on their usual format—that is, the design, conduct, and analysis of the results. Readers of epidemiology papers at any level will be assisted in their task by Last's A Dictionary of Epidemiology, an essential guide to all. Assessing the quality of epidemiological studies equates to assessing their validity.

CONCEPT OF VALIDITY AND ITS THREATS

According to Last, validity is the "degree to which the inference drawn from a study is warranted when account is taken of the study methods, the representativeness of the study sample, and the nature of the population from which it is drawn." The concept of validity was further developed in the 1950s by Campbell when he introduced the distinction between external and internal validity:

- Internal validity is the extent to which systematic error is minimised during all stages of data collection.
- External validity is the extent to which results of trials provide a correct basis for generalisation to other circumstances; this is regarding patients, treatment regimens, setting, modalities of outcome, which include definition of outcomes and duration of follow up.

Every step in a study should be undertaken in such a way as to maximise its validity. There are three threats to validity: bias, confounding, and chance.

Bias

Bias is a systematic error. Sackett has listed dozens of biases that can distort the estimation of an epidemiological measure. The distinction among these is occasionally difficult to discern but there are two general types of bias that should be remembered: selection bias and information bias. Selection bias is error due to systematic differences in characteristics between those who take part in a study and those who do not. Information bias, also called measurement bias, is systematic error arising from inaccurate measurement (or classification) of subjects on study variable(s). Measurement bias can arise from the choice of tools one uses to measure as well as the assessor’s attitude and the cooperation of the participant, if it is a human based study.

Bias in studies does not necessarily mean that they become scientifically unacceptable and should be disregarded. A first step must be to assess the probable impact of the described biases on study results—that is, the direction in which each bias is likely to affect outcome, and its magnitude. The magnitude should not be so great that the results are changed to make the relationship stronger or weaker than that observed. Unfortunately, there is no simple formula for assessing biases: each must be
considered on its own merit in the context of the study population.

Confounding
Confounding is a type of bias but it is often considered as its own entity. According to Last:

"Confounding bias is a distortion of the estimated effect of an exposure on an outcome, caused by the presence of an extraneous factor associated both with the exposure and the outcome, that is, confounding is caused by a variable that is a risk factor for the outcome among non-exposed persons and is associated with the exposure of interest, but is not an intermediate step in the control pathway between exposure and outcome".

Confounding is illustrated in fig 1. Another way of viewing confounding is as a confusion of effects. The distortion introduced by a confounding factor can be large and it can lead to overestimation or underestimation of an effect, depending on the direction of the associations that the confounding factor has with exposure and disease. Confounding can even change the apparent direction of an effect.

Methods to prevent confounding include randomisation, restriction, and matching. Random allocation, not to be confused with haphazard assignment, can be used in trials. It follows a predetermined plan and aims, within the limits of chance variation, to make the control and experimental groups similar at the start of an investigation, thus minimising any unbalanced relationship between known and unknown confounders and other studied variables. This is because confounding cannot occur if potential confounding factors do not vary across groups. In a similar manner, restriction and matching also try to make the study group and comparison group comparable with respect to extraneous factors but this time by specifically selecting subjects according to their “confounder-bearing” status. For instance, continuing with the example above, the study groups could be chosen in such a way as to include only non-smokers or only smokers. Confounding can also be adjusted for during the statistical analysis phase of the study with stratified analysis and multivariate analysis techniques. Stratification is a technique that involves the evaluation of association between the exposure and disease within homogeneous categories or strata of the confounding variable. The results from the above study can be analysed according to smoking history: never smoker, ex-smoker 10+ years, ex-smoker <10 years, current smoker. Multivariate analysis involves the construction of a mathematical model and allows for the efficient estimation of measures of association while controlling for a number of confounding factors simultaneously, even in situations where stratification would fail because of insufficient numbers.

The reader can thus assess confounding by considering whether any important factors have not been taken into account in the design and/or analysis phase of a study, based on an understanding of the natural history of a disease.

**CHANCE**
Inevitably, because studies cannot include entire populations and continue indefinitely in time, some chance factor may result in study outcomes not representing the ultimate true values, even if bias and confounding are non-existent. Investigators adjust for the chance factor using statistics in the analysis phase of the study. However, variations from the true values will be minimised the larger and/or the longer the study.

**ASSESSING THE DESIGN AND CONDUCT OF AN EPIDEMIOLOGICAL STUDY**

**Choice of study design**
Which study design was chosen and was it appropriate?
Researchers have a choice of several study designs for their investigation and a judgment must be made as to whether their choice is reasonable in relation to the question they wish to consider. Table 1 lists epidemiological study designs and specific goals these can help achieve.

The more appropriate the study design, the more convincing the evidence that will be produced. Conclusions from a case-control study assessing the efficacy of a surgical procedure will be stronger than that of an observational cohort study and will be weaker than that of a well conducted randomised controlled trial.

The reader must beware not to accept what the study claims to be without going through the description of its design. Particularly interventional studies that are described as randomised controlled trials do not always stand up to careful scrutiny. This may be because there is pressure to overclaim the design of a study considered to be the gold standard in epidemiological investigations, which is difficult to conduct in a valid way.

**Choice of study population**
Has the population been sufficiently described?
It is important that researchers report the sociodemographic characteristics of the study population to allow readers to see the possibilities of generalisation to other populations. Furthermore, it allows physicians to judge whether they can apply the results to particular patients. In some instances of case-control studies and trials, the description...
of the group allows assessment for selection bias—that is, differences in the two groups at baseline, which may account for effects observed in the analysis phase. This assessment must also be done even in randomised trials where systematic bias is eliminated. Randomisation does not necessarily produce perfectly balanced groups with respect to prognostic factors and differences due to chance remain in the intervention groups. Assessment of selection bias is crucial and if identified will need to be controlled for during the analysis phase, although in some examples this will not be possible.

What is the source population?

The source of the population is known to have an impact on the conclusions of a study. For example, selection bias introduced by referral of patients from care centres can affect profoundly the results of clinical and epidemiological studies. This is because referral is influenced by more than the severity of the disorder itself and has much to do with the way that communities contain and deal with aberrant behaviour. Referral may differ according to burden of symptoms, access to care, popularity of disorders and institutions. Another example is with participants recruited via the media. Those who volunteer to participate are likely to differ from non-participants in a number of important ways, including basic levels of motivation and attitudes towards health. As a further example, the readers should judge whether recruiting via telephone or door knocking or whether incentives were given to take part in a study will affect the final results, and if so, in what direction.

How were the participants selected?

The inclusion and exclusion criteria for subject participation and the ways in which they were applied must be clearly defined. This is to show minimum tampering of subject participation by researchers. A common error is defining studies as population based. However, as long as participants have not been recruited from all subgroups of a population, one cannot consider the study to be community based. For example, solely recruiting from health registries would only be acceptable in a country where health care is universal and free. Another important source of effect on the outcome is whether subclinical cases have been included. The readers must always consider how non-included members may affect the results of the study.
Have the investigators strived for high participation rates?
Investigators must strive for high participation rates. If, for example, the researcher contacts an initial target population and manages to recruit 65% to take part in his/her study, one must assess whether these 65% are representative of the initial population. In addition, the investigator must assess whether the numbers recruited are large enough to make statistically viable conclusions. As mentioned earlier, the role of chance on results can be minimised and the generalisability can be maximised in larger and/or longer trials.

Has attrition been high enough to change the main characteristics of the study and control groups?
In the same manner, any attrition or loss to follow up should be reported with an attempt to explain what differences this makes to conclusions.

Has there been any participant exclusion after recruitment?
Exclusion numbers should be reported. Exclusion is acceptable if study personnel made errors in the implementation of eligibility criteria or if patients never received the intervention in an experimental study. However, in no circumstance should exclusion be accepted if it appears to be dependent on the treatment given. In trials, post-randomisation exclusion acceptability really depends on whether the goal of study is to address an explanatory (efficacy) or management (effectiveness) question. Not excluding participants who did not follow their intended treatment will allow an answer to an effectiveness investigation on an intention to treat basis. Only 13% of all randomised trials published in the New Zealand Medical Journal between 1943 and 1995 provided evidence that final analyses were conducted on an intention to treat basis. Investigators should clearly state the number of patients recruited but not included in the primary analysis of data and explain the circumstances under which such patients were enrolled but excluded from the analysis.

Which comparison group?
Any differences between the exposed and control group during the study should be assessed in relation to their potential effect on outcomes observed. Unless this is done adequately, any analysis will be dangerously misleading.

Some investigators feel that the closer the identity of the compared groups with respect to all measurable factors, the greater the validity, since some factors may affect disease incidence without the investigator’s awareness. Matching unexposed to exposed subjects in cohort studies can prevent confounding of the crude risk difference and ratio because such matching prevents the association between exposure and the matching factors among the study subjects at the start of the follow up. Matching in cohort studies though is rarely done. In practice much of the controlling in cohort studies occurs in the analysis phase where complex statistical adjustment is made for baseline differences in key variables. Matching in case-control studies may introduce bias and thus matching on a factor may still necessitate its control in the analysis phase. If controls are selected to match cases on a factor that is correlated with the exposure, than the crude exposure frequency in controls will be distorted in the direction of similarity to that of the cases, creating a risk of over matching.

The choice of comparison groups can also introduce error in experimental studies. For example in a meta-analysis showing that research sponsored by the drug industry was more likely to produce results favouring the product made by the company sponsoring the research than studies funded by other sources, it was shown that this might be due to inappropriate comparators or publication bias rather than the reported quality of methods. It was found that in trials of psychiatric drugs, the comparator drug is often given in doses outside the usual range. Similarly, research funded by the company marketing fluconazole compared it with the oral amphoterinc B, a drug known to be poorly absorbed, thereby creating a bias in favour of fluconazole.

Often the comparison is a placebo controlled group, meaning that the control participants were given an inert medication or procedure that is intended to give them the perception that they are receiving treatment for their complaint. This is thought to control for the power of suggestion by a medical adviser. Hrobjartsson and Gotzsche investigated patient reported and observer outcomes and found no evidence that placebo interventions in general have clinically important effects, except possibly on subjective continuous outcomes, such as pain, where the effect could not be clearly distinguished from bias. The placebo effect can thus help compare the validity of the methods of investigation in experimental studies. In a review of trials looking at the treatment of irritable bowel syndrome (IBS), the placebo response was extremely variable and high, most frequently between 40% and 70%. Differences of this magnitude reflect not only the nature of the patients enrolled in a trial but also the methods used to determine treatment response. It is a useful way to compare methods and results across studies.

If necessary, has the method of randomisation and allocation concealment been reported?
The non-reporting of the method of randomisation and allocation concealment is one of the main errors in articles reporting randomised trials. For example, a review reported that the mechanism used to allocate interventions was omitted in reports of 93% of trials in dermatology, 89% of trials in rheumatoid arthritis, 48% of trials in obstetrics and gynaecology journals, and 45% of trials in general medical journals. Unless stated clearly in the paper, one cannot be assured that randomisation was correctly done. Correct randomisation is dependent on proper allocation concealment—that is, random allocation without foreknowledge of treatment assignments. Methods of concealment include sequentially numbered, opaque, sealed envelopes or containers, can be pharmacy controlled, or completed by central randomisation. However, each may not be sufficient. Elements convincing of concealment must be reported in the study paper. This is crucial as results of four empirical investigations reported by Schulz and Grimes have shown that trials that used inadequate or unclear allocation concealment compared with those that used adequate concealment, yielded up to 40% larger estimates of effect.

Choice of exposure and outcome measures
One major source of error in studies, especially in cohorts, is in the degree of accuracy with which respondents have been classified with respect to their exposure and disease status—that is, measurement bias. Choosing what and how measurements will be collected, whether it be exposure, outcome and other auxiliary variables, determines the validity of the study. If the mis-measurement is random, the misclassification of a dichotomous exposure is always in the direction of the null value. Although it is generally considered acceptable to underestimate effects rather than overestimate them, this type of error may account for some discrepancies amongst studies.

Has potential bias from the choice of tools for data collection been dealt with?
Two types of data can be used for epidemiological studies: routine data and data which have been collected specifically

www.postgradmedj.com
for the study. Creating new knowledge versus using routine
data has a great impact on any study. Routine data have the
advantage of being collected independently of the study and
thus an automatic blinding of assessors is in place. However,
routine data are often incomplete and not necessarily
appropriate for answering the study question.

There are many tools for collecting data. These include
open group discussions, self rating, direct examination
interviews, and biological marker measurement. Data should
be collected in as objective, reliable, accurate, and reproduc-
ibly fashion as possible. Different data collection methods
are prone to different errors of measurement. Hence the use
of well recognised standards or validated tools is a positive
point. Validity here is an expression of the degree to which a
measurement measures what it purports to measure.1

Correct case classification can involve varying effort. For
example, the clinical diagnosis of Alzheimer’s disease is one
of exclusion. Cerebrospinal fluid and blood analyses and
imaging are used to differentiate Alzheimer’s disease from
other illnesses that may cause the same clinical symptoms.
Possibly, the more the tests carried out, the less likely a
participant would be classified as having Alzheimer’s disease.

Has enough or too much information been collected?
Correct case classification can involve varying effort. For
example, the clinical diagnosis of Alzheimer’s disease is one
of exclusion. Cerebrospinal fluid and blood analyses and
imaging are used to differentiate Alzheimer’s disease from
other illnesses that may cause the same clinical symptoms.
Possibly, the more the tests carried out, the less likely a
participant would be classified as having Alzheimer’s disease.

Has potential bias from the participants been dealt with?
The use of standardised questionnaires or laboratory proto-
cols does not always prevent observer variation. Discrepancies
between repeated observations by the same observer and between
different observers are to be expected.1

This variation is measured by the kappa factor, which allows
for chance association. Reporting of kappa values shows a
will for validity by the investigators. The higher the factor,
the higher the concordance is between measurements.

Incorrect recording of the results. Misinterpretation of data
can be due to the pre-judgment and expectancy of what
results should be. This highlights the importance of ‘blind-
ing’ the measurer to the probable caseness of the measured
subject and of observing quality controls in carefully agreed
protocols. Although often considered free of bias, molecular
work is not immune to measurement bias. For example,
while comparing tangle determination in the CERAD protocol
for neuropathologically diagnosing Alzheimer’s disease,
Mirra and co-workers found that only 66% of raters from 15
laboratories showed internal consistency.15 It is difficult to
assess whether low inter-rater/intrarater reliability can have
an effect other than random on the results. However, a
minimum aim is to report this reliability for readers to assess
the validity of the results.

Has potential bias from the participants been dealt with?
Bias can result from inaccurate reporting by participants. This
is particularly so in case-control studies as the information on
exposure is often provided by the participant after the onset
of disease. Recall bias can occur when cases differ with
respect to their exposure response due to the disease
experience relative to controls. For instance, those who have
suffered from food poisoning may remember their meals
differently from those who did not suffer similarly. The use
of memory aids can help reduce recall bias.

There are also circumstances where participants perceive
social pressures to report fittingly. This is especially so when
dealing with self report on drinking,9 smoking,20,21 drug
taking, and sexual habits.22 For example, the reader should
judge how self report over the telephone to monitor
overweight and obesity in populations can be affected by
social desirability. It was found that body mass index, based
on measured weights and heights, classified 62% of males
and 47% of females as overweight or obese, compared with
39% and 32% respectively from self report.23 Blinding of the
participants to study goals and participants’ classification
status to any interests may help.

Box 3: Study design and conduct

• The main aspects of the study design and conduct that
need to be assessed include: choice of study design,
choice of study population, and the choice of exposure
and outcome measures.
the degree of variability in the factor being investigated and
the limited size of the study: the wider the confidence
intervals, the less powerful the study is.

Often, a p value under or equal to a probability of 1 in 20 or
0.05 is considered statistically significant, however, signifi-
cance does not mean that the results make biological sense.
Results can be statistically significant without being biologi-
cally/sociologically significant. For example, a very large
clinical trial can provide a significant result on the effect of a
specific drug that increases the concentration of haemoglobin
by 1 g/100 ml blood. The readers should consider whether
this is plausible and whether this can have a useful medical
effect.

Often the need for large sample sizes to achieve sufficient
power and thus precision to answer study hypotheses can
lead to combination of broad categories of cases. This can
cause heterogeneity in the cases groups, which can be
inappropriate. This happens in cohort studies and the result
is that it can obscure effects on more narrowly defined
diseases. However, such non-differential misclassification of
exposure, even if substantial, only underestimates associa-
tions, provided that the misclassification probabilities apply
uniformly to all subjects.

**Have the results shown a cause-effect relationship?**

Showing that an exposure is strongly associated with a
disease does not necessarily imply that there is a cause-effect
relationship. Hill described a series of conditions, which if
completed will prove a cause-effect relationship.26 These are:

- A sufficient strength of association.
- A temporal relationship between exposure and outcome.
- A dose-response relationship.
- Consistency.
- Biological plausibility.
- Coherence.
- Specificity.
- Analogy.

Temporality is particularly difficult to demonstrate in case-
control studies where all data are collected at once. Table 2
shows these criteria illustrated using the cause effect
relationship described for the human papillomavirus and
cervical cancer.27

**What are the policy implications?**

The final phase of assessing epidemiological studies is
determining whether it has any policy implications. Although a
consistency and magnitude of effect can be
demonstrated, the impact of any intervention must also be
considered. This is also known as the generalisability of the
results and is directly dependent on the study participants’
characteristics.

To assess the impact of an intervention, the reader should
also think in terms of attributable risk rather than relative
risk. Attributable risk is the proportion of a disease or other
outcome in exposed individuals that can be attributed to the
exposure. This measure is derived by subtracting the rate of

---

**Table 2** Example of the cause effect relationship with the
human papillomavirus (HPV) and cervical cancer
(adapted from Bosch et al27)

<table>
<thead>
<tr>
<th>Criterion</th>
<th>HPV-cervical cancer example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sufficient strength of assoc.</td>
<td>The association between HP DNA in cervical specimens and cervical cancer is one of the strongest ever observed for a human cancer. HPV-16 accounts for almost 50% of the types identified in cervical cancer. The cancer risk for any one of at least 10 HPV types or for any combination of HPV types does not differ significantly</td>
</tr>
<tr>
<td>Temporal relationship between exposure and outcome</td>
<td>HPV infections precede cervical precancerous lesions and cervical cancer by a substantial number of years. The epidemiology and the dynamics of HPV infection in populations satisfy previous observations that related cervical cancer to a sexually transmitted disease</td>
</tr>
<tr>
<td>Dose-response relationship</td>
<td>The risk of cervical cancer may be related to estimates of viral load. The technology to estimate viral load is being developed and compliance with the biological gradient requirement needs to be further validated</td>
</tr>
<tr>
<td>Consistency</td>
<td>The association between HPV DNA in cervical specimens and cervical cancer is consistent in a large number of investigations in different countries and populations. There are no published studies with observations challenging the central hypothesis on causality</td>
</tr>
<tr>
<td>Biological plausibility and coherence</td>
<td>The association of HPV DNA in cervical specimens and cervical cancer is plausible and coherent with previous knowledge. This includes in vitro experiments and observations in humans. Novel criteria of causality are being proposed and tested as molecular technology develops and is introduced into epidemiological research protocols</td>
</tr>
<tr>
<td>Specificity</td>
<td>The association of type specific HPV DNA and cervical cancer is significantly different from random. Systematic patterns of HPV type and cervical cancer histology suggest a fair degree of specificity. Patterns are also observed when the scope of HPV and cancer expands to include the full spectrum of HPV types and the large number of additional cancer sites that have been investigated</td>
</tr>
<tr>
<td>Analogy</td>
<td>The HPV and cervical cancer model is analogous to many other examples of papillomavirus induced papillomas and carcinomas and cancers caused by other viruses</td>
</tr>
</tbody>
</table>

**Table 3** Relative and attributable risks of mortality from lung cancer and coronary heart disease among cigarette smokers in a cohort study of British male physicians (adapted from Doll and Peto28)

<table>
<thead>
<tr>
<th></th>
<th>Lung cancer</th>
<th>Coronary heart disease</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cigarette smokers</td>
<td>140</td>
<td>669</td>
</tr>
<tr>
<td>Non-smokers</td>
<td>10.0</td>
<td>413</td>
</tr>
<tr>
<td>Relative risk</td>
<td>14.0</td>
<td>1.6</td>
</tr>
<tr>
<td>Attributable risk</td>
<td>130/10^5/year</td>
<td>256/10^5/year</td>
</tr>
</tbody>
</table>

---

**Box 4: Aspects of study analysis to be assessed**

- Aspects of the study analysis phase which need to be assessed are the statistical and biological interpretation of the results, the generalisability of the findings, and whether they show a cause-effect relationship between the factors under investigation.
the outcome (usually incidence or mortality) among the unexposed from the rate among the exposed individuals. It is assumed that causes other than the one under investigation have had equal effects on the exposed and unexposed groups. This is different to the relative risk, which is the ratio of the risk of disease or death among the exposed to the risk among the unexposed. The relative risk provides information that can be used in making a judgment of causality. However, once causality is assumed, from the perspective of public health policy making, measures of association based on absolute differences in risk between exposed and non-exposed individuals assume far greater importance. This is illustrated with the example in table 3.

**CONCLUSION**

There are many subjective elements to the interpretation of epidemiological studies; however, minimum standards in the conduct of a study ensure that any conclusion reached is appropriate. The reader must bear in mind that assessing an epidemiological study not only implies knowing how to look for key information in its paper but also in its “comments” and “corrections”. These are listed along with the paper reference in Medline. Bias, confounding, and chance can threaten the validity of a study at all its stages. Thus, the methodology must be well thought-out and this must be reflected in the study paper. It is understood that all the details about choices made by investigators cannot be published; nevertheless, the printed information should provide sufficient details so as to rule out alternative interpretations of the results. Investigators must show that they planned to minimise bias and account for confounding while also describing statistical methods. More importantly though, they must report any potential impact of limitations on the results found. Many reviewers when assessing study validity take a “guilty until proved innocent approach”, where one assumes that the quality is inadequate unless the information to the contrary is provided in the text. This can be a dangerous tactic and may exclude many valid studies. The reader should take the same approach as described for dealing with potential bias and confounding and balance any missing information with its potential impact on the conclusions of the report.

**REFERENCES**


ANSWERS
1. T; 2. F (it is a particular problem in case-control studies); 3. F; 4. F; 5. F (matching is used to control for confounding).

Clinical Evidence—Call for contributors

Clinical Evidence is a regularly updated evidence based journal available worldwide both as a paper version and on the internet. Clinical Evidence needs to recruit a number of new contributors. Contributors are health care professionals or epidemiologists with experience in evidence based medicine and the ability to write in a concise and structured way.

Currently, we are interested in finding contributors with an interest in the following clinical areas:

- Altitude sickness
- Autism
- Basal cell carcinoma
- Breast feeding
- Carbon monoxide poisoning
- Cervical cancer
- Cystic fibrosis
- Ectopic pregnancy
- Grief/bereavement
- Halitosis
- Hodgkins disease
- Infectious mononucleosis (glandular fever)
- Kidney stones
- Malignant melanoma (metastatic)
- Mesothelioma
- Myeloma
- Ovarian cyst
- Pancreatitis (acute)
- Pancreatitis (chronic)
- Polymyalgia rheumatica
- Post-partum haemorrhage
- Pulmonary embolism
- Recurrent miscarriage
- Repetitive strain injury
- Scoliosis
- Seasonal affective disorder
- Squint
- Systemic lupus erythematosus
- Testicular cancer
- Varicocele
- Viral meningitis
- Vitiligo

However, we are always looking for others, so do not let this list discourage you.

Being a contributor involves:

- Appraising the results of literature searches (performed by our Information Specialists) to identify high quality evidence for inclusion in the journal.
- Writing to a highly structured template (about 2000–3000 words), using evidence from selected studies, within 6–8 weeks of receiving the literature search results.
- Working with Clinical Evidence Editors to ensure that the text meets rigorous epidemiological and style standards.
- Updating the text every eight months to incorporate new evidence.
- Expanding the topic to include new questions once every 12-18 months.

If you would like to become a contributor for Clinical Evidence or require more information about what this involves please send your contact details and a copy of your CV, clearly stating the clinical area you are interested in, to Claire Folkes (cfolkes@bmjgroup.com).

Call for peer reviewers

Clinical Evidence also needs to recruit a number of new peer reviewers specifically with an interest in the clinical areas stated above, and also others related to general practice. Peer reviewers are health care professionals or epidemiologists with experience in evidence based medicine. As a peer reviewer you would be asked for your views on the clinical relevance, validity, and accessibility of specific topics within the journal, and their usefulness to the intended audience (international generalists and health care professionals, possibly with limited statistical knowledge). Topics are usually 2000–3000 words in length and we would ask you to review between 2–5 topics per year. The peer review process takes place throughout the year, and our turnaround time for each review is ideally 10–14 days.

If you are interested in becoming a peer reviewer for Clinical Evidence, please complete the peer review questionnaire at www.clinicalevidence.com or contact Claire Folkes (cfolkes@bmjgroup.com).