Suggested measures to ensure drug safety before definite licensing of a drug

- Legal requirement for drug companies to register all randomised controlled trials prospectively
- Legal requirement for drug companies to make all data on serious adverse events from clinical studies publicly available immediately after study completion
- Continuously updated systematic reviews of adverse events based on published and unpublished data from randomised controlled trials and observational studies
- Phased introduction of new interventions in independent, large scale, randomised trials before definite drug licensing
- Clear cut financial firewalls between pharmaceutical companies and researchers performing systematic reviews and clinical studies

has to be reconsidered. Thirdly, we must find ways of preventing further similar episodes (box).

Single phase III drug trials are simply not big enough to detect relatively uncommon but important adverse events, which may affect large numbers of people in routine clinical use.[6] The potential public health impact of previously undetected drug related adverse events is likely to be made worse if widely marketed new drugs are prescribed haphazardly and rapidly to large numbers of people. Within five months of the launch of rofecoxib, more than 42 000 patients had been prescribed the drug in England,[11] even though newly marketed drugs carry a black triangle warning, indicating an incomplete safety profile. Unfortunately, postmarketing surveillance is not a panacea to determine safety, as methodological flaws may produce inaccurate results.

We therefore recommend that drug companies are legally required to make all data on serious adverse events from clinical studies available to the public immediately after completion of the research. This will allow independent, timely, and updated systematic reviews of serious adverse events. In addition, we advocate the phased introduction of new interventions through randomised trials, which are independent from the pharmaceutical industry and are large enough to study rare outcomes, together with systematic, more robust, and comprehensive approaches to pharmacovigilance.[12] Although these measures will not be popular with pharmaceutical companies, they will limit the numbers of patients exposed unsystematically to unknown hazards and provide robust and unbiased evidence on adverse events before a drug is licensed fully.

Paul A Dieppe *director*

MRC Health Services Research Collaboration, University of Bristol, Bristol BS8 2PR
(p.dieppe@bristol.ac.uk)

Shah Ebrahim *head of department*
Richard M Martin *senior lecturer in epidemology and public health*

Department of Social Medicine, University of Bristol

Peter Jüni *senior research fellow in clinical epidemiology*

Departments of Social and Preventive Medicine and Rheumatology, University of Berne, CH-3012 Berne, Switzerland

Competing interests: SE is coordinating editor of the Cochrane Heart Group, which produces systematic reviews not funded by the pharmaceutical industry of the effects of interventions for heart diseases. RMM worked at the Drug Safety Research Unit between 1997 and 1999, during which time the unit received unconditional donations from pharmaceutical companies.

1 *Merck announces voluntary worldwide withdrawal of VIOXX®*. News release. Whitehouse Station, NJ: Merck, 2004. www.vioxx.com/vioxx/documents/english/vioxx_press_release.pdf (accessed 30 Sep 2004).
2 Jüni P, Dieppe P. Older people should NOT be prescribed "coxibs" in place of conventional NSAIDs. *Age Ageing* 2004;33:100-4.
3 Smith R. Medical journals and pharmaceutical companies: uneasy bedfellows. *BMJ* 2003;326:1202-5.
4 Taggart HM, Alderdice JM. Fatal cholestatic jaundice in elderly patients taking benoxaprofen. *BMJ (Clin Res Ed)* 1982;284:1372.
5 Kargman S, Charleson S, Cartwright M, Frank J, Riendeau D, Mancini J, et al. Characterization of Prostaglandin G/H Synthase 1 and 2 in rat, dog, monkey, and human gastrointestinal tracts. *Gastroenterology* 1996;111:445-54.
6 Dieppe P, Bartlett C, Davey P, Doyal L, Ebrahim S. Balancing benefits and harms: the example of non-steroidal anti-inflammatory drugs. *BMJ* 2004;329:31-4.
7 Silverstein FE, Faich G, Goldstein JL, Simon LS, Pincus T, Whelton A, et al. Gastrointestinal toxicity with celecoxib vs nonsteroidal anti-inflammatory drugs for osteoarthritis and rheumatoid arthritis. The CLASS study: a randomized controlled trial. *JAMA* 2000;284:1247-55.
8 Jüni P, Rutjes AW, Dieppe PA. Are selective COX 2 inhibitors superior to traditional non steroidal anti-inflammatory drugs? *BMJ* 2002;324:1287-8.
9 Bombardier C, Laine L, Reicin A, Shapiro D, Burgos-Vargas R, Davis B, et al. Comparison of upper gastrointestinal toxicity of rofecoxib and naproxen in patients with rheumatoid arthritis. VIGOR Study Group. *N Engl J Med* 2000;343:1520-8.
10 Mukherjee D, Nissen SE, Topol EJ. Risk of cardiovascular events associated with selective COX-2 inhibitors. *JAMA* 2001;286:954-9.
11 Layton D, Riley J, Wilton LV, Shakir SA. Safety profile of rofecoxib as used in general practice in England: results of a prescription-event monitoring study. *Br J Clin Pharmacol* 2003;55:166-74.
12 Waller PC, Evans SJ. A model for the future conduct of pharmacovigilance. *Pharmacoepidemiol Drug Saf* 2003;12:17-29.

The scandal of poor epidemiological research

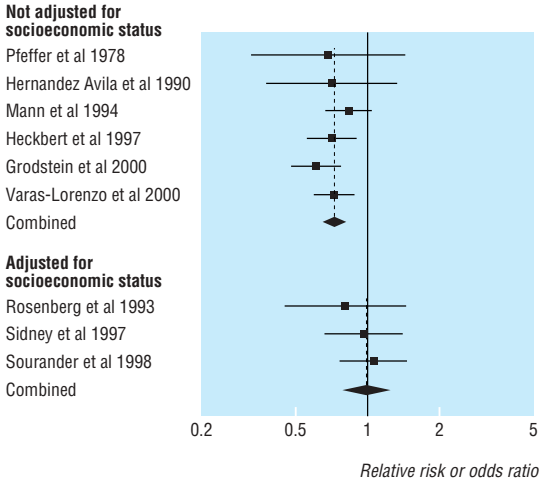
Reporting guidelines are needed for observational epidemiology

Something surely must be wrong with epidemiology when the new editors of a leading journal in the field entitle their inaugural offering, "Epidemiology—is it time to call it a day?"[1] Observational epidemiology has not had a good press in recent years. Conflicting results from epidemiological studies of the risks of daily life, such as coffee, hair dye, or hormones, are frequently and eagerly reported in the popular press, providing a constant source of anxiety for the public.[2 3] In many cases deeply held beliefs, given credibility by numerous observational studies over long periods of time, are challenged only when contradicted by randomised trials. In the most recent example, a Cochrane review of randomised trials shows that antioxidant vitamins do not prevent gastrointestinal cancer and may even increase all cause mortality.[4 5]

Now Pocock et al describe the quality and the litany of problems of 73 epidemiological studies published in January 2001 in general medical and specialist journals (p 883).[6] Perhaps the most relevant findings relate to how investigators dealt with confounding, multiple comparisons, and subgroup analyses.

Papers p 883

BMJ 2004;329:868–9



Meta-analysis of cohort studies and case-control studies of hormone replacement therapy and coronary heart disease. There is little evidence for a protective effect when analyses are adjusted for, in contrast to studies not adjusted for, socioeconomic status. Adapted from Humphrey et al, reference 7

Confounding, the situation in which an apparent effect of an exposure on risk is explained by its association with other factors, is probably the most important cause of spurious associations in observational epidemiology. For example, a recent meta-analysis of observational studies shows how confounding could have been responsible for the belief that hormone replacement therapy provides protection against cardiovascular disease.[7] A protective effect of hormone replacement therapy was evident in studies that did not control for socioeconomic status, but not in studies that did (figure).[7] Higher socioeconomic position is strongly associated with both more frequent use of hormone replacement therapy and lower risk of coronary heart disease. In the large (unconfounded) Women's Health Initiative randomised trial hormone replacement therapy had no beneficial effect on cardiovascular disease.[8]

Worryingly, Pocock et al find that the rationale behind the choice of confounders is usually unclear, and that the extent of adjustment varies greatly. They also confirm that observational studies often consider several exposures, outcomes, and subgroups. This results in multiple statistical tests of hypotheses and a high probability of finding associations that are statistically significant but spurious. In many studies 20% or more of the findings may be erroneous, rather than the expected 5% false positive associations ($P < 0.05$).[9]

Epidemiology is not alone: and there is hope. As Altman pointed out 10 years ago, much medical research is of poor quality.[10] Efforts so far to remedy what Altman described as the "scandal of poor medical research" have focused on controlled clinical trials. Empirical research has shed light on aspects of methodological quality that are crucial to prevent bias,[11] and Consolidated Standards for Reporting Trials (CONSORT) have been developed and endorsed widely.[12]

Clearly, such guidelines are also required for observational epidemiology. A month ago a group of epidemiologists, statisticians, and editors met in Bristol to work on a first draft of STROBE, the Standards for the Reporting of Observational Studies in Epidemiology. STROBE will provide guidance on the reporting of cohort studies, case-control studies, and cross sectional studies; a supporting document will give explanations and examples. Outcomes from the Bristol workshop will be posted on a dedicated website (www.strobe-statement.org), and you and everyone else will be invited to comment and suggest improvements before a revised version is published some time next year.

Can STROBE do for observational epidemiology what CONSORT does for randomised clinical trials? Not exactly. Both guidelines aim to promote comprehensive reporting of important methodological detail and facilitate appraisal of study quality. In the case of a large high quality randomised trial, this means that the results can be assumed to provide an unbiased estimate of the treatment effect in the population studied. This is not so in observational epidemiology. A well conducted case-control or cohort study might still produce misleading results if, for example, important confounders were not known, not measured, or imprecisely measured.

Importantly for observational studies, STROBE will also pay considerable attention to what investigators should write in the discussion section of their paper by suggesting the inclusion of statements on why methodological approaches were chosen, and why results are interpreted the way they are. The skilful interpretation of epidemiological evidence, bearing in mind theoretical considerations, resisting being seduced by possible mechanisms, and suspending beliefs to allow healthy scepticism will remain the great challenge and joy of working in epidemiology.[8] More transparent and complete reporting of epidemiological studies, coupled with more thoughtful interpretation of their results, will help restore the reputation of a discipline that has contributed importantly to improving the health of the public, and will continue to do so in the future.

Erik von Elm *research fellow*

(vonelm@ispm.unibe.ch)

Matthias Egger *professor*

(egger@ispm.unibe.ch)

Department of Social and Preventive Medicine, University of Berne, Finkenhubelweg 11, CH-3012 Berne, Switzerland

Competing interests: EvE and ME are members of the STROBE group.

1 Davey Smith G, Ebrahim S. Epidemiology—is it time to call it a day? *Int J Epidemiol* 2001;30:1-11.
2 Taubes G. Epidemiology faces its limits. *Science* 1995;269:164-9.
3 Bartlett C, Sterne J, Egger M. What is newsworthy? Longitudinal study of the reporting of medical research in two British newspapers. *BMJ* 2002;325:81-4.
4 Bjelakovic G, Nikolova D, Simonetti RG, Gluud C. Antioxidant supplements for prevention of gastrointestinal cancers: a systematic review and meta-analysis. *Lancet* 2004;364:1219-28.
5 Lawlor DA, Davey SG, Kundu D, Bruckdorfer KR, Ebrahim S. Those confounded vitamins: what can we learn from the differences between observational versus randomised trial evidence? *Lancet* 2004;363:1724-7.
6 Pocock SJ, Collier TJ, Dandreo KJ, de Stavola BL, Goldman MB, Kalish LA, Kasten LE, McCormack VA. Issues in the reporting of epidemiological studies: a survey of recent practice. *BMJ* 2004;329:883-7.
7 Humphrey LL, Chan BK, Sox HC. Postmenopausal hormone replacement therapy and the primary prevention of cardiovascular disease. *Ann Intern Med* 2002;137:273-84.
8 Petitti D. Commentary: hormone replacement therapy and coronary heart disease: four lessons. *Int J Epidemiol* 2004;33:461-3.
9 Ottenbacher KJ. Quantitative evaluation of multiplicity in epidemiology and public health research. *Am J Epidemiol* 1998;147:615-9.
10 Altman DG. The scandal of poor medical research. *BMJ* 1994;308:283-4.
11 Jüni P, Altman DG, Egger M. Assessing the quality of controlled clinical trials. *BMJ* 2001;323:42-6.
12 Moher D, Schulz KF, Altman DG, for the CONSORT Group. The CONSORT statement: revised recommendations for improving the quality of reports of parallel group randomized trials. *JAMA* 2001;285:1987-91.