Good Epidemiology Practice (GEP)

Author : P N Lee

Date : 12.2.96.

1. The need for GEP

Epidemiology is the examination of relationships between a putative causative agent and a disease by observing human populations. There is general agreement that reliable human evidence should take precedence over animal evidence from toxicological studies when forming public policy and making regulatory decisions concerning risks to humans. However, drawing inferences from epidemiological data is much more difficult than drawing inferences from toxicological data. Toxicological studies, unlike epidemiological studies, can rigidly control exposure conditions, so avoiding any problems in defining exposure or measuring it accurately. Again unlike epidemiological studies, they can assign animals randomly to experimental groups, so avoiding problems of bias due to confounding by other risk factors. Toxicological studies can also more easily ensure standardization of measurement of health outcomes. Study design and statistical analysis are also more complex for epidemiological than for toxicological studies. It is clear that epidemiological studies are less easy to do well than is the case for toxicological studies. It is also clear that epidemiological studies are often badly conducted, with their reported findings open to criticism and controversy.

Good Laboratory Practice (GLP) has been with us for many years and has helped to ensure that toxicological studies are well done. Given the greater difficulty of doing epidemiology well and the greater perceived relevance of its findings, there would seem to be a stronger case for GEP than for GLP.

Provided suitable GEP guidelines are prepared, and provided epidemiologists keep to them, GEP would seem to have a number of advantages:

(i) it would enhance the quality of epidemiological research,

(ii) it would rationalize the relationship between epidemiology and public policy formulation,

(iii) it would improve public confidence in epidemiology and reduce public misapprehension and fear caused by deficient study conduct or reporting of

results, and

(iv) it would help towards a more optimal use of research resources.

Appropriate guidelines should also endeavour to:

(v) ensure that all relevant findings are made available, so avoiding problems of publication bias and improving the validity of "meta-analyses" summarizing the overall evidence on an association from multiple studies.

2. Outline of guidelines

This document presents proposed GEP guidelines, which are outlined in this section, and presented in more detail, with justification as appropriate, in the sections that follow.

The first part of the guidelines (see section 3) relates to the fact that a formal protocol should be prepared at the outset, describing not only the study design and objectives, but also giving full details of all aspects of it, including the study investigators, other organizations involved, the sponsors, the reason for doing the study, the time the study will take, how the data will be processed, statistically analyzed, interpreted, reported and made available, and demonstrating that adequate facilities and resources are available.

This protocol should be peer-reviewed by independent scientists and considered by an ethical review board (section 4).

Protocol departures, unintentional or planned, should be documented and any effect that these departures have on the study integrity should be clarified (section 5).

Standard operating procedures (SOPs) should be written to describe how the various tasks in the study are to be carried out, and an independent quality assurance (QA) unit should review the adequacy of the SOPs, ensure that the SOPs and the protocol are kept to, and satisfy themselves that the data are recorded accurately, with any data modifications justified (section 6).

While, on occasion, the final results can be adequately reported in a paper for publication, there will usually be a need for a more extensive report to describe the study findings in adequate detail. This report should, inter alia, describe the research objectives and study rationale, the study methods and procedures, all relevant results of statistical analysis, and give the interpretation of the findings. The analysis and interpretation should make clear which sources of bias were identified and the extent to which they were controlled for, indicate the uncertainty of results arising both from sampling variation and from uncontrolled bias, and make clear how the results fit in with existing literature and affect overall conclusions. The report, the format of which is described more fully in section 7, should be peer-reviewed.

All relevant study materials should be archived in a secure location. Copies of the full report should be made available to independent scientists and relevant regulatory authorities. The full data should be held on computer files which should be made available to independent scientists, subject to any ethical, privacy or professional considerations (section 8).

GEP principles should not only apply to individual epidemiological studies but also to formal meta-analyses of existing study data. Here, the protocol should also make it clear how studies are to be identified and selected for consideration, the aspects of study design to be taken account of in analysis, the statistical methods used to calculate overall summary statistics and to test for heterogeneity of results and the action to be taken if heterogeneity is detected (section 9).

The guidelines outlined here are principally intended to provide suitable criteria for conduct of epidemiological studies aimed at formally testing specific research hypotheses developed in advance of the study. It is recognized that a number of epidemiological studies are aimed wholly or partly at developing new hypotheses, and that here it may not be possible to define fully in advance in the protocol the hypotheses, the statistical analyses and the reports to be produced. Indeed, some studies may be concerned only with collection of data to be made available for other epidemiologists and

statisticians to analyze and interpret in due course. In these cases, only those parts of the guidelines which are applicable should be considered. The part of section 3 relating to statistical analysis makes some comments concerning the interpretation of results from multiple hypotheses, not defined a priori.

3. The study protocol

The study should have a written protocol. This protocol must be approved before the study begins (see section 4). The protocol should include the following:

3.1 Title A title descriptive of the study.

3.2 Organization The company (or individual) responsible for the study.

3.3 Abstract A brief summary of the main features of the protocol.

3.4 Study investigators The names, titles, degrees, addresses and affiliations of the principal investigator and all the main co-investigators.

3.5 Other organizations involved Where part of the study is sub-contracted to another organization, e.g. interviewing to a market research company, the protocol should contain the names, titles, degrees, addresses and affiliations of the main persons responsible for the research in that organization.

Note that personnel engaged in epidemiological research and related activities should have the education, training and experience necessary to perform the assigned functions in a competent manner. Organizations involved in GEP research should maintain and update periodically summaries of training and experience and a job description for each individual engaged in or supervising relevant activities.

3.6 Sponsors The sponsors of the study, if not the organization responsible for it, should be named, with their address given.

3.7 Facilities and resources Adequate facilities and resources should be available to all those engaged in epidemiological research and related activities so as to ensure a proper and timely completion of the study and to be able to store research materials in a safe and secure environment. The protocol should give enough detail to ensure the reader that this is so.

3.8 Timing The protocol should make clear when it was approved, when the study is to start, when any periodic progress reviews are planned, when the various stages of data collection are scheduled for completion, and when the study report is expected to be completed.

3.9 Hypotheses and rationale It should normally be made clear what are the main and any subsidiary hypotheses to be tested. It should be clearly stated whether any hypothesis is one-sided or two-sided, and, if appropriate, what the alternative hypothesis is. Thus one might be testing whether the agent of interest affects risk of cancer (two-sided hypothesis, with the null hypothesis as the alternative) or whether the agent of interest increases risk of cancer by more than 20% (one-sided hypothesis, against the alternative hypothesis that it increases risk by no more than 20%).

The protocol should also make it clear why the study is being conducted, e.g. to address a previously unanswered question or to attempt to corroborate or confirm previous findings (perhaps using what are felt to be more appropriate techniques).

Exceptionally, a study may be exploratory, collecting data for examination of a variety of possible hypotheses (e.g. a prospective study collecting data on an extensive list of risk factors, for correlation ultimately with risk of a range of causes of death). Here the protocol need only state in general terms the intention of the study.

3.10 Literature review The protocol should include the findings from a review of relevant literature. This should encompass pertinent animal experiments, clinical studies, vital statistics and previous epidemiologic studies. The review should concern the main and all subsidiary hypotheses and be of sufficient depth to identify all the potential

confounding variables and effect modifiers that need to be considered in the study. It will not always be necessary for the study investigators to have conducted the review themselves. However, if they place reliance on published reviews, they should be able to satisfy the reader that the review fulfils its main two purposes, viz. ensuring that resources are not being spent on unnecessary duplication of previous research, and identifying areas where new knowledge is needed.

3.11 Study design The protocol should describe, in adequate detail to allow replication of the study, all relevant aspects of the study design. For a multicentre study any differences between centres should be noted. These should include:

3.11.1 **Study type** The protocol should clarify the type of study, e.g.

(a) *prospective* (longitudinal or cohort) in which information on exposure to the agent (and other relevant risk factors) is collected from subjects who are initially disease-free and onset-rates of subsequent disease or death are compared in groups with different levels of exposure;

(b) *case-control*, in which current and past exposure to the agent is compared between a group of patients with the disease of interest (cases) and a group without the disease (controls);

(c) *cross-sectional*, in which a sample of the population is taken at one point in time and data on exposure and presence of disease are recorded; or

(d) *ecological*, in which data on average exposure and average risk are obtained for a number of different populations, for example in different countries, and an attempt is made to ascertain whether average risk varies according to average exposure.

3.11.2 **Method of sample selection** It should be made clear which populations the various study samples derive from (e.g. a study of the general population, of nurses in Yorkshire, of coal-miners working in certain pits in a given period), as well as how the samples are to be selected from these populations. Exclusion criteria should be defined. For case-control studies, any matching variables should be stated, as well as the method of matching.

3.11.3 **Sample size** The planned numbers in the sample should be given, by age, sex and other demographic variables as appropriate. For multi-centre studies the sample size should be given for each centre.

3.11.4 **Data collection** The data to be collected and the data collection methods to be used should be specified clearly. It will be convenient to subdivide this section according to:

(a) *disease endpoints* Are these based on post-mortem examination, death certificates, clinical examination, histopathological investigation, biochemical tests and/or questionnaires? What are the criteria used for a positive diagnosis?

(b) *exposure measures* Are these based on monitoring devices, biochemical tests and/or questionnaires?

(c) *potential confounding variables* Which variables are to be considered and how are they to be measured? These may include not only variables that may affect risk of the disease endpoint(s) considered, but also variables that may affect the validity of measurement of disease, exposure and/or other confounding variables.

(d) *other sources of bias* Bias in an epidemiological study derives from the following possible sources; errors in determining disease endpoints, errors in determining exposure measures, confounding, errors in determining confounding variables, errors in selecting subjects into the study groups, and unrepresentativeness of the study groups for the populations they are intended to be relevant to. Any data collected relating to these sources of bias should be described. For example the study might include:

(i) attempts to obtain information on non-responders to assess the representativeness of those who do respond;

(ii) analysis of duplicate food samples in a subset of subjects to allow correction to be made to results based on dietary questionnaire data for all subjects;

(iii) review of cancer diagnoses by independent pathologists; or

(iv) measurement of cotinine in saliva samples to allow confirmation of self-reported nonsmoking status in a study intended to be restricted to nonsmokers.

The protocol should make it clear that the following points have been taken into account in data collection:

(i) Any potential health risk to participants, or any other form of disadvantage or embarrassment to them should be identified. This should be made clear not only in the protocol, but also to the study subjects in a suitably worded consent form.

(ii) The confidentiality of data collected specific to each subject should be maintained as far as possible. Normally subjects should be identified by reference number and not by name. Where it is necessary for the name to be recorded, e.g. for subsequent identification of health endpoint data from medical records to be linked to exposure data already collected, procedures should be adopted to ensure that computer files or data listings should not allow ready association of the subject's name with the data collected for them.

(iii) Any data should be collected, as far as possible, by techniques which minimize error. It is particularly important that blind measurement should be used if it can be. Ideally, exposure data should be collected in the absence of knowledge of disease, as knowledge of disease, by the respondent, the interviewer (or even the laboratory technician) may affect data collected on exposure. Similarly data on health endpoints should be collected, as far as possible, in the absence of knowledge of the exposure(s) of interest.

(iv) Missing data should be clearly identified. One should be able to distinguish data which are missing for different reasons. For example, a blood chemistry measurement might be missing because the subject refused to give blood, because the sample taken was inadequate or because of a failure in chemical analysis. The study should be designed to minimize missing data.

(v) Information should be obtained as far as possible on the validity and reliability of any data recorded. The accuracy of any relative risk estimate obtained will be affected by errors in the determination of health endpoints, exposures and confounding variables, and the extent of such inaccuracy can only be estimated if data are available on the magnitude of these errors. Such information may be obtained from, e.g., carrying out duplicate assays on the same biochemical samples, asking subjects the same questions at two points in time, obtaining

corroborative information from next-of-kin, etc. Often such extra data need only be collected for a sub-sample and on occasion reliance may have to be based on published literature on validity or reliability from special studies.

3.12 Rationale for study design There are usually many possible study designs available to test a given hypothesis. The protocol should make it clear why the particular study design was selected and, in general, justify any of the main choices made as regards study type, method of sample selection, data to be collected and method of data collection. It is important also to give reasons why the particular sample size was chosen, with results of power calculations presented to determine the magnitude of effect that can be detected subject to stated false-positive and false-negative probabilities.

3.13 Limitations of study design Although the study should be designed to be as valid as possible, there will inevitably be some limitations of the chosen study design. These should be explained clearly. In particular any sources of bias that cannot be controlled directly in the study design and/or statistical analysis should be identified.

3.14 Ending the study Any factors that affect the time at which the study is completed should be referred to. For example a prospective study might continue until a defined number of the original population has died from the disease of interest, or a case-control study may continue until a defined number of cases fitting certain criteria have successfully been interviewed. In some studies a decision to continue may depend on the results of interim statistical analyses of the main hypothesis.

Conditions that would result in premature abandonment of the study should also be mentioned. For example a study might be abandoned if the response rate is found to be below some defined minimum value for fear that the subjects interviewed may be inadequately representative of the population of interest. Or it may become clear that it is impossible to obtain adequate numbers of cases in a reasonable length of time.

3.15 Pilot study Whether a study design will work properly in practice can often better be judged by carrying out a pilot study first with a small number of subjects. Sometimes the

pilot study will concern only certain aspects of the planned study. The protocol for the study proper should include reference to any details of any relevant pilot studies that have been previously conducted.

Especially where the planned study is large and costly and involves some areas where previous information is limited, pilot studies are to be recommended in principle, to avoid the possibility of wasting substantial effort on a study that ultimately proves unsatisfactory or has to be abandoned.

3.16 Data processing Methods of data entry to computer and any checks to be made for data accuracy and consistency should be described in general terms. Where the data are not entered directly into computer, e.g. interviewers using lap-top computers, care should be taken to ensure that errors are not introduced by transfer of data from paper (e.g. questionnaires) to computer. Ideally, and particularly for the more important data, there should be independent dual data entry, with any discrepancies detected and corrected. (An alternative procedure might be to conduct dual entry only on a sample of data initially, with provision to extend to full dual entry if error rates from the sample exceed a defined minimum.) The data being transferred to computer may also be in error and software should detect out-of-range values or values that are inconsistent with other data. Such erroneous or inconsistent values (and also any unreadable values) should be referred back to the person responsible for producing the original data sheets and errors rectified as far as possible. Any changes to data should be clearly marked, signed and dated with the reason for the change given.
Criteria for acceptability of data of various types should be defined and stated, and the software to be used in data processing should be named in the protocol.

Data processing should have particular regard to data confidentiality. Where possible the identity of the subject and the data collected for that subject should be kept separate by use of subject reference numbers.

3.17 Statistical analysis The protocol should contain adequate details of the planned statistical analyses. It is important (except perhaps for a hypothesis generating study)

that these plans should be laid down in advance in order to avoid the possibility that investigators choose to conduct numerous analyses and then only present the results of those that fit in best with their favourite hypothesis. Choice of exact health endpoint, precise exposure measure or confounding variables included can often affect the strength of the relative risk observed and its statistical significance.

The protocol should include:

3.17.1 **Methodology** A broad description of the statistical techniques to be used. Where possible these should be the standard accepted techniques, for example as defined in the IARC Statistical Monographs by Breslow and Day on case-control studies (1980) and on cohort studies (1987).

3.17.2 **Software** Any software that is planned to be used should be named, although it is appreciated that in some studies software may be developed during the study.

3.17.3 **Statistical significance and confidence levels** As noted earlier (section 3.9), the hypothesis should make it clear whether one-sided or two-sided tests are to be used. Criteria for statistical significance should be stated; however, it is not acceptable just to present results as, say, "$p<0.05$" or "not significant" as this gives inadequate information as to the strength of the evidence in favour of the hypothesis. While it is not necessary always to give the precise p value, it will be preferable to give p values in ranges such as $p<0.001$, $p<0.01$, $p<0.05$, $p<0.1$ or $p \geq 0.1$. Confidence intervals should also be presented.

It should be made clear whether exact p values are calculated or whether these are approximations, e.g. based on asymptotic chisquared statistics. Computing power has developed sufficiently for exact tests to be used in more analyses than hitherto.

In some studies, particularly hypothesis-generating studies, multiple analyses may be conducted. The protocol should describe the precautions to be taken in the statistical analysis to avoid over-interpretation of isolated statistically significant findings. This

condition does not necessarily imply that formal statistical procedures should be adopted that correct for multiple testing. Especially if one wishes to extract data from the study relevant to a specific comparison, for integration into a review or meta-analysis, it can be very unhelpful if the p-value for that comparison is adjusted upwards to account for other comparisons that happened to be investigated in the study.

3.17.4 **Main comparisons of interest** Especially for the main hypothesis to be tested, the protocol should make clear the exact definition of the health endpoint(s) and the exposure indices to be considered. If a dose-related trend analysis is envisaged, the protocol should define in advance not only the index of exposure but also how it is to be categorized into groups.

3.17.5 **Biases to be adjusted for** Data will have been collected relevant to various sources of bias and confounding. The protocol should describe how these are to be taken into account in analysis.

3.17.6 **Expression of uncertainty of results** Many published papers present results of statistical analyses relating a risk factor to a disease, adjusted for various confounding factors, in the form of a relative risk with 95% confidence limits. In the discussion of the paper, various uncontrolled sources of bias are considered in general terms and the authors state their opinion as to whether these are important or not. Often some potential sources of bias are not referred to at all.

This procedure, though convenient, is not totally satisfactory, as it gives little quantitative feel for the uncertainty of the relative risk estimate. In particular the confidence limits only indicate sampling variation and do not express at all the extent of uncertainty arising from uncontrolled sources of bias. GEP studies should move towards an improvement of this situation and should state in the protocol the procedures to be adopted to quantify uncertainty from such sources. These procedures should consider the possibilities of bias due to:

(i) sample selection, e.g. arising from non-response, or in errors in variables used as criteria for entry into the study,

(ii) errors in determining health endpoints,

(iii) errors in measuring exposure,

(iv) failure to include relevant confounding variables, and

(v) errors in measuring confounding variables that were included (so-called "residual confounding").

For each source of bias, the general approach should be to quantify its likely magnitude based on a range of plausible assumptions. The range should be guided, as far as possible, by evidence collected in the study and available in the literature. There should also be an attempt to quantify the likely magnitude of the overall bias, from all sources considered.

3.17.7 **Interim analyses** The protocol should make it clear whether there is to be a single statistical analysis at the end of the study or whether analyses are planned at various stages, and if so whether they have implications for the termination of the study.

3.17.8 **Multi-centre studies** For a multi-centre study the protocol should describe whether analyses are to be conducted overall, or centre by centre, any procedures to test for heterogeneity of findings over centre, and how heterogeneity will be dealt with when describing overall results.

3.27.9 **Extrapolations** On occasion study investigators will wish to extrapolate their findings to a wider population, e.g. to conclude that x thousand deaths per year in a country from a particular cause can be attributed to a given agent. If such analyses are to be conducted, the protocol should make it clear how account is to be taken of relevant differences in distribution of disease, risk factors or confounding variables between the study population and the population to which the results are to be extrapolated, any assumptions involved in the extrapolation, and how many uncertainties in the extrapolation are to be quantified.

3.18 Interpretation of analyses Provided the statistical analysis adjusts for as many potential sources of bias and confounding as possible, adequately expresses the uncertainty resulting from uncontrolled sources of bias and confounding, and gives appropriate

criteria for statistical significance, it should be relatively easy in many cases to determine whether or not any observed relative risk can be explained by chance, bias or confounding. Other criteria need to be considered in causal inference, however, including biological plausibility, consistency of findings with other reported evidence, and exclusion of the possibility that the disease affected exposure rather than vice versa.

The protocol should describe the types of evidence to be taken into account in making causal inferences.

3.19 Reporting of results The protocol should make it clear that there is an intention to report the findings regardless of the results. Even if the study has to be abandoned there should be a brief report giving the reasons why.

The protocol should also make it clear what types of report are planned, including:

(i) Any interim reports,
(ii) Full study report,
(iii) Papers for publication in journals,
(iv) Making results known to the media and to study participants.

(See section 7 for details about content of reports.)

3.20 Availability of reports and data The protocol should also mention provisions made to archive study materials and to make available reports and data to interested parties.

(See also section 8.)

3.21 References Bibliographic references of cited papers should be included.

3.22 Signed and dated The principal investigator should sign and date the protocol.

3.23 Appendices Attached as appendices to the protocol should be such material as relevant correspondence, agreements between collaborating organizations, approvals from

institutions, reports from those reviewing the protocol (see section 4), detailed findings from any pilot studies, samples of any documents to be used in the study (such as consent forms and questionnaires) and anything else of particular relevance to the study.

4. Protocol Review

The study protocol should receive scientific review by independent person(s) not part of the investigative team. The reviewer should ensure that the protocol is written according to GEP guidelines and that the methods to be used in the study are appropriate. The reviewer should be aware that budget or time constraints may rule out the use of some scientifically superior alternative techniques, but should be ready to point out any technical improvements that are felt to be feasible. The reviewer should also point out any deficiencies in the literature review or in the arguments put forward in the protocol as to why certain procedures have or have not been employed. Generally the review should attempt to ensure that the study is capable of addressing adequately the hypotheses of the study.

The ethical aspects of the study should also be reviewed. This review should consider the following issues:

(i) the health and welfare of the study subjects and investigators,
(ii) the need for informed consent by subjects and investigators,
(iii) maintenance of confidentiality and privacy,
(iv) avoidance of conflicts of interest,
(v) adherence to the law,
(vi) promotion of public confidence in epidemiological research,
(vii) dissemination of the findings to all relevant individuals as well as any other issues which may affect obligations to the research subjects, those conducting the study, the sponsors and employers, colleagues and society in general.

Sponsors, contractors and any third parties involved should review the protocol in order to ensure that adequate resources are available to complete the study on time and in appropriate fashion.

5. Protocol departures

Where a protocol requires extensive change, e.g. if the scientific review board make large numbers of suggestions, a new version of the protocol can be prepared. In these circumstances it should make clear which version it supersedes and why.

More usually any protocol departures should be documented by protocol amendments. Any protocol amendment should describe the changes to be made, explain why they were necessary and clarify any effect they may have on the integrity of the study. Protocol amendments should be signed and dated by the principal investigator and kept with the original protocol.

6. Quality assurance (QA)

6.1 Standard operating procedures (SOPs) are written detailed descriptions of routine procedures involved in performing epidemiological studies. Use of SOPs improves reproducibility, accuracy and validity of findings. A designated individual should develop, review and update SOPs in his or her area of responsibility. New and updated versions should be signed as approved by appropriate management. A manual of SOPs, including all revisions, should be available to all relevant personnel.

SOPs should include a statement of the purpose of the procedure and a detailed description of it, as well as the training level required to perform the procedure. It should also include the date at which it became effective, the issue/revision number, the signature of the author, and the signature of the person reviewing and authorizing it.

SOPs can be established for a variety of activities relevant to epidemiological studies, including data collection, data validation, data coding, assessment of error rates in data processing, various aspects of statistics analysis, chemical analytic methods and archive management.

6.2 QA unit GEP envisages an independent check of work done by a QA unit. The unit may be within the organization conducting the study or outside. A QA auditor from within the QA unit shall have specific responsibility for the study and shall prepare, at

intervals, a report reviewing study compliance with procedures. The QA auditor should attempt to ensure:

(i) that there are an adequate number of SOPs to describe all the relevant procedures in the study;

(ii) that the SOPs themselves are adequate;

(iii) that the SOPs are being kept to;

(iv) that the study protocol itself is being kept to;

(v) that changes to raw data are signed and adequately justified; and

(vi) that any software is appropriate and has been adequately validated.

7. **Reporting**

Limitations of space in a scientific journal usually means that a study cannot be reported in sufficient detail to provide adequate coverage of methods and results. Although, on occasion, it may be sufficient to use a published paper as the final report, it will more usually be appropriate to prepare a separate full report to satisfy GEP requirements. The full report should include the following items:

7.1 Title A title descriptive of the report.

7.2 Abstract A brief summary of the hypothesis, study design and results.

7.3 Study investigators The names, titles, degrees, addresses and affiliations of the principal investigator and all the main co-investigators.

7.4 Other organizations involved Where part of the study is sub-contracted to another organization, the protocol should contain the names, titles, degrees, addresses and affiliations of the main persons responsible for the research in that organization.

7.5 Sponsors The sponsor of the study, if not the organization responsible for it, should be named, with their address given.

7.6 Timing The dates on which the study was initiated and completed, separately for major

parts of the study if appropriate.

7.7 Introduction This should include the hypothesis to be tested, the reason for doing the study and a summary of relevant literature from the review that was conducted (see section 3.10). Results of any pilot studies should be referred to.

7.8 Study methods Relevant aspects of study design (study type, method of sample selection, sample size, data to be collected, methods of data collection), data processing, and statistical methods should be included. Any departures from the original protocol should be referred to, together with any effect they might have on the study quality.

7.9 Results All relevant results of the statistical analyses carried out should be shown, regardless of whether or not statistically significant relationships were seen. It should be made clear which sources of bias were identified and the extent to which they were controlled. If appropriate, extensive results should be summarized in the main body of the text and presented in detail in Appendix tables.

7.10 Expression of uncertainty Results of sensitivity analyses should be presented, describing the likely magnitude of bias resulting from the single and joint effects of any sources of bias that could not be controlled for fully in the statistical analyses. The assumptions behind the sensitivity analyses should be presented clearly.

7.11 Interpretation of the findings This section should attempt to make it clear whether or not relevant associations can be explained by chance, taking account multiplicity of analyses where relevant, and whether or not these associations can be explained by biases controlled for in the main statistical analyses or by other biases considered in the sensitivity analyses. Biological plausibility should also be considered and possible mechanisms discussed. How the results of the study fit in with existing literature from other studies on the hypothesis of interest should be considered, as well as how the new findings affect the overall conclusions to be drawn. Limitations of the study should be indicated, with, if appropriate, suggestions made as to further studies that are ongoing or might usefully be conducted.

7.12 References Bibliographic references of cited papers should be included.

7.13 Signed and dated The principal investigator should sign and date the protocol.

7.14 Appendices Appendices to the report should give either raw data listings (identified by reference number not person name) plus summary statistics or, in the case of extensive data, details of the computer data file(s) holding the data plus summary statistics.

If appropriate, appendices should give more detailed results from statistical analyses summarized in section 7.9.

Other appendices should give other relevant material such as questionnaires and other data collection forms, consent forms, important correspondence, and institution approvals.

7.15 Peer-review It is important that the report should be independently peer-reviewed. Often this will be by the same epidemiologist(s) who reviewed the study protocol.

8. Availability of data and findings

8.1 Archiving Secure archives should be available for orderly storage and retrieval of all study related material. The contents should be described by an index which should identify their location within the archive as well as identifying by name and location any materials held elsewhere.

Access to the archives should be limited to authorized personnel. Special procedures may need to be implemented to ensure that confidentiality of information about study subjects is protected.

The archive should contain, or at least refer to the whereabouts of, the following materials:

(i) the study protocol and protocol departures,

(ii) the final study report,

(iii) all the source data and, where feasible, specimens,

(iv) a printout of a sample of the computer data file(s),

(v) a disk (or equivalent) containing the master data file(s),

(vi) sufficient documentation to identify all computer programs and statistical procedures used,

(vii) a copy of all output, as hard copy or on disk, relevant to any tabulations or statistical analyses appearing or referred to in the final report,

(viii) any laboratory/research notebooks relevant to the study,

(ix) standard operating procedures,

(x) copies of all quality assurance reports and audits,

(xi) all completed informed consent forms, and

(xii) any other relevant correspondence, reviewer reports, agreements, etc., not already included as appendices to the final report.

8.2 Availability of report Copies of the final report should be made available (at reasonable cost) to independent scientists and relevant regulatory authorities requesting it. The existence of the report should be made clear in papers for publication, or if this is some time after the report has been finalized, by an announcement in a suitable medical or epidemiological journal. Alternatively the final report may be made available through a national auxiliary publishing service if one is available.

8.3 Availability of data Copies of the computer data file(s) should also be made available (again at reasonable cost) to independent scientists and regulatory authorities subject to any ethical, privacy or professional considerations. The study investigators should be expected to grant permission to publish any analyses of these data provided that:

(i) they themselves are conducted according to GEP guidelines,

(ii) the study investigators and sponsors are acknowledged for their assistance, and

(iii) they are given a copy of the publication.

9. **Meta-analyses**

GEP principles should also apply to meta-analyses of existing data from multiple

studies, though of course some of the specific sections described above will not be relevant. Particular features of meta-analyses that should be referred to in the protocol (and the final report) are described below.

9.1 Identification of studies

It should be made clear how the investigators intend to identify studies of potential interest. A common procedure is likely to be literature search followed by acquiring all other relevant papers cited in the original search (or cited in these subsequently acquired papers). However, it is often advantageous to try to obtain unpublished reports (so as to minimize publication bias) and any methods to do this should be described in the protocol.

9.2 Criteria for selection of studies

Criteria should be laid down by which seriously scientifically deficient studies should be excluded. It will also be necessary to exclude results for studies where updated results have been published later. Except perhaps for exclusion of extremely small studies, e.g. involving 5 cases or less of the disease of interest, studies should not be selected on sample size. Unless the hypothesis under consideration is extremely specific, insistence on exact comparability of studies is not to be recommended, as it limits the generality of the findings. Rather it is preferable to allow variability in study design aspects and consider in the statistical analysis whether the association between the agent and disease of interest is dependent on any of these aspects.

9.3 Obtaining the full data

The protocol should make it clear whether any attempts are to be made to obtain the full data. Provided the full data can be obtained for all, or a substantial majority of studies, this is beneficial in allowing more detailed statistical analyses to be attempted. A single study, for example, may not have sufficient power to look meaningfully at whether an observed association varies within age groups but a combined analysis may allow this. Having the full data may also more readily allow comparable statistical analyses to be conducted for each study.

9.4 Results to be used

Different papers may present results for a variety of differing analyses. The protocol should make clear in advance which ones the meta-analysis is concerned with.

9.5 Study characteristics

In principle the data can be seen as a comparable set of statistics for each study concerning the association of interest (e.g. relative risk plus 95% confidence limits), together with a number of covariates describing characteristics which vary from study to study, such as location, time to study, study type, methods of determining exposure, method of determining diagnosis, confounding variables accounted for, and various indicators of study quality. The data to be collected as covariates, the study characteristics, should be defined at the outset.

9.6 Statistical methods

The protocol should make clear the techniques used to compete overall summary statistics and to test for heterogeneity of the association. It should also describe the action to be taken if heterogeneity is detected, preference being given to stratification or regression methods allowing for expression of variation in association according to the different study characteristics rather than using "random-effect" summaries. The protocol should also describe any methods used to detect, or adjust for, publication bias.

9.7 Study quality

It would not be appropriate to insist on only including studies adhering to GEP guidelines in the meta-analysis or indeed on only including peer-reviewed studies. Insistence on the former would mean meta-analyses could not be envisaged for many years to come, while insistence on the latter would promote publication bias. Whether a study is GEP, peer-reviewed but not GEP, or not peer-reviewed could be considered a study characteristic in statistical analysis.