Histol Histopathol (1999) 14: 203-215

http://www.ehu.es/histol-histopathol

Histology and Histopathology

Invited Review

Systematic review and meta-analysis in anatomic pathology: the value of nuclear DNA content in predicting progression in low grade CIN, the significance of the histological subtype on prognosis in cervical carcinoma

M.K. Heatley

Department of Cellular Pathology, Taunton and Somerset NHS Trust, Taunton, UK

Summary. Meta-analysis, though increasingly popular in clinical medicine, has not found acceptance in anatomic pathology. This paper argues that, in combination with a systematic review of the literature, meta-analysis may be usefully applied to pathological research and two examples drawn from gynaecological pathology (the value of nuclear DNA quantitation in predicting progression in low grade cervical intra-epithelial neoplasia and the difference in prognosis between squamous cell carcinoma and adenocarcinoma of the cervix) are included to illustrate the methods used and to demonstrate some of the difficulties associated with these techniques.

Key words: Meta-analysis, Cervical intraepithelial neoplasia, Aneuploidy, Adenocarcinoma, Squamous cell carcinoma

Introduction

Overviews of research have long been necessary when studies were inconclusive, produced conflicting results or when definitive studies were deemed to be impossible (Sacks et al., 1987). Contemporary approaches to providing such overviews include the techniques of systematic review and meta-analysis. Systematic review permits data from a number of different studies to be aggregated allowing the results to be interpreted in the context of a given clinical problem. Meta-analysis is a quantitative synthesis of data from several studies, usually involving the use of statistical methods, which may have been identified on systematic review.

Meta-analysis has not found acceptance in pathological research and a medline search combining the text words meta-analysis with anatomic pathology, surgical pathology and pathology generated no matches. This paper argues that the techniques employed in meta-analysis can and should be applied to pathological research.

Examples of a systematic review of nuclear DNA quantitation in cervical intraepithelial neoplasia (CIN) and a meta-analysis of the prognostic implications of glandular differentiation in low stage cervical carcinoma demonstrate the methods used and illustrate some of the problems which may be encountered in performing this kind of review.

Why carry out reviews?

Much medical research consists of assessing the effect of an intervention either in the healthy individual, for example, the epidemiological effect of cigarette smoking in healthy adults and the subsequent development on carcinoma of the lung and coronary heart disease, or in the diseased state. The latter, therapeutic research, most commonly assesses how drugs modify the course of disease and uses the randomised controled trial as its gold standard for assessing new drugs.

On occasion, existing evidence may be limited to a number of studies based on only a few cases. Further studies may not be feasible because the disease is rare preventing the accumulation of sufficient numbers of cases to allow the effect of the intervention and the effect of chance to be balanced (Chalmers et al., 1992) or because withholding an intervention in a condition with a dismal outlook may be considered unethical

Offprint requests to: Dr. Mark K. Heatley, Department of Pathology, Royal Liverpool and Broadgreen University Hospitals NHS Trust, 5th Floor Duncan Building, Prescot Street, Liverpool L7 8XP, UK. Fax: 0151-706-5859

(Chalmers et al., 1992). To make matters worse a number of studies may have produced apparently conflicting results. Finally the results obtained in a sample of one age range, racial group or in one of the sexes may not necessarily be applicable to older or younger subjects or to patients of different ethnic origin or gender. Since it would be difficult to organise a study which allowed every possible permutation to be assessed, there has always been a need to synthesise the evidence from existing studies allowing a more rational approach to patient management. Traditionally this has been attempted using the narrative review in which an individual attempts to draw meaningful conclusions from the, sometimes conflicting results of existing studies. Systematic review with or without meta-analysis is an alternative to this approach.

Advantages of systematic review

A systematic review seeks to impose a structured approach to the review of existing data with regard to the methods of locating studies, combining data and assessing outcome (Thacker, 1988; Thacker et al., 1996). By combining the data from a range of existing studies carried out in different localities a systematic review resembles a multicentre trial and the review process has features in common with primary research (Chalmers et al., 1992; Oxman, 1994; Peipert and Bracken, 1997). By a simple increase in the size of the sample under study, meta-analysis of this data increases the power of the statistical analyses (Sacks et al., 1987; Peipert and Bracken, 1997) allowing summary descriptive statistics to be generated across the range of studies (Thacker, 1988). It allows a more precise assessment of the impact of treatment or risk factors for a disease (Peipert and Bracken, 1997) and improves estimates of the magnitude of the effect of interventions (Sacks et al., 1987; Thacker et al., 1996) such as differences in the drug dosages or degree of exposure to an environmental pollutant. The relationships between different aetiological features and therapeutic modalities may be examined (Thacker et al., 1996). Since the subgroups contained within an individual study may be too small to permit meaningful statistical analysis, combining studies may generate subgroups which are large enough to permit a statistical manipulation (Thacker, 1988). The source studies may have originated in a variety of ethnic groups in different countries and as a consequence the results are more generalisable than those of a single centre or multicentre study confined to a single country or continent (Thacker, 1988; Thacker et al., 1996). Furthermore since a variety of investigators are involved in planning and executing the studies, investigative bias and subjectivity is mitigated (Thacker, 1988). By balancing their outcomes with those of larger studies meta-analysis reduces the impact of studies consisting of small numbers of patients which have "statistically significant" results. As is discussed later such studies are more likely to be submitted for publication than studies in which the results are "not significant" (Koren et al., 1989; Easterbrook et al., 1991; Dickersin et al., 1992) and may give an exaggerated impression of the effectiveness of a therapy but seem likely to continue to be performed whilst researchers are judged in terms of their output of papers in peer review journals at the expense of their contribution to multicentre studies (Newcombe, 1987)

Ongoing meta-analysis

By initiating a systematic review soon after the introduction of a new intervention accompanied by regular updates, with the results of new studies followed by the statistical evaluation of a meta-analysis, the value of the intervention becomes more rapidly obvious. This has obvious advantages in therapeutics where the early recommendation of a reliable therapy may be made and where a therapy demonstrated to be useless, harmful or obsolete can be rapidly identified preventing its further use (Chalmers et al., 1992). Thus meta-analyses prevent a continued needless use of resources after existing data has demonstrated the effectiveness of an intervention (Chalmers et al., 1992). Finally the data available at the end of a meta-analysis may permit questions which had not been posed at the beginning of the individual trials to be addressed (Sacks et al., 1987; Peipert and Bracken, 1997) and the results of the meta-analysis may provide indications for future lines of research and help with the planning and design of studies (Altman and Elbourne, 1988; Huque, 1988)

Limitations of meta-analysis

Like any review, meta-analyses require suitable studies to be available for inclusion (Altman and Elbourne, 1988; Thacker et al., 1996). The quality of available studies will be variable but (Altman and Elbourne, 1988; Thacker et al., 1996) as many studies as possible should be identified to maximise the data available for review (L'abbe et al., 1987; Oxman 1994). It is also essential to ensure that the studies are "combinable", in terms of ensuring that they employ similar methodologies and examine samples drawn from similar populations.

Levels of evidence

Careful randomization is designed to eliminate bias between study groups by ensuring that investigators do not determine which patients enter the study or control groups thus determining the characteristics of the groups and influencing outcomes (Chalmers et al., 1983). As a result any difference between the groups is explained by chance (Altman, 1991). Variation in the quality of randomization has been acknowledged (Schultz et al., 1994). The random allocation of subjects for study in laboratory experiments is achievable (Altman, 1991) although much research in anatomical pathology consists of non-randomised cohort or case control studies and is

Table 1. Types of study design in decreasing order of like hood of bias level of evidence of treatment (Chalmers et al. 1987; Altman, 1991; Easterbrook et al., 1991; Oxman, 1994).

1. Randomised control trial - lower level of confidence limit exceeds significant benefit.
2. Randomised control trial - lower level of confidence limit does not exceed significant benefit.
3. Non-randomised cohort study.
4. Non-randomised cohort study - historical study.
5. Case control study.

centred on how particular markers differ in benign as opposed to malignant disease or in how they assist in predicting prognosis and response to therapy.

Levels of evidence for the effectiveness of study design have been proposed (Table 1). According to these, observational studies such as case control or cohort studies are deemed to be less persuasive (Chalmers et al., 1987; Altman, 1991; Oxman, 1994) and at a greater risk of bias than randomized studies. It is argued that they should be analysed separately from randomized studies and possibly should never be combined at all (Spitzer, 1991; Shapiro, 1994; Feinstein, 1995; Peipert and Bracken, 1997). It has also been suggested that since standard statistical techniques assume randomisation of the study sample, studies should be randomly selected for inclusion in a meta-analysis (Thacker, 1988).

Narrative review versus systematic review and meta-analysis

Conventional narrative reviews are usually constructed by experts with a wide knowledge of the field of interest who select the most appropriate, but by no means all, of the published data, for inclusion. The basis of this selection is not only on the standing of the journal in which the article has been published but is based on the expert's knowledge of the trustworthiness of the researchers and the reputation of their laboratories. Furthermore since they are familiar with research in the field, they are also aware of pitfalls in interpreting the data such as those which result from variation in experimental techniques or design (Eysenck, 1994). The most obvious danger in employing reviewers who are active in a particular field of research is that they will select papers which provide evidence to support a hypothesis which they favour. In addition it has been suggested that experts when asked to write a review spend less time preparing it than non-experts and write a review of inferior quality (Oxman and Guyatt, 1993). Whilst the final conclusion may not differ from those of a narrative review in which no further statistical tests have been applied, (Thacker, 1988) it has been suggested that unwarranted validity may be applied to the results of a meta-analysis because it appears to be a more objective and "scientific" pursuit (Thacker, 1988). Methods of defining the quality of meta-analysis (Thacker, 1988) and standardised reporting of these studies have been proposed (Sacks et al., 1987) and are outlined below.

Location and presentation of data

In common with any research paper the source of data, methods of identifying it, the data itself and methods of statistical analysis etc should be presented in the final written report in sufficient detail to permit the reader to repeat the study (Altman and Elbourne, 1988; Chalmers et al., 1992; Peipert and Bracken, 1997).

Organisation of systematic review and meta-analysis

Hypothesis

Meta-analyses are performed to test hypotheses (Thacker et al., 1996). In this report we will test two hypotheses. The first hypothesis is that a polyploid nuclear DNA content in cervical intraepithelial neoplasia grade I is not associated with progression to higher grade CIN or invasive carcinoma. This hypothesis has been chosen because it has been previously suggested as a suitable pathological topic for systematic review (Heatley, 1995, 1998) and because a small and manageable number of primary studies, which highlight some of the difficulties associated with systematic review and the process of attempting a meta-analysis, are available. Most of the article will deal with the methods used in examining this hypothesis. Ultimately none of the polypoid cases of CIN progressed. A second hypothesis that glandular differentiation in cervical carcinoma is associated with a poor prognosis therefore complements the first by demonstrating other aspects of the techniques of meta-analysis.

Systematic reviews and meta-analyses confined to small numbers of studies and patients are not unusual (Schulz et al., 1995).

Identification of studies

Meta-analysis requires the combination of existing studies. The systematic identification and selection of such studies using previously arranged inclusion and exclusion criteria is a laborious affair which like any primary research requires advance planning (Thacker et al., 1996).

Variations in the results of meta-analysis are often dependent on variations in their methodology and in particular on which studies the reviewer decides to include (Chalmers et al., 1987). The identification of all studies or as many as possible is limited if confined to standard computer searches such as the National Library of Medicine's database "Medline", which it is estimated permit the identification of only about 50% of studies, even though many more will have been published as papers in peer review journals (Chalmers et al., 1992), due to inadequacies in indexing and the failure of

Table 2. Criteria for inclusion in systemic review.

Prospective studies.
Colposcopy and histological confirmation of diagnosis.
In-situ squamous neoplasia of the cervix.
Follow-up period sufficient to assess potential for regression/ progression.
Quantitative analysis of nuclear DNA content.

Table 3. Journals handsearched in performing systemic review.

Journal of Clinical Pathology
Journal of Pathology
Histopathology
British Journal of Obstetric and Gynaecology
Americal Journal of Obstetrics and Gynaecology
Obstetrics and Gynaecology

authors to describe their own research methods. When performing meta-analyses many authors therefore extend their searches using a variety of methods such as hand searching of journals and following up reference lists in known studies. Problems are encountered with both methods. Handsearching of journals is time consuming and if practicable should be limited to a small number of relevant journals. Gotzsche (1987) mentions that in his study of double blind trials of two or more non steroidal anti-inflammatory drugs used in rheumatoid arthritis he found 200 relevant articles which had been published in 63 different journals! He was thus unable to refine a list of journals which should be targeted for further detailed hand searching.

Multiple publication of the same data in a series of references (for example, 60 papers and 20 abstracts have been traced to a single study (L'abbe et al., 1987)) or selection of only a proportion of available studies and failure to mention studies similar to the index paper adds to the potential for bias when searching reference lists.

Identification of studies of nuclear DNA content in CIN

The criteria developed to decide whether a study should be included in this review are listed in Table 2. Papers were identified using a medline search. In addition the index pages of three obstetrics and gynaecology and three anatomical pathology journals, six in all, published between 1987 and 1996 (Table 3) were hand searched for relevant papers. These journals were available in this institution and although the number of journals and the interval searched was restricted, the process took a total of 17 hours. Hand searching was facilitated in those journals where a list of papers published in each issue of the journal was included with the author and subject index at the end of each volume and where details of the methodology and size of the study was included in the article title. Whilst recognising the value of declarative titles (Underwood, 1997) the alternative of providing a descriptive title outlining methods used, the numbers of patients and the results saves the analyst much time in obtaining unhelpful references. Perhaps the use of an explanatory sentence after the title as demonstrated in the Contents page of journals such as "Obstetrics and Gynecology" and "Gynecological Oncology" represents a suitable compromise. In addition the reference lists of known papers were scoured.

Papers likely to be relevant to the review were obtained on inter-library loan.

Studies not published as full papers in peer review journals

A further difficulty in the identification of suitable data is that many studies are never published as full papers in peer review journals. The data may however be available in what is termed the "grey literature" - in short reports, abstracts and conference proceedings, technical reports, reports submitted by drug companies in pursuit of product licences or in government reports and in master and doctoral theses and dissertations (Thacker, 1988; Chalmers et al., 1990, 1992; Easterbrook et al., 1991; Dickersin et al., 1992). The inclusion of short reports and letters, which often do not have summaries included on the medline or in published articles is important as only 36% of these, proceed to full reports, the comparable figure for abstracts is about 50% (Goldman and Loscazlo, 1980; Chalmers et al., 1990; Scherer et al., 1994). Although some authors confine their reviews to published data (Chalmers et al., 1987) it is recognised that this under-represents all the data available in any field and that failure to include these studies from the "grey literature" may lead to radically different conclusions. For example, reviews of only published papers suggest that multi-drug cytotoxic therapy for ovarian carcinoma is of greater efficacy than single agent treatment. Inclusion of all available studies however shows no significant difference between the two methods of treatment (Simes, 1986). Some authors argue that data which has not been published in peer review journals is unreliable (Sacks et al., 1987).

A number of investigations have sought to explain why certain studies are ultimately published as full papers and why some are not. An association between the publication of results in peer review journals or acceptance for presentations at society meetings on the one hand and statistically significant or positive findings on the other has been described (Chan et al., 1982; Thacker, 1988; Dickersin et al., 1992; Scherer et al., 1994). Although other authors have failed to find an association between statistically significant results and publication, significant results were associated with a greater chance of multiple publication or publication in journals with a higher citation index (Easterbrook et al., 1991). Other factors which increase the probability of an

Table 4. Societies whose Abstract were reviewed in performing the systemic review.

Tha Pathological Society of Great Britain and Ireland.
The American Gynaecologial and Obstetrical Society.
The Blair Bell Research Society.
The British Gynaecological Cancer Society.
The Central Association of Obstetricians and Gynecologists.
The Pacific Coast Obstetrical and Gyecological Society.
The Society of Gynecological Surgeons.
The South Atlantic Association of Obstetricians and Gynecologists.

article's publication include having a sample size greater than the median (Easterbrook et al., 1991; Scherer et al., 1994), external funding; especially from government agencies such as the United States National Institute of Health (Dickersin et al., 1992; Dickersin and Min, 1993) and multiple data recruitment sites (Dickersin et al., 1992; Dickersin and Min, 1993). No association with the presence of a comparison group (Dickersin et al., 1992) or study type - observational versus clinical trial (Dickersin et al., 1992) has been demonstrated.

Although in one study papers with negative results were found to have a greater chance of being rejected than those with statistically significant results, despite often superior methodologies (Koren et al., 1989), follow up by a number of authors indicates that most studies which are presented for peer review are ultimately published (Easterbrook et al., 1991; Dickersin et al., 1992; Scherer et al., 1994) although submission to up to six journals may be needed (Scherer et al., 1994). In one series only 6 of 124 (4.8%) (Dickersin et al., 1992) and in another only 9% of repeatedly submitted papers were ultimately unsuccessful (Easterbrook et al., 1991). Thus, under-reporting in the peer review literature may be a function of the failure of authors to consider their data to be of importance (Easterbrook et al., 1991) and to submit it for consideration (Dickersin et al., 1987; Chalmers et al., 1992) rather than a reflection of publication bias by editors. Reasons cited by authors for failure to submit papers include methodological problems, failure to analyse the data (Easterbrook et al., 1991), the need to perform further statistical analyses (Dickersin and Min, 1993), negative results following statistical analysis (Easterbrook et al., 1991), results which the authors perceive as uninteresting and unlikely to be accepted for publication (Koren et al., 1989; Dickersin and Min, 1993) - a high importance rating by the investigator is an independent statistically significant variable on multivariate analysis in determining if the paper is published (Easterbrook et al., 1991), - problems with co-investigators (Dickersin and Min, 1993) and lack of time to prepare the report (Dickersin and Min, 1993). In some cases the study may have been performed only to enable the pharmaceutical company to gain a product licence (Easterbrook et al., 1991). Studies which had formed part of a doctoral or masters thesis were no more likely to be published as peer review papers (Dickersin et al., 1992).

Aside from suggesting that there is an ethical obligation to publish data (Chalmers, 1990; Dickersin et al., 1992) steps which will facilitate the identification of studies which have not been published include the establishment of the Cochrane Centre, and the Online Journal of Current Clinical Trials (Dickersin and Min, 1993; Scherer et al., 1994). Potential future developments include a suggestion that protocols for all drug trials should be published in peer review journals before commencement of the study (Piantadosi and Byar, 1988) and that all trials should be centrally registered (Chalmers et al., 1988; Meinert, 1988). National research registers, in some instances based on the ethical committees which exist in most countries and institutions have been established or proposed (Easterbrook, 1987; Dickersin et al., 1992; Dickersin and Min, 1993; Scherer et al., 1994). These initiatives could easily be extended to pathological research.

Identification of studies not published as full papers to peer review journals

In an attempt to identify additional sources of material the proceedings of a number of learned societies were screened (Table 4). No additional studies were identified.

Some authorities suggest that locating additional material may be facilitated by approaching individuals with an acknowledged interest in the field at scientific meetings or that investigators who have published papers on related subjects should be contacted in an effort to identify further unpublished material or to clarify details of published studies (Oxman, 1994). In my experience neither is a fruitful source of enquiry, in particular a total of 20 letters sent to the authors of publications, reporting nuclear DNA ploidy analysis in CIN lesions, requesting follow up information, yielded only four replies!

Selection criteria for inclusion of studies

Quality issues are not confined to data in the grey literature. Concerns regarding data from papers published in peer review journals include study size and the quality of study design especially in older studies (Gillman and Runyan, 1984). It has been shown that the chances of poorly planned and executed trials being published as full papers are increased if they show extreme results. As a result a large number of statistically significant studies each composed of a small number of patients may be published and may give a distorted impression of the effect of an intervention (Newcombe, 1987). Confining meta-analyses to studies which have been published in peer review journals will not remove the problems associated with including poor quality data. There is therefore a case for being selective in deciding which papers to include in the meta-analysis and excluding less rigorously executed studies, the results of which may be less generalizable (Thacker,

1988).

In many instances it is not possible for a single investigator or group to obtain a sizeable sample for study. Failure to perform and publish such studies would however deprive the literature of the data which they provide. An alternative to refusing to publish these papers with statistically significant results based on small number of patients is to ensure that studies with small samples which have yielded negative findings are also published (Dickersin et al., 1987).

Since it has been suggested that most systematic reviews or meta-analyses consist of data originating from a small number of large studies, the inclusion or exclusion of studies comprising small numbers of patients has little effect on the ultimate conclusion of the review (Chalmers et al., 1987). Indeed some authorities recommend that the number of unpublished studies or cases which would be necessary to convert statistically significant data to statistical insignificance should be established. If the number is small then there would be cause for concern (L'abbe et al., 1987).

Determining quality assessment criteria

To overcome all of these problems it is recommended that a quality assessment protocol should be generated (Peipert and Bracken, 1997) which consists of the minimum criteria necessary for a study to be included in the systematic review or meta-analysis (Newcombe, 1987). To some extent this reflects the situation with primary research where data may be excluded because it is incomplete or because subjects are unsuitable - they may be too old, too young or of the wrong sex for inclusion in the study. Quality criteria for the inclusion/exclusion of such studies have been proposed (Peipert and Bracken, 1997). To avoid confirmatory bias, that is the tendency to only select those papers which confirm one's own views (Thacker, 1988; Koren et al., 1989), it is usually recommended that the selection of papers be based on scrutiny of the Materials and methods sections. The criteria for inclusion in the study are determined (Altman and Elbourne, 1988; Thacker, 1988; Oxman, 1994; Thacker et al., 1996; Peipert and Bracken, 1997) and the protocol prepared (Sacks et al., 1987; Peipert and Bracken, 1997) before the papers are read. Some investigators go so far as to anonymise photocopies of the Materials and methods sections of the papers under consideration (L'abbe et al., 1987; Altman and Elbourne, 1988; Thacker, 1988; Oxman, 1994; Peipert and Bracken, 1997) by removing any means of identification except for a code number and having a single reviewer (Schultz et al., 1995) or two independent reviewers assess the paper's suitability for inclusion (Chalmers, 1987; Altman and Elbourne , 1988; Thacker, 1988; Thacker et al., 1996; Peipert and Bracken, 1997). Discrepancies or differences are resolved in conference.

By avoiding any opportunity to view the results the reviewer is not unduly influenced by apparently promising results and only methodologically sound papers are included (Altman and Elbourne, 1988). Whilst this council of perfection is to be applauded it is dependent upon the discipline of authors and editors in ensuring that the Materials and methods section of every paper is a complete account of how the study was performed and an accurate predictor of what results can be expected. In reviewing the papers for this study I found that this was not the case and that it was often necessary to turn to the summaries and Results sections to glean further information thus avoiding the loss of useful data gathered in a methodologically sound study.

Details of the criteria for inclusion/exclusion of studies should be included in the final report along with a list of all the papers screened. The reasons for the inclusion or exclusion of each individual paper should be stated (Peipert and Bracken, 1997).

Inclusion criteria for this study

The criteria for inclusion of studies in this review are listed (Table 2). Following a computerised literature search using the Medline database handsearching six journals (Table 3) published between 1987 and 1996, and searches of the reference lists of the papers thus identified, 45 studies of potential interest were located (Table 5).

Outcomes

As recommended all the papers reviewed and the reasons for rejecting any are given in Table 5 (Sacks et al., 1987; Peipert and Bracken, 1997). Careful screening of the 45 papers enabled the exclusion of the great majority. In most instances the papers dealt exclusively with invasive carcinomas and were outside the scope of the study. In papers dealing with CIN lesions many did not include a sufficiently long period of follow-up between diagnosis of the patient's CIN lesion and treatment to enable an assessment as to whether the disease was likely to resolve or not. In some instances the studies had been carried out on definitive therapeutic specimens. It should be borne in mind however that many of the papers assessed in this review described comparisons between different groups and were not primarily designed to provide follow-up data (Oxman, 1994).

At the end of this preliminary selection exercise four studies had met the basic criteria for inclusion. Further assessment was carried out to ensure that the studies were similar enough to be combined (Peipert and Bracken, 1997). Since different studies may use a variety of methodologies (Peipert and Bracken, 1997) it was necessary to ensure that it was possible or appropriate to pool the studies for statistical analysis (Thacker et al., 1996). Of the four candidate studies which met the initial criteria, two were large studies one of 100 patients and the other of over 300 patients. The third was a study of 32 patients and the fourth a study of 59 patients in which

Table 5. Studies considered for inclusion in systemic review.

AUTHOR	SELECTED/EXCLUDED	REASON FOR INCLUSION(COMMENTS
1. Barry Walsh et al., 1993	Exclude	Invasive carcinoma - case report
2. Bibbo et al., 1989	Select	Follow up of 302 in-situ cases
3. Bocking et al., 1986	Exclude	Reviews established invasive cases. Follow up less than 3 months
4. Chacho et al., 1990	Exclude	No follow up available
5. Clavel et al., 1992	Exclude	No follow up available
6.Connor et al., 1993	Exclude	Invasive carcinoma - series
7. De Vita et al., 1990	Select	Follow up in 4 of 59 cases
8. Dudzinsky et al., 1987	Exclude	No follow up available for CIN patients.
9. Elias Jones et al., 1986	Exclude	No follow up/Technical paper
10. Evans and Monaghan 1983	Exclude	No follow up
11. Fletcher et al., 1991	Exclude	Based on study of cell lines. No follow up
12. Freni, 1975	Exclude	No follow up
13. Fu et al., 1978	Select	
14. Fu et al., 1981	Select	
15. Fu et al., 1982a,b	Exclude	Invasive adenocarcinoma series
16. Fu et al., 1983	Exclude	No follow up available
17. Fujii et al., 1984	Exclude	No follow up available
18. Hanselaar et al., 1988	Exclude	Review of established invasive cases
19. Hanselaar et al., 1990	Exclude	Compares CIN 3 with and without invasion. Not a follow up study
20. Hanselaar et al., 1991	Exclude	No follow up
21. Hanselaar et al., 1992	Exclude	Study based on cytological findings - no colposcopic/histologic verification
22. Hendy Ibbs et al., 1987	Exclude	No follow up
23. Hasegawa et al., 1975	Exclude	No follow up
24. Hughes et al., 1987	Exclude	No follow up data
25. Jacobsen et al., 1983a	Exclude	Invasive carcinoma
26. Jacobsen et al., 1983b	Exclude	No follow up
27. Kaern et al., 1992	Exclude	Study includes invasive carcinomas from sites not confined to the uterine cervix
28. Kirkland et al., 1967	Exclude	Study based on chromosomal counts, not DNA quantitation
29. Kristensen, 1995	Exclude	Invasive carcinoma
30. Minagawa et al., 1993	Exclude	Invasive carcinoma
31. Nasiell et al., 1979	Select	
32. Naslund et al., 1987	Unsuitable	Animal study
33. Ng and Atkin, 1973	Exclude	Invasive carcinoma
34. Pascale Segers et al., 1995	Exclude	No follow up
35. Rihet et al., 1996	Exclude	No follow up
36. Sandritter et al., 1996	Exclude	Invasive carcinoma
37. Schevchuk and Richart, 1986	Exclude	No follow up
38. Steinbeck et al., 1995	Exclude	No follow up
39. Van Driel Kulke and Ploem-Zaaijer, 1989	Exclude	No follow up
40. Van Leuwen et al., 1996	Exclude	No follow up
41. Tsou et al., 1984	Exclude	No follow up
42. Watts et al., 1987	Exclude	No follow up
43. Wilkbanks et al., 1967	Exclude	No follow up
44. Winkler et al., 1984	Exclude	No follow up
45. Zolzer et al., 1995	Exclude	Invasive carcinoma

follow-up is available for only 4.

Whilst there is no reason for excluding the study with follow up in 4 patients on the basis of a small sample size alone, other variables such as the fact that follow-up was available in a minority of patients all of whom had high grade CIN lesions placed it outside the scope of the hypothesis. The importance of ensuring that patients with diseases which have different causal factors are not compared is also important (Eysenck, 1994). The study of 32 patients by Fu et al. (1978) was based on patients who were exposed to diethylstilbestrol (DES) in utero. Since DES is known to predispose to malignancy, especially adenocarcinomas, of the female genital tract there would appear to be a prima facie objection to including these data. Although inclusion of this study would increase the number of women with polyploid CIN 1 to 60, an approximately twofold increase in incidence of squamous CIN has been described in DES exposed women (Robboy et al., 1984). For these biological rather than statistical reasons this study was also excluded. A further potential objection is that since this study shares some of the authors of another study (Fu et al., 1981), it is possible that the same patients had been included in both study groups. As a result their effects or findings could be duplicated exaggerating bias. On examination of the latter study the authors indicate that it consisted of a distinct sample of patients. The common authorship was not therefore a reason to exclude one or other study.

Although there was some variation in terminology

Table 6. Numbers of cases of polyploid and aneuploid cases of low and high grade CIN which persisted/regressed of progressed to higher grade disease in three studies.

		POLYPLOID		ANEUPLOID	
		Persisted/regressed	Advanced	Persisted/regressed	Advanced
Fu et al., 1978	CIN 1	18	0	1	0
	CIN 2/3	1	0	3	0
Fu et al., 1981	CIN 1	30	0	6	14
	CIN 2/3	2	0	30	18
Bibbo et al., 1989	CIN 1	12	0	17	0
	CIN 2/3	41	18	85	24

used to describe the grades of intraepithelial neoplasia, consistency of grouping (Sacks et al., 1987; Eysenck, 1994) was possible in each paper by classifying the patients studied into high grade (moderate or severe dysplasia, CIN 2 or 3) and low grade (mild dysplasia, CIN 1). Review indicated that similar methods of DNA quantitation were used in both papers. Although it is recognised that studies using different techniques may need to be included in meta-analyses to increase their statistical power and provide additional opportunities to study the effects of interventions, Hanselaar et al., (1991) found that data may not be interchangeable between techniques of DNA quantitation. Peipert and Bracken (1997) have highlighted the problems which may be caused if studies follow different methodologies and emphasises the importance of ensuring that protocols are combinable (Sacks et al., 1987; Peipert and Bracken, 1997) with objective measurements of heterogeneity being advised (Sacks et al., 1987; Thacker et al., 1996).

None of the 42 (0%) cases of CIN 1 which was polyploid in the two remaining studies progressed to a higher grade lesion although 14 of 37 (38%) aneuploid cases did (Table 6). Confidence intervals of 0-7.1% for the percentage of polyploid CIN 1 lesions which may be expected to advance can be calculated using this data. (Had the DES exposed patients been included the confidence interval would be 0-5%). It is possible also to calculate that it would be necessary to find no evidence of progression in a sample size in excess of about 300 cases of polyploid CIN 1 to predict that the risk of progression in the general population was under 1% (ie confidence interval 0-1%) (Table 6).

Other methods for presenting results

The pooling of data from the studies in a meta-analysis does not necessarily involve the combination of the actual data but is accomplished by calculating differences of effect for each study in the form of an odds ratio with confidence intervals followed by the construction of a weighted estimate of the effects of the intervention (Altman and Elbourne, 1988). As a result data is presented study by study rather than as for a combined sample. This may be presented graphically either in the form a linear plot graph or more usually in the form of a bar chart (often termed a "blobogram") with evidence of a positive correlation on one side and evidence of a negative correlation on the other side of a midline indicator of no effect (L'abbe et al., 1987; Altman and Elbourne, 1988; Thacker et al., 1996; Peipert and Bracken, 1997). Reflecting a process often adopted in narrative review, simple vote counting, that is the description of the number of studies with evidence of a positive correlation as opposed to the number with a negative correlation, may be utilized. Most meta-analyses however include a statistical analysis, the most usual test performed being the Mantel Haenszel test (L'abbe et al., 1987; Altman and Elbourne, 1988; Peipert and Bracken, 1997) but with Chi-squared, regression analysis and others being performed on occasion (L'abbe et al., 1987).

Several factors may have an effect on outcomes, and in this situation, multivariate analysis may be more appropriate than univariate analysis (Eysenck, 1984, 1994; L'abbe et al., 1987). The appropriateness of the statistical tests used has been questioned since these tests are designed for analysis of primary data (L'abbe et al., 1987; Thacker, 1988; Eysenck, 1994; Peipert and Bracken, 1997) and the development of statistical methods specifically for meta-analyses (in particular tests which allow combination of studies other than randomized control trials) has been proposed (Thacker, 1988).

The description of confidence intervals or standard deviations, as in the above example, is also encouraged (L'abbe et al., 1987; Newcombe, 1987; Sacks et al., 1987; Altman and Elbourne, 1988; Thacker, 1988, 1996; Oxman, 1994) although those in which the confidence interval for the size of the effect excludes zero are likely to be preferred (Newcombe, 1987).

Confounding factors in meta-analysis

Tests of heterogeneity between studies are advisable to ensure one is not combining fundamentally different studies, a process which has been likened to adding apples and oranges (Eysenck, 1984, 1994; Thacker, 1988; Peipert and Bracken, 1997). Statistical analyses should also address the extent to which the study results depend upon variations in characteristics (L'abbe et al., 1987). Interventions are likely to vary from study to

Table 7. Sensitivity analysis examines the effect of:

- Changing inclusion criteria
- Effects of including/excluding trials which do not satisfy selection criteria
- Effects including/excluding unpublished studies
- Reanalysing data using a reasonable range of results where the data are not available
- Reanalysis using alternative statistical methods

Table 8. Criteria for inclusion of series in a meta-analysis of survival in squamous cell versus adenocarcinoma and adenosquamous carcinoma of the cervix.

Data availabe for squamous cell and adeno/adenosquamous carcinoma.
Mucin stains performed in cases which morphologically appear to be squamous cell carcinomas to identify squamous carcinomas with mucin secretion (adenosquamous carcinomas).
Survival/mortality data available for five or more years after diagnosis.
Data presented in numerical form in text or table.
Details of tumour stage included.

study, for example in a pathological study involving immunohistochemistry, the concentration of primary antibody or the incubation conditions may vary. Attempts to correlate results with variations in such study characteristics and to determine how sensitive the results are to the way in which the study is carried out are advised (Oxman, 1994; Peipert and Bracken, 1997). This is termed sensitivity analysis and the points which should be included in such an analysis are listed in Table 7 (Chalmers et al., 1987; Sacks et al., 1987; Oxman, 1994; Thacker et al., 1996; Peipert and Bracken, 1997).

Practical demonstration of tests of heterogeneity, graphical representation of results and the Mantel-Haenszel test

Squamous cell and adenocarcinoma of cervix

To demonstrate a situation in which there was no evidence of heterogeneity, a review of published series of cervical carcinoma was carried out to examine the hypothesis that the prognosis in squamous cell carcinoma in which there is no evidence of glandular differentiation is better than in adenocarcinomas and in squamous cell carcinomas in which there is evidence of glandular differentiation as demonstrated morphologically or with mucin stains (squamous carcinoma with mucin secretion/adenosquamous carcinomas). The criteria for inclusion are presented in Table 8 the studies identified are detailed in Table 9 with the reasons for exclusion if appropriate. The results of the three series studied are given in Table 10. Although one of these has a relatively high survival rate the standard test for heterogeneity is nonsignificant indicating that it is not unreasonable to conduct a meta-analysis. The relative risks for mortality are illustrated in the Figure 1. The

Table 9. Reasons for excluding studies of prognostic features in adenocarcinoma of the cervix.

Kilgore et al., 1988	Mucin stains not used, data not presented in numerical/tabular form
Langlois et al., 1996	Data not given in numerical/tabular form
Hale et al., 1991	Results not presented as table, stages 1b/2a combined.
Harrison et al., 1993	Only 67% of patients had 5 year survival
Terada et al., 1988	Mucin stains not used
Colgan et al., 1993	No follow up data
Fu et al., 1982	Data for squamous cell carcinoma not included
Schorrock et al., 1990	No follow up or staging data
Gallup et al., 1985	Follow up available for only 2 years in some cases
Greer et al., 1989	Follow up less than 5 years in some cases
Hopkins et al., 1988	Mucin stains not done
Ireland et al., 1987	Follow up for 3 years only
Randal et al., 1988	Mucin stain not performed
Saigo et al., 1986	Data for squamous carcinoma not included
Yazigi et al., 1990	Deals with disease free not absolute survival rate
Buckley et al., 1988	Deals with regional nodal and distant recurrence not surevival rate

Table 10. Five year survival rate in patients with squamous cell carcinoma compared to patients with glandular differentiation in their carcinoma.

	SQUAMOUS CELL CARCINOMA			GLANDULAR DIFFERENTIATION		
	Survived	Died	Total	Survived	Died	Total
Bethwaite et al., 1992	77	26	103	20	19	39
Benda et al., 1985	32	3	35	16	1	17
Wheeless et al., 1970	242	191	433	24	34	58

Mantel-Haenszel weighted relative risk of mortality (squamous relative to adeno) for the three studies combined is 0.69 (95% confidence intervals 0.56 - 0.86). Thus the risk of mortality for patients with squamous carcinoma in 69% of that of patients with adenosquamous carcinoma.

Discussion

The discussion of any meta-analysis should complete the analysis begun with the presentation of data and their statistical analysis, in the same manner as the discussion of any research paper. As with a conventional paper, these should be discussed in the context of the existing literature on the subject (Thacker, 1988; Oxman, 1994; Thacker et al., 1996). Reflecting the hybrid nature of these papers, there are advantages in having an individual with recognized expertise in the field participate in the construction of the discussion as he can evaluate methodological weaknesses and assess the validity of the data in the context of the existing

literature (Eyseenck, 1994). His expertise will eg provide a knowledge of relevant basic science research and may provide an explanation for the findings (Oxman, 1994) or indicate alternative explanations for them (Thacker, 1988).

Many of the factors which are usually included in a discussion section have already been dealt with in the methods and results sections of this paper. In this study a strict adherence to preselected criteria reduced the number of studies available for inclusion to two in the first review and three in the second. Whilst less rigid criteria or a less strict adherence to the criteria may have increased the number of studies available for inclusion they would also have increased the potential for bias. One of the papers recovered (Hanselaar et al., 1992) consisted of retrospective reviews of cytology in patients who had had established invasive carcinomas of the cervix diagnosed some years later. Although these data might have been included in a narrative review they were excluded from this study because they did not meet the criteria outlined in Table 2, in particular the prospective nature of the surveys, for inclusion. Furthermore the study was based on cytological specimens originally diagnosed as negative or showing low grade abnormalities in patients found later to have micro-invasive stage 1A or stage 1B squamous cell carcinoma, an average of 30 to 33 months later. Colposcopy had not been carried out at the time of initial presentation for screening and the full range of abnormalities on the cervix may not have been visualized. Thus it is possible that the original smear may not have sampled an existing high grade or even invasive lesion.

Although subgroup analysis may be of value, interpretation of these results requires particular care since it may result in some patients being denied an intervention or because it may misdirect research resulting in the waste of resources (Oxman and Guyatt, 1992). Some authors advocate reliance on the overall results to indicate the likely effect on a particular subgroup. Criteria for subgroup analyses have been proposed (Yusuf et al., 1991). Any conclusions and recommendations for future changes in practice should be in proportion to the strength of the evidence provided (Oxman, 1994) and should include an assessment of all

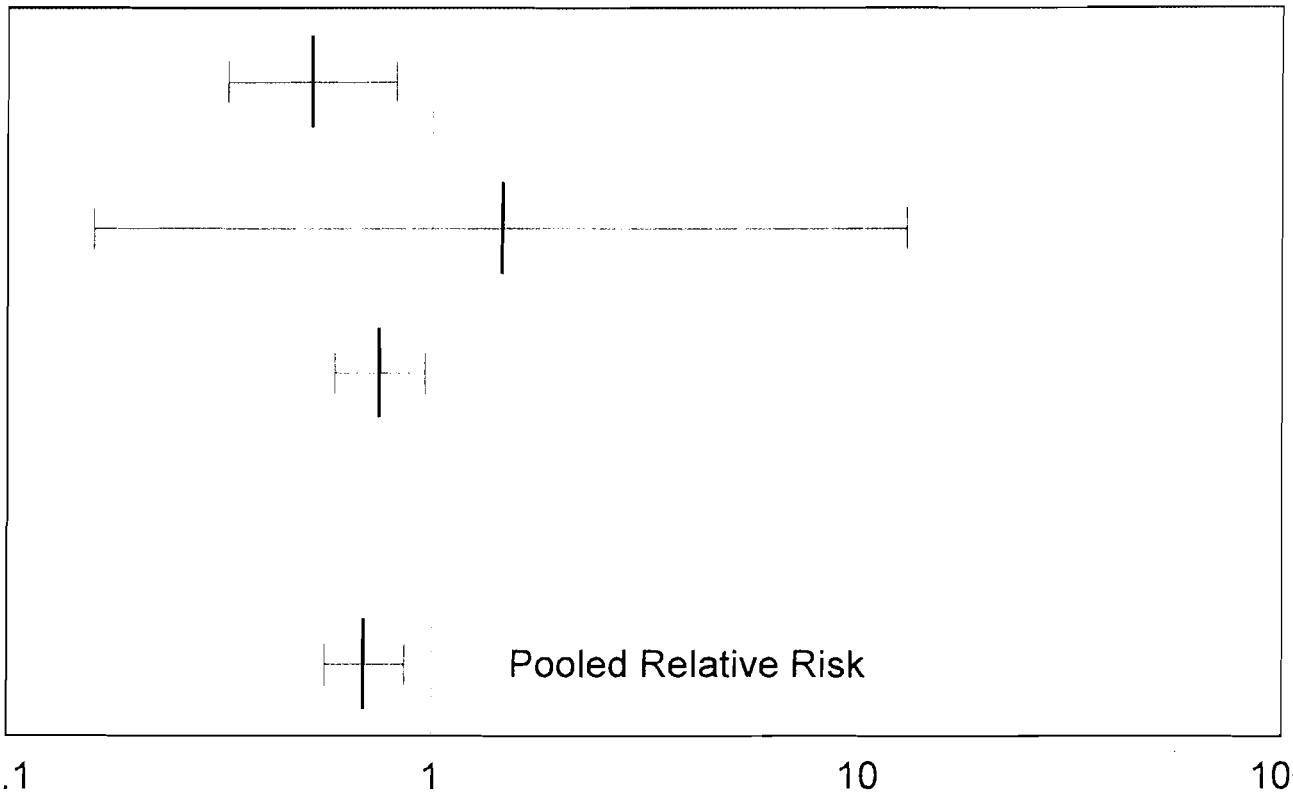

Fig. 1. This "blobogram" demonstrates the relative risk of death for patients described in three studies (Wheless et al., 1970; Benda et al., 1985; Bethwaite et al., 1992). The index line at the centre of the diagram identifies an equal risk, the point at which series in which the risk of death for patients with squamous cell carcinoma was the same as that for adenocarcinoma would lie. In two studies the risk of death in patient with squamous cell carcinoma was less than that for patients with adenocarcinoma as demonstrated by the heavy line being drawn to the left of the index line. In the remaining study in which the risk of death was greater for patients with squamous cell carcinoma is denoted by a heavy line to the right of the index line. Confidence intervals are given for each study. The overall risk when the studies have been combined is denoted as the pooled relative risk.

likely side effects of the intervention (Oxman, 1994) eg in this instance a prospective study of at least 300 patients with polyploid CIN 1 should be followed to ensure that there is no evidence of disease progression.

Finally whilst meta-analyses may include the synthesis of the best available data, it must be remembered that this may not be universally applicable outside a study setting especially in community practice (Thacker, 1988).

Acnowledgements. I am grateful to Mrs. C. Davies for typing the manuscript, to Mr. P. Ewings of the Research and Development Unit for statistical assistance and to the staff of the Nuffield Library, Taunton and Somerset NHS Trust for providing the many references used in preparing th paper.

References

Altman D.G. (1991). Randomisation. Essential for reducing bias. Br. Med. J. 302, 1481-1482.

Altman D.G. and Elbourne D. (1988). Commentary. Combining results from several clinical trials. Br. J. Obstet. Gynaecol. 95, 1-2.

Barry Walsh C., Kay E., Prendiville W., Turner M. and Leader M. (1993). Lymphoepithelioma-like carcinoma of the uterine cervix with c-erbB-2, p53 oncoprotein expression and DNA quantification. Histopathology 23, 592-593.

Benda J.A., Platz C.E., Buchsbaum H. and Lifshitz S. (1985). Mucin production in defining mixed carcinoma of the uterine cervix: a clinicopathologic study. Int. J. Gynecol. Pathol. 4, 314-327.

Bethwaite P., Yeong M.L., Holloway L., Robson B., Duncan G. and Lamb D. (1992). The prognosis of adenosquamous carcinomas of the uterine cervix. Br. J. Obstet. Gynecol. 99, 745-750.

Bibbo M., Dytch H.E., Alenghat E., Bartels P.H. and Wied G.L. (1989). DNA ploidy profiles as prognostic indicators in CIN lesions. Am. J. Clin. Pathol. 92, 261-265.

Bocking A., Hilgarth M., Auffermann W., Hac-Werdier C., Fischer-Becker D. and von Kalkreuth G. (1986). DNA-cytometric diagnosis of prospective malignancy in borderline lesions of the uterine cervix. Acta Cytol. 30, 608-615.

Buckley C.H., Beards C.S. and Fox H. (1988). Pathological prognostic indicators in cervical cancer with particular reference to patients under the age of 40 years. Br. J. Obstet. Gynecol. 95, 47-56.

Chacho M.S., Eppich E., Wersto R. and Koss L.G. (1990). Influence of human papillomavirus on DNA ploidy determination in genital condylomas. Cancer 65, 2291-2294.

Chalmers T.C., Celano P., Sacks H.S. and Smith H. (1983). Special article. Bias in treatment assignment in controlled clinical trials. N. Engl. J. Med. 309, 1358-1361.

Chalmers T.C., Berrier J., Sacks H.S., Levin H., Reitman D. and Nagalingam R. (1987). Meta-analysis of clinical trials as a scientific discipline: II: Replicate variability and comparison of studies that agree and disagree. Stat. Med. 733, 744.

Chalmers T.C., Simes R.J., Dickersin K., Boissel J-P. and Meinert C.L. (1988). Report from the panel on the case for registers of clinical trials at the eighth annual meeting of the Society for Clinical Trials. Control. Clinic. Trials 9, 76-81.

Chalmers I. (1990). Under-reporting research in scientific misconduct. JAMA 263, 1405-1408.

Chalmers I., Adams M., Dickersin K., Hetherington J., Tarnow-Mordi W., Meinert C., Tonascia S. and Chalmers T.C. (1990). A cohort study of summary reprots of controlled trials. JAMA 263, 1401-1405.

Chalmers I., Dickersin K. and Chalmers T.C. (1992). Getting to grips with Archie Cochrane's agenda. Br. Med. J. 305, 786-787.

Chan S.S., Sacks H.S. and Chalmers T.C. (1982). The epidemiology of unpublished randomized control trials. Clin. Res. 30, 16.

Clavel C., Zerat L., Binninger I., Boutterin M-C., Polette M., Monsonego J. and Birembaut P. (1992). DNA content measurement and in situ hybridization in condylomatous cervical lesiosn. Diagn. Mol. Pathol. 3, 180-184.

Colgan T.J., Auger M. and McLaughlin J.R. (1993). Histopathologic classification of cervical carcinomas and recognition of mucin secreting squamous caricnoma. Int. J. Gynecol. Pathol. 12, 64-69.

Connor J.P., Miller D.S., Bauer K.D., Murad T.M., Rademaker A.W. and Lurain J.R. (1993). Flow cytometric evaluation of early invasive cervical cancer. Obstet. Gynecol. 81, 367-371.

De Vita R., Calugi A., Maggi F., Mauro F., Montevecchi L. and Vecchione A. (1990). Flow cytometric DNA analysis of the human cervix affected by human papillomavirus and/or intraepithelial neoplasia. Anal. Quant. Cytol. Histol. 12, 306-313.

Dickersin K. and Min Y.I. (1993). Publication bias: the problem that won't go away. Ann. NY Acad Sci. 703, 135-148.

Dickersin K., Min Y-I. and Meinert C.L. (1992). Factors influencing publication of research results. JAMA 267, 374-378.

Dickersin K., Chan S., Chalmers T.C., Sacks H.S. and Smith H. Jr. (1987). Publication bias and clinical trials. Control. Clinic. Trials. 8, 343-353.

Dudzinski M.R., Haskill S.J., Fowler W.C., Currie J.L. and Walton L.A. (1987). DNA content in cervical neoplasia and its relationship to prognosis. Obstet. Gyneocl. 69, 373-377.

Easterbrook P. (1987). Reducing publication bias. Br. Med. J. 295, 1347.

Easterbrook P., Berlin J.A., Gopalan R. and Matthews D.R. (1991). Publication bias in clinical research. Lancet 337, 867-872.

Elias-Jones J., Hendy-Ibbs P., Cox H., Evas G.I. and Watson J.V. (1986). Technical method - cervical brush biopsy specimens suitable for DNA and oncoprotein analysis using flow cytometry. J. Clin. Pathol. 39, 577-581.

Evans A.S. and Monaghan J.M. (1983). Nuclear DNA content of normal, neoplastic and "wart-affected" cervical biopsies. Anal. Quant Cytol. 5, 112-116.

Eysenck H.J. (1984). Meta-analysis: an abuse of research integration. The Journal of special education 18, 41-59.

Eysenck H.J. (1994). Meta-analysis and its problems. Br. Med. J. 309, 789-792.

Feinstein A.R. (1995). Meta-analysis: statistical alchemy for the 21st century. J. Clin. Epidemiol. 48, 71-79.

Fletcher S., Neill W.A. and Norval M. (1991). Seminal polyamines as agents of cervical carcinoma: production of aneuploidy in squamous epithelium. J. Clin. Pathol. 44, 410-415.

Freni S.C. (1975). Non-visual prescreening of cervical smears with a flow-through cytophotometer. Acta Cytol. 19, 448-452.

Fu Y.S., Robboy S.J. and Prat J. (1978). Nuclear DNA study of vaginal and cervical squamous cell abnormalities in DES exposed progeny. Obstet. Gynecol. 52, 129-137.

Fu Y.S., Reagan J.W. and Richart R.M. (1981). Definition of precursors. Gynecol. Oncol. 12, S220-S231.

Fu Y.S., Reagan J.W., Fu A.S. and Janiga K.E. (1982a).

Adenocarcinoma and mixed carcinoma of the uterine cervix - II Prognostic value of nuclear DNA analysis. Cancer 49, 2571-2577.

Fu Y.S., Reagan J.W., Hsiu J.G., Storaasli J.P. and Wentz W.B. (1982b). Adenocarcinoma and mixed carcinoma of the uterine cervix 1: a clinicopathologic study. Cancer 49, 2560-2570.

Fu Y.S., Braun L., Shah K.V., Lawrence W.D. and Robboy S.J. (1983). Histologic, nuclear DNA, and human papillomavirus - studies of cervical condylomas. Cancer 52, 1705-1711.

Fujii T., Crum C.P., Winkler B., Fu Y.S. and Richart R.M. (1984). Human papillomavirus infection and cervical intraepithelial neoplasia: histopathology and DNA content. Obstet. Gynecol. 63, 99-104.

Gallup D.G., Harper R.H. and Stock R.J. (1985). Poor prognosis in patients with adenosquamous cell carcinoma on the cervix. Obstet. Gynecol. 65, 416-421.

Gillman M.W. and Runyan D.K. (1984). Bias in treatment assignment in controlled clinical trials. N. Engl. J. Med. 310, 1610-1611.

Goldman L. and Loscazlo A. (1980). Fate of cardiology research originally published in abstract form. N. Engl. J. Med. 305, 255-259.

Gotzsche P. (1987). Reference bias in reports of drug trials. Br. Med. J 295, 654-656.

Greer B.E., Figge D.C., Tamini H.K. and Cain J.M. (1989). Stage 1B adenocarcinoma of the cervix treated by radical hysterectomy and pelvic lymph node dissection. Am. J. Obstet. Gynbecol. 160, 1509-1514.

Hale R.J., Wilcox F.L., Buckely C.H., Tindall V.R., Ryder W.D.J. and Logue J.P. (1991). Prognostic factors in uterine cervical carcinoma: a clinicopathological analysis. Int. J. Gynecol. Cancer 1, 19-23.

Hanselaar A.G.J.M., Vooijs G.P., Oud P.S., Pahlpatz M.M.M. and Beck J.L.M. (1988). DNA ploidy patterns in cervical intraepithelial neoplasia grade III, with and without synchronous invasive squamous cell carcinoma. Cancer 62, 2537-2545.

Hanselaar A.G.J.M., Vooijs G.P., Van't Hof-Grootenboer A.E. and Pahlplarz M.M.M. (1990). Cytophotometric analysis of cervical intraepithelial neoplasia grade III, with and without synchronous invasive squamous cell carcinoma. Cytometry 11, 901-906.

Hanselaar A.G.J.M., Vooijs G.P., Van't Hof-Grootenboer A.E., Gemmink J.H., De Leeuw H. and Pahlplatz M.M.M. (1991). Cytophotometric analysis of corresponding cytological and histological cervical intraepithelial neoplasia grade III specimens. Cytometry 12, 1-9.

Hanselaar A.G.J.M., Vooijs G.P., Mayall B.H., Pahlplatz M.M.M. and Van't Hof-Grootenboer A.E. (1992). DNA changes in progresive cervical intraepithelial neoplasia. Anal. Cell Pathol. 4, 315-324.

Harrison T.A., Sevin B.U.S., Koechli O., Nguyen H.N., Averette H.F., Penalver M., Donato D.M. and Nadji M. (1993). Adenosquamous carcinoma of the cervix: prognosis in early stage disease treated by radical hysterectomy. Gynecol. Oncol. 50, 310-315.

Hasegawa T., Tsutsui F. and Kurihara S. (1975). Cytomorphologic study on the atypical cells following cryosurgery for the treatment of chronic cervicitis. Acta Ctyol. 19, 533-537.

Heatley M.K. (1995). Proliferation in the normal cervix and in preinvasive cervical lesions. J. Clin. Pathol. 49, 957.

Heatley M.K. (1998). What is the value of proliferation markers in the normal and neoplastic cervix. Histol. Histopathol. 13, 249-254.

Hendy-Ibbs P., Cox H., Evan G.I. and Watson J.V. (1987). Flow cytometric quantitation of DNA and c-myc oncoprotein in archival biopsies of uterine cervix neoplasia. Br. J. Cancer 55, 275-282.

Hopkins M.P., Schmidt R.W., Roberts J.A. and Morley G.W. (1988). The prognosis and treatment of Stage 1 adenocarcinoma of the cervix. Obstet. Gynecol. 72, 915-921.

Hughes R.G., Neill W.A. and Norval M. (1987). Nuclear DNA analysis of koilocytic and premalignant lesions of the uterine cervix. Br. Med. J. 294, 267-269.

Huque M.F. (1988). Experiences with meta-analysis in NDA submissions. 1988. Proceedings of the Biopharmaceutical Section for the American Statistical Society 2, 28-33.

Ireland D., Cole S., Kelly P. and Monaghan J.M. (1987). Mucin production in cervical intraepithelial neoplasia and in stage 1b carcinoma of cervix with pelvic lymph node metastases. Br. J. Obstet. Gynecol. 94, 467-472.

Jakobsen A., Nielsen K.V. and Ronne M. (1983a). DNA distribution and chromosome number in human cervical carcinoma. Anal. Quant. Cytol. 5, 13-18.

Jakobsen A., Kristensen P.B., Poulsen H.K. (1983b). Flow cytometric classification of biopsy specimens from cervical intraepithelial neoplasia. Cytometry 4, 166-169.

Kaern J., Wetteland J., Trope C.G., Farrants G.W., Juhng S-W., Pettersen E.O., Reith A. and Danielsen H.E. (1992). Comparison between flow cytometry and image cytometry in ploidy distribution assessments in gynecologic cancer. Cytometry 13, 314-321.

Kilgore L.C., Soong S.-J., Gore H., Shingleton H.M., Hatch K.D. and Partridge E.E. (1988). Analysis of prognostic features in adenocarcinoma of the cervix. Gynecol. Oncol. 31, 137-148.

Kirkland J.A., Stanley M.A. and Cellier K.M. (1967). Comparative study of histologic and chromosomal abnormalities in cervical neoplasia. Cancer 20, 1934-1951.

Koren G., Shear H., Graham K. and Einarson T. (1989). Bias against the null hypothesis: the reproductive hazards of cocaine. Lancet 2, 1440-1442.

Kristensen G.B., Kaern J., Abeler V.M., Hagmar B., Trope C.G. and Pettersen E.O. (1995). No prognostic impact of flow-cytometric measured DNA ploidy and S-phase fraction in cancer of the uterine cervix: a prospective study of 465 patients. Gynecol. Oncol. 57, 79-85.

L'Abbe K.A., Detsky A.S. and O'Rourke K. (1987). Meta-analysis in clinical research. Ann. Intern. Med. 107, 224-233.

Langlois N.E.I., Ellul B. and Miller I.D. (1996). A study of the value and prognostic significance of mucin staining in squamous cell carcinoma of the uterine cervix. Histopathology 28, 175-178.

Meinert C.L. (1988). Towards prospective registration of clinical trials. Control. Clinic. Trials 9, 1-5.

Minagawa Y., Kigawa I., Kanamori I., Itamochi H. and Terakawa N. (1993). Tumor cell kinetics in elderly patients with cervical cancer. Obstet. Gynecol. 81, 610-614.

Nasiell K., Auer G., Nasiell M. and Zetterberg A. (1979). Retrospective DNA analyses in cervical dysplasia as related to neoplastic progression or regression. Anal. Quant. Cytol. 11, 103-106.

Naslund I., Rubio C.A. and Auer G.U. (1987). Nuclear DNA changes during pathogenesis of squamous carcinoma of the cervix in 3,4-benzopyrene-treated mice. Anal. Quant. Cytol. Histol. 9, 411-418.

Newcombe R.G. (1987). Towards a reduction in publication bias. Br. Med. J. 295, 656-659.

Ng A.B.P. and Atkin N.B. (1973). Histological cell type and DNA value in the prognosis of squamous cell cancer of uterine cervix. Br. J. Cancer 28, 322-331.

Oxman A.D. (1994). Checklists for review articles. Br. Med. J 309, 648-651.

Oxman A.D. and Guyatt G.H. (1992). A consumer's guide to subgroup analyses. Ann. Intern. Med. 116, 78-84.

Oxman A.D. and Guyatt G.H. (1993). The science of revieweing research. Ann. N. Y. Acad. Sci. 703, 125-134.

Pascale Segers P., Haesen S., Castelain P., Amy J.J., De Sutter P., Van Dam P. and Kirsch-Volders M. (1995). Study of numerical aberrations of chromosome 1 by fluorescent in situ hybridization and DNA content by densitometric analysis on (pre)-malignant cervical lesions. Histochem J. 27, 24-34.

Peipert J.F. and Bracken M.B. (1997). Systematic reviews of medical evidence: the use of meta-analysis in obstetrics and gynecology. Obstet. Gynecol. 89, 628-633.

Piantadosi S. and Byar D.P. (1988). A proposed for registering clinical trials. Control. Clinic. Trials 9, 82-84.

Randal M.E., Constable W.C., Hahn S.S., Kim J-A. and Mills S.E. (1988). Results of the radiotherapeutic management of carcinoma of the cervix with emphasis on the influence of histologic classification. Cancer 62, 48-53.

Rihet S., Lorenzato M. and Clavel C. (1996). Oncogenic human papillomaviruses and ploidy in cervical lesions. J. Clin. Pathol. 49, 892-896.

Robboy S.J., Noller K.L., O'Brien P., Kaufman R.H., Townsent D., Barnes A.B., Gunderson J., Lawrence D., Bergstrahl E., McGorray S., Tilley B.C., Anton J. and Chazen G. (1984). Increased incidence of cervical and vaginal dysplasia in 3980 diethylstilbestrol-exposed young women. JAMA 252, 2979-2983.

Sacks H.S., Berrier J., Reitman D., Ancona-Berk V.A. and Chalmers T.C. (1987). Meta-analyses of randomized controlled trials. N. Engl. J. Med. 316, 450-455.

Saigo P.E., Cain J.M., Kim W.S., Gaynor J.J., Johnson K. and Lewis J.L. (1986). Prognostic factors in adenocarcinoma of the uterine cervix. Cancer 57, 1584-1593.

Sandritter W., Carl M. and Ritter W. (1966). Cytophotometric measurements of the DNA content of human malignant tumors by means of the Feulgen reaction. Acta Cytol. 10, 26-30.

Scherer R.W., Dickersin K. and Langenberg P. (1994). Full publications of results initially presented in abstracts. A meta-analysis. JAMA 272, 158-162.

Schulz K.F., Chalmers I., Grimes D.A. and Altman D.G. (1994). Assessing the quality of randomization from reports of controlled trials published in obstetrics and gynecology journals. JAMA 272, 125-128.

Schulz K.F., Chalmers I., Hayes R.J. and Altman D.G. (1995). Empirical evidence of bias. Dimensions of methodological quality associated with estimates of treatment effects in controlled trials. JAMA 273, 408-412.

Shapiro S. (1994). Meta-analysis/Shmeta analysis. Am. J. Epidemiol. 9, 771-778.

Shevchuk M.M. and Richart R.M. (1982). DNA content of condyloma acuminatum. Cancer 49, 489-492.

Shorrock K., Johnson I. and Johnson I.R. (1990). Epidemiological changes in cervical carcinoma with particular reference to mucin-secreting subtypes. Histopathology 13, 53-57.

Simes R.J. (1986). Publication bias: the case for an international registry of clinical trials. J. Clin. Oncol. 4, 1529-1541.

Spitzer W.O. (1991). Meta-meta analysis: unaswered questions about aggregating data. J. Clin. Epidemiol. 44, 103-107.

Steinbeck R.G., Heselmeyer K.M., Moberger H.B. and Auger G.U. (1995). The relationship between proliferating cell nuclear antigen (PCNA), nuclear DNA content and mutant p53 during genesis of cervical carcinoma. Acta Oncol. 34, 171-176.

Terada K.Y., Morley G.W. and Roberts J.A. (1988). Stage 1B carcinoma of the cervix with lymph node metastases. Gynecol. Oncol. 31, 389-395.

Thacker S.B. (1988). Meta-analysis. A quantitative approach to research integration. JAMA 259, 1685-1689.

Thacker S.B., Peterson H.B. and Stroup D.F. (1996). Meta-analysis for the obstetrician-gynecologist. Am. J. Obstet. Gynecol. 174, 1403-1407.

Tsou K-C., Hong D-H., Varello M., Giuntoli R., Wheeler J.E., Atkinson B.F., Mangan C. and Kikuta J. (1984). Flow cytometric DNA analysis as a diagnostic acid for cervical condolyloma and cancer. Cancer 54, 1778-1787.

Underwood J.C.E. (1997). Editorial. Histopathology 31, 1-2.

van Driel-Kulker .M.J. and Ploem-Zaaijer J.J. (1989). Image cytometry in automated cervical screening. Anal. Cell Pathol. 1, 63-77.

van Leeuwen A.M., Ploem-Zaaijer J.J., Pieters W.J.L.M., Hollema H. and Burger M.P.M. (1996). The suitability of DNA cytometry for the prediction of the histological diagnosis in women with abnormal cervical smears. Br. J. Obstet. Gynaecol. 103, 359-365.

Watts K.C., Husain O.A.N., Campion M.J., Lorriman F., Butler E.B., McCance D., Jenkisn D. and Singer A. (1987). Quantitative DNA analysis of low grade cervical intraepithelial neoplasia and human papillomavirus infection by static and flow cytometry. B. Med. J. 295, 1090-1092.

Wheeless C.R., Graham R. and Graham J.B. (1970). Prognosis and treatment of adenoepidermoid carcinoma of the cervix. Obstet. Gynecol. 35, 928-932.

Wilkbanks G.D., Richart R.M. and Terner J.Y. (1967). DNA content of cervical intraepithelial neoplasia studied by two-wavelength Feulgen cytophotometry. Am. J. Obstet. Gynecol. 98, 792-799.

Winkler B., Crum C.P., Fujii T., Ferenczy A., Boon M., Braun L., Lancaster W.D. and Richart R.M. (1984). Koilocytotic lesions of the cervix. The relationship of mitotic abnormalities to the presence of papillomarivus antigens and nuclear DNA content. Cancer 53, 1081-1087.

Yaziqi R., Sandstad J., Munoz A.K., Choi D.J., Nguyen P.D. and Risser R. (1990). Adenosquamous carcinoma of the cervix: prognosis in stage 1b. Obstet. Gyneol. 75, 1012-1015.

Yusuf S., Wittes J., Probstfield J. and Tyroler H.A. (1991). Analysis and interpretation of treatment effects in subgroups of patients in randomized clinical trials. JAMA 266, 93-98.

Zolzer F., Alberti W., Pelzer T., Lamberti G., Hulskamp F.H. and Streffer C. (1995). Changes in S-phase fraction and micronucleus frequency as prognostic factors in radiotherapy of cervical carcinoma. Radiother. Oncol. 36, 128-132.

Accepted June 5, 1998