Commentary

An agenda for UK clinical pharmacology Pharmacoepidemiology

Stephen J. W. Evans

Department of Medical Statistics, The London School of Hygiene & Tropical Medicine, Keppel Street, London WC1E 7HT, UK

Current concerns over the safety of medicines once they have been marketed mean that pharmacoepidemiology is of increasing importance. There are three main areas in which further research is needed.

- 1 To improve the methods used to make causal inference of effects of medicines and to raise the quality of the reporting and critical appraisal tools, so that the strengths and weaknesses of the new methods can be judged.
- 2 To apply the methods in areas where randomized trials cannot easily be done, such as in pregnancy.
- 3 To use electronic health records as fully as possible, using linkage between different databases, ensuring the data are of as high quality as possible.

Public health and public perceptions mean that much of pharmacoepidemiology must be done using non-industry funding sources.

Introduction

Pharmacoepidemiology is rarely taught to medical students and the amount of time devoted to it in clinical training is limited. However, the past few years have seen an emphasis on the safety of medicines and evidence-based prescribing that seems to be greater than at any time since the immediate aftermath of the thalidomide disaster 50 years ago.

Pharmacoepidemiology as a discipline is probably less than 50 years old. The word was first used only just over 25 years ago [1], and before that, 'pharmaceutical epidemiology' was a phrase used by Jan Venulet, with the same sense [2]. It can be defined simply as the application of epidemiological methods to the effects of medicines, including vaccines and cell-based or biological treatments. As such it has strong links with, even if it is not part of, clinical pharmacology. Epidemiological methods generally use studies of large or extremely large numbers of individuals and so pharmacoepidemiology seems removed from the usual concerns of clinical pharmacology. However, the clinical aspects of the effects of medicines are often only well estimated in large numbers of people and sufficiently large randomized trials using clinical end points are not as frequent as those (smaller trials) that use surrogates. For example, it requires smaller numbers to show statistically significant reductions in cholesterol concentrations than reductions in the numbers of myocardial infarctions or in

overall mortality. As a general point, pharmacoepidemiology shows some signs of losing its links to pharmacology and it is vital that it does not do so. Concentration on methods is good, but not to the exclusion of the molecular basis for the actions of medicines.

While it is clear that randomized trials generally have the greatest inferential power to determine causal effects, they are expensive and might be seen as unethical when there is some convincing evidence of benefit. Clinically relevant harms may be rarer than clinically relevant benefits, but they can be serious enough to alter the benefit-harm balance of a marketed medicine. In some instances individual cases of harms can be used to show a truly causal effect [3], but individual cases cannot easily be used to estimate frequency of adverse reactions and hence will rarely affect the benefit-harm balance for a marketed medicine. They can assist the patient and physician in knowing that a drug causes a reaction, so that stopping it or using some other intervention can reduce or prevent the harm. Pharmacoepidemiology comes into its own when there is limited information regarding harms and their frequencies. Increasingly, it is concerned with measuring the comparative effectiveness of two treatments, each with acknowledged benefit. This will be a challenge.

So there is clearly a place for pharmacoepidemiology, especially in confirming clinical benefits in real practice and looking for harms, in the hope of failing to find them or at least showing that their rate is low, so that the goal of

safety (absence of important harms) can be assured. It can be conducted in 'real world' situations, in contrast to the often rarefied atmosphere of randomized controlled trials (RCTs), especially those used to obtain marketing authorizations for medicines. The usual RCTs have limitations, because they have too narrow entry criteria in terms of age (children and the elderly often being excluded), gender (women of child-bearing age or pregnant women being excluded), co-morbidities (multiple co-morbidities and patients with renal or liver disease being excluded) and co-prescriptions. Furthermore, the power of a study is calculated for efficacy and is too low for studying harms.

This has been set out clearly by Rawlins in his Harveian Oration [4], although his advocacy of observational research has been misinterpreted by some as suggesting that randomized trials are unnecessary. Some (extreme?) Bayesians have suggested that RCTs are illogical [5], but they seem to misunderstand the problems inherent to observational data, and such a view is not shared by this author.

Here I shall concentrate on issues that relate to medicinal products, but the methods and requirements for advances apply equally in other areas, such as vaccines and cell-based therapies.

Areas in which advances are required

The overall areas under consideration can be divided into three groups:

- 1 Problems of inferring causality when methods are the focus.
- 2 Application areas and related clinical questions.
- 3 The sources and quality of the available data.

Priorities for research – epidemiological methods and causality

Observational studies are often criticized because there is uncertainty over the strength of causal inferences that can be drawn, and a mantra acknowledging such limitations often appears in the Discussion sections of papers. Epidemiologists have used the classical cohort and case-control study designs to study drug effects for many years and they continue to have their place. Many, if not most, of them in the 21st century are conducted in databases, but the methods all involve observing what has occurred under ordinary conditions of use. This means that forming a reasonable comparison group is not easy, and there will usually be differences between those treated and those used as a comparison group. This is becoming more difficult as guidelines for treatment are implemented. When the differences are also associated with the outcome of interest, this results in confounding. Classical methods of adjustment for such differences will never perfectly adjust,

and attempts to improve things are a vital component when there are priorities for research in methods for pharmacoepidemiology. Other biases can also affect interpretation, even of clinical trials, but confounding by indication is a major problem in this area.

One of the important new methods to address the problems of causality is the use of propensity scores [6]. These model the probability (called the propensity score, PS) that individuals have of being in the treatment group (as opposed to the comparison group). Then a more valid comparison can be made, using comparisons between those with equal values of the PS, i.e. stratifying by PS. Alternatively, a weighted analysis using the inverse of the PS or regression adjustment using the PS can be done. The propensity score depends on characteristics of the individuals potentially or actually receiving the treatment, but not on their association with the outcome(s) that may be studied. These methods can be used in various ways, but at best they yield only limited gains compared with classical methods. They are useful when dealing with rare outcomes and the approach has been used a great deal in pharmacoepidemiology [7].

One use of PSs has been to check if the results of randomized trials that look at a relatively frequent outcome can be reproduced using the method, so that the same method can be used to look at less frequent outcomes. In contrast, regression methods require a different model for each outcome to be studied. This means that the results may be more trustworthy than a regression approach for the rarer outcomes [8].

Priority 1: Further research is necessary to elucidate whether the results, particularly of studies using so-called high-dimensional propensity scores, give reliable answers.

A second major approach to improving causal inference has been the use of marginal structural models [9]. These have been used in pharmacoepidemiology, but we do not yet know whether they have general applicability. A key aspect is understanding which variables should be used to adjust for confounding. Variables on the causal pathway that are incorrectly used in a regression model can lead to biased results. It is certainly possible that some of the discrepant results found in observational research may be due to such incorrect adjustments. Instrumental variables and 'Directed Acyclic Graphs' have been used to try to improve causal inference, but they have not yet been shown to be as helpful as was originally hoped.

Pure methodologists rather than pharmacoepidemiologists (although there are some who are both) are likely to carry out the advances in these types of methods in general. The main priority is in making appropriate use of the methods in pharmacoepidemiology and ensuring that the application is done correctly. Critical appraisal guidelines are well known for cohort and case-control studies and the STROBE initiative [10] encourages clear reporting of the methods of observational studies. However, the average reader of, say, this journal, who can form a reason-

able view of the analysis of a cohort study, has limited tools for critical appraisal or dissection of a study using these more complex methods. A useful survey of applications of propensity scores was done by Weitzen *et al.* [11], but further guidance on critical appraisal is required.

Priority 2: A research priority is to make better evidence-based guidelines for potential users of complex methods, including propensity scores and marginal structural models, and better tools for readers to understand them.

A major new epidemiological study design with an associated new method for analysis was invented in 1995 [12]. This method, the 'Self-Controlled Case Series' (SCSS), has had major use in assessing the safety of vaccines, for which it was originally designed [13]. More recently it has been used to study drug effects [14, 15], and while it has some similarities to the case-crossover design [16] it has several advantages and new additions to the method [17, 18] have increased its utility. Its major advantage, implied in the name, is that it avoids fixed confounders, because there is no comparison group. It uses as their own controls those who have been exposed and have had the outcome. It estimates the incidence of the outcome in different periods around the time of exposure, defining time periods as being 'risk' or 'control', before and usually also after the exposure periods. The incidence rate ratio then compares risk and control rates. While it is not applicable to every pharmacoepidemiology question of interest, it does have wider applicability than was initially thought. Clear tutorials on its use have been published, but there is a need to have more comprehensive guidelines on when it can and cannot be used in pharmacoepidemiology. A very useful guide with regard to vaccines has been published [19], but further work on guidelines for medicines research would be useful.

Priority 3: Work needs to be done in setting standards for reporting both SCSS and the other self-controlled designs (they are not yet mentioned in the STROBE guidelines, nor are there any clear evidence-based guidelines for critical appraisal in the way that they have been derived for RCTs and other designs).

Miettinen [20] and Vandenbroucke [21] have argued that observational studies are less biased when they are looking at *unintended* effects. A useful recent paper has shown that in looking at harms (in contrast to benefits), observational data and RCTs have more similar results than has often been thought [22].

Whether the use of observational studies can be extended to provide reliable results in comparative effectiveness is a very hot topic. In the USA a major effort is being put into comparative effectiveness and Congress has set up an independent Patient-Centred Outcomes Research Institute [23], which has, as a major component, a Methodology Committee that includes pharmacoepidemiology expertise as well as specialists in RCTs. The Health technology appraisal process does not seem to have progressed as much in this direction in Europe in general. There has

perhaps been more of a separation between RCT and pharmacoepidemiology expertise in setting up research.

Priority 4: To extend the work on unintended effects and to identify whether there are characteristics of observational studies that are associated with reliable results for benefits as well as harms. This needs to include designs and methods of analysis.

Priorities for research – which applications and which clinical questions?

As noted above, comparative effectiveness is a major area, but the methodological aspects will need to be addressed before confidence in pharmacoepidemiology in this context can be said to provide definitive answers.

An aspect of pharmacoepidemiology that is not usual in general epidemiology is in describing the epidemiology of morbidities associated with particular disease or age groups. These are not risk factors that are of interest in finding causes of disease, a major interest of general epidemiology, but in obtaining background rates of adverse events that might be suspected as being associated with a drug or vaccine, while in fact they are coincidental events. An example was that the introduction of bupropion as an anti-smoking intervention was followed by large numbers of reports of seizures, although earlier studies, when it was first introduced as an antidepressant, had also shown such an association [24]. It was clear at first that there was no knowledge of how frequent seizures were in smokers in general. This type of knowledge is not very exciting, but it is of great importance in setting spontaneous reports of suspected adverse reactions in context. Major efforts were made in this type of approach, both before the introduction of the human papilloma virus (HPV) vaccine [25], but also for specific events such as Guillain-Barré syndrome before the use of the pandemic influenza vaccine [26]. The results of such systematic studies, particularly if they are done across countries, cannot be obtained extremely rapidly, so careful planning is required.

Priority 5: A research priority is to describe adverse reactions in potentially treated populations before the introduction of a new medication (drug or biological product), either in a novel indication or in a new age group.

The SCSS method, as noted above, has been extended to allow for events that curtail follow-up after the event, and the range of questions that might be addressed using the method has increased. It is of interest that it can be used with data derived from conventional cohort or case-control studies. When this has been done, it has been shown that the conventional methods may be affected by unmeasured confounding. An interesting example has been a comparison of the risks of fractures in patients taking tricyclic or SSRI antidepressants [27]. This has shown that, in contrast to conventional analyses, the increased risks are similar in the two classes.

Descriptive pharmacoepidemiology has other uses in studying drug utilization. The priorities are in detecting

possible inappropriate prescribing or in demonstrating that changes in recommendations have been carried out in the community. Studies of the use of marketed medicines in unlicensed indications can be very helpful. Perhaps the major priority is in studying medicines used in pregnancy [28]. There is obviously a great deal of fear among both patients and prescribers about such uses and it is sensible to avoid unnecessary use. However, knowing what the true benefits and harms are in this population is vital.

Priority 6: Pharmacoepidemiology will probably be the only way of studying drugs in pregnancy and is a high priority for further research, not just to identify harms but to examine the overall benefit–harm balance for treatments in pregnancy.

In the more general population the benefits of diagnostic test monitoring are not known with confidence, and it occupies a great deal of resources in practice [29], but further research in this area is also required.

The limitations of spontaneous reporting of suspected adverse reactions are well known. While ADRs with rapid onset and those with a low background incidence can be detected relatively easily, those with delayed onset or that have a high background incidence, like myocardial infarction, are more difficult to detect. With the advent of large routine databases, the potential for scanning them to detect ADRs has been suggested [30] and results are beginning to be seen. Initiatives in Europe and the USA have begun recently and this continues to be a high priority for research. It is not yet clear which pharmacoepidemiological methods of design and analysis will be most useful. Clearly, many false-positive associations may be found when looking at large numbers of combinations of drugs and adverse events. There are many problems to solve, and the key process of estimating false-positive and false-negative rates in real data is made very difficult by the absence of a gold standard. Simulated data, for which the 'truth' is known, may also not reflect the patterns of data seen in practice.

This is a very high priority and the UK has not been at the forefront of this, although some groups have been involved in the European initiatives. The possibility of incorporating knowledge of basic pharmacology in signal detection is a field of current interest, and this should be a priority, with the need for collaboration between several disciplines. It has been used to examine potential risks in drug discovery, but not yet in signal detection or pharmacoepidemiology [31].

A general area of research is in studying the longer term effects of older drugs. It is often assumed that they are safer, but there is now evidence that, for example, diclofenac may confer as much of an increased risk of myocardial infarction as COX-2 inhibitors. The large randomized trials that brought into question the cardiovascular safety of the COX-2 inhibitors have not been carried out for older NSAIDs, and it is pharmacoepidemiological methods that will provide the answers, since it is unlikely that RCTs

will be done to answer such questions. The adage that 'absence of evidence is not evidence of absence' applies particularly to knowledge of the long term safety of many drugs.

Priority 7: The use of good methods to improve signal detection using electronic health records rather than just spontaneous reporting. This should utilize knowledge of pharmacology and look at longer term reactions as well as immediate ones.

Priorities for research – sources and quality of the available data

With increasing use of databases there is a danger that errors in the recorded data, missing data for expected variables and the complete absence of some variables may not be fully recognized. Data that are recorded for routine care may be interpretable, even in the presence of errors, but may make research difficult or may produce misleading results. There is a need for more research on the impact of missing data in pharmacoepidemiology. A Bayesian approach has been suggested for estimating the effect of unmeasured confounding [32], but it has not yet been shown to have wide validity.

While databases provide relatively quick answers, studies that are set up to collect data on a specific question still have a role. The prospective collection of data allows all the relevant variables to be collected and has the potential to provide reliable answers. A group led by Abenhaim has shown that this can provide relatively rapid answers, with a useful paper on Guillain–Barré syndrome and its relation to influenza virus and immunization [33]. It is important to ensure that pharmacoepidemiology does not simply become analysis of databases. Their limitations must be acknowledged and the need to allow for a variety of data sources to answer questions is important.

Diagnoses and drug prescriptions have been validated in many databases, most notably in the General Practice Research Database [34], and it is important that such validation continues. It cannot be assumed that quality of data remains high over many years. It is not exciting, but as a basis for good research it is a priority that checks continue to be made.

Priority 8: The possibility of linkage with other data sources, such as cancer registries, allows for validation of data and such developments are to be welcomed and will bring new challenges.

Conclusions

Pharmacoepidemiology is a fascinating field, with major intellectual challenges. The research priorities that are listed above do not require major resources, but the general issue of funding is itself a priority. There is a danger that whether or not industry funding leads to bias, the perception among the public and some journal editors is

that it is likely to lead to results favourable to industry. There are financial pressures and pressures from patients to bring new drugs to market early, but there has not been the same pressure in the past to have effective post-marketing monitoring. New legislation in Europe (and in the USA) is changing this. The European Union has made funding available in recent years for pharmacoepidemiology, but the UK record has not been as good. It is always easier, it seems, to obtain funding for molecular biology, so much research has had to rely on industry. Drug safety leads to major public concerns, but these concerns have not been followed by notable public funding. There is a need for funding of the methods and infrastructure that allow groups to be able to carry out studies relatively rapidly. The pressure to have almost immediate answers is great, but this must not lead to a reduction in quality.

Competing Interests

There are no competing interests to declare.

I am grateful to Professor Liam Smeeth and Dr Sara Thomas, who made very helpful comments on an early draft of this paper. They are not responsible for errors that remain.

REFERENCES

- 1 Lawson DH. Pharmacoepidemiology: a new discipline. *Br Med J (Clin Res Ed)* 1984; 289: 940–1.
- 2 Venulet J. From experimental to social pharmacology. Natural history of pharmacology. *Int J Clin Pharmacol* 1974; 10: 203–5.
- 3 Hauben M, Aronson JK. Gold standards in pharmacovigilance: the use of definitive anecdotal reports of adverse drug reactions as pure gold and high-grade ore. *Drug Saf* 2007; 30: 645–55.
- 4 Rawlins M. De testimonio. On the evidence for decisions about the use of therapeutic interventions. *Lancet* 2008; 372: 2152–61.
- 5 Urbach P. Randomization and the design of experiments. *Philos Sci* 1985; 52: 256–73.
- 6 Rubin DB. Estimating causal effects from large data sets using propensity scores. *Ann Intern Med* 1997; 127 (8 Pt 2): 757–63.
- 7 Rassen JA, Avorn J, Schneeweiss S. Multivariate-adjusted pharmacoepidemiologic analyses of confidential information pooled from multiple health care utilization databases. *Pharmacoepidemiol Drug Saf* 2010; 19: 848–57.
- 8 Smeeth L, Douglas I, Hall AJ, Hubbard R, Evans S. Effect of statins on a wide range of health outcomes: a cohort study validated by comparison with randomized trials. *Br J Clin Pharmacol* 2009; 67: 99–109.
- 9 Robins JM, Hernán MA, Brumback B. Marginal structural models and causal inference in epidemiology. *Epidemiology* 2000; 11: 550–60.
- 10 von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP. STROBE Initiative. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *Epidemiology* 2007; 18: 800–4.
- 11 Weitzen S, Lapane KL, Toledano AY, Hume AL, Mor V. Principles for modeling propensity scores in medical research: a systematic literature review. *Pharmacoepidemiol Drug Saf* 2004; 13: 841–53.
- 12 Farrington CP. Relative incidence estimation from case series for vaccine safety evaluation. *Biometrics* 1995; 51: 228–35.
- 13 Farrington CP. Control without separate controls: evaluation of vaccine safety using case-only methods. *Vaccine* 2004; 22: 2064–70.
- 14 Smeeth L, Thomas SL, Hall AJ, Hubbard R, Farrington P, Vallance P. Risk of myocardial infarction and stroke after acute infection or vaccination. *N Engl J Med* 2004; 351: 2611–8.
- 15 Douglas IJ, Evans SJ, Pocock S, Smeeth L. The risk of fractures associated with thiazolidinediones: a self-controlled case-series study. *PLoS Med* 2009; 6: e1000154.
- 16 Maclure M. The case-crossover design: a method for studying transient effects on the risk of acute events. *Am J Epidemiol* 1991; 133: 144–53.
- 17 Farrington CP, Whitaker HJ, Hocine MN. Case series analysis for censored, perturbed or curtailed post-event exposures. *Biostatistics* 2009; 10: 3–16.
- 18 Farrington CP, Anaya-Izquierdo K, Whitaker HJ, Hocine MN, Douglas I, Smeeth L. Self-controlled case series analysis with event-dependent observation periods. *J Am Stat Assoc* 2011; 106: 417–26.
- 19 Weldelessie YG, Whitaker HJ, Farrington CP. Use of the self-controlled case-series method in vaccine safety studies: review and recommendations for best practice. *Epidemiol Infect* 2011; 18: 1–13.
- 20 Miettinen OS. The need for randomization in the study of intended effects. *Stat Med* 1983; 2: 267–71.
- 21 Vandenbroucke JP. Observational research, randomised trials, and two views of medical science. *PLoS Med* 2008; 5: e67.
- 22 Golder S, Loke YK, Bland M. Meta-analyses of adverse effects data derived from randomised controlled trials as compared to observational studies: methodological overview. *PLoS Med* 2011; 8: e1001026.
- 23 Clancy C, Collins FS. Patient-Centered Outcomes Research Institute: the intersection of science and health care. *Sci Transl Med* 2010; 2: 37.
- 24 Van Wyck Fleet J, Manberg PJ, Miller LL, Harto-Truax N, Sato T, Fleck RJ, Stern WC, Cato AE. Overview of clinically significant adverse reactions to bupropion. *J Clin Psychiatry* 1983; 44 (5 Pt 2): 191–6.

- 25** Siegrist CA, Lewis EM, Eskola J, Evans SJ, Black SB. Human papilloma virus immunization in adolescent and young adults: a cohort study to illustrate what events might be mistaken for adverse reactions. *Pediatr Infect Dis J* 2007; 26: 979–84.
- 26** Black S, Eskola J, Siegrist CA, Halsey N, Macdonald N, Law B, Miller E, Andrews N, Stowe J, Salmon D, Vannice K, Izurieta HS, Akhtar A, Gold M, Oselka G, Zuber P, Pfeifer D, Vellozzi C. Importance of background rates of disease in assessment of vaccine safety during mass immunisation with pandemic H1N1 influenza vaccines. *Lancet* 2009; 374: 2115–22.
- 27** Hubbard R, Farrington P, Smith C, Smeeth L, Tattersfield A. Exposure to tricyclic and selective serotonin reuptake inhibitor antidepressants and the risk of hip fracture. *Am J Epidemiol* 2003; 158: 77–84.
- 28** Petersen I, Gilbert RE, Evans SJ, Man SL, Nazareth I. Pregnancy as a major determinant for discontinuation of antidepressants: an analysis of data from The Health Improvement Network. *J Clin Psychiatry* 2011; 72: 979–85.
- 29** Coleman JJ, McDowell SE, Evans SJ, Gill PS, Ferner RE. Oversight: a retrospective study of biochemical monitoring in patients beginning antihypertensive drug treatment in primary care. *Br J Clin Pharmacol* 2010; 70: 109–17.
- 30** Waller PC, Evans SJ. A model for the future conduct of pharmacovigilance. *Pharmacoepidemiol Drug Saf* 2003; 12: 17–29.
- 31** Vanderwall DE, Yuen N, Al-Ansari M, Bailey J, Fram D, Green DV, Pickett S, Vitulli G, Luengo JJ, Almenoff JS. Molecular clinical safety intelligence: a system for bridging clinically focused safety knowledge to early-stage drug discovery – the GSK experience. *Drug Discov Today* 2011; 16: 646–53.
- 32** McCandless LC, Douglas IJ, Evans SJ, Smeeth L. Cutting feedback in Bayesian regression adjustment for the propensity score. *Int J Biostat* 2010; 6: Article 16. Available at <http://www.bepress.com/ijb/vol6/iss2/16> (last accessed March 2012).
- 33** Grimaldi-Bensouda L, Alperovitch A, Besson G, Vial C, Cuisset JM, Papeix C, Lyon-Caen O, Benichou J, Rossignol M, Abenhaim L, The GBS-PGRx Study Group. Guillain-Barré syndrome, influenza-like illnesses, and influenza vaccination during seasons with and without circulating A/H1N1 viruses. *Am J Epidemiol* 2011; 174: 326–35.
- 34** Herrett E, Thomas SL, Schoonen WM, Smeeth L, Hall AJ. Validation and validity of diagnoses in the General Practice Research Database: a systematic review. *Br J Clin Pharmacol* 2010; 69: 4–14.

RECEIVED

18 January 2012

ACCEPTED

30 January 2012

ACCEPTED ARTICLE PUBLISHED ONLINE

23 February 2012

CORRESPONDENCE

Professor Stephen J. W. Evans, Department of Medical Statistics, The London School of Hygiene & Tropical Medicine, Keppel Street, London WC1E 7HT, UK.
Tel.: +44 20 7927 2960
Fax: +44 20 7637 2853
E-mail: Stephen.Evans@lshtm.ac.uk