Review Article | **Open Access**

Randomised Controlled Trials and 'Unexpected' Adverse Events Associated with Newly Released Drugs: Improvements in Pharmacovigilance Systems are Necessary for Real-Time Identification of Patient Safety Risks

Margaret T. Whitstock[1]*, Christopher M. Pearce[2] and Elizabeth J. Eckermann[1]

[1]School of Humanities and Social Sciences, Faculty of Arts & Education, Deakin University, Australia
[2]Department of General Practice, Faculty of Medicine, Nursing and Health Sciences, Monash University, Australia

Abstract

Adverse drug events are one of the major causes of morbidity in developed countries, yet the drugs involved in these events have been trialled and approved on the basis of randomised controlled trials (RCTs), regarded as the study design that will produce the best evidence.

Though the focus on adverse drug events has been primarily on processes and outcomes associated with the use of these approved drugs, attention needs to be directed to the way in which the RCT study design is structured. The implementation of controls to achieve internal validity in RCTs may be the very controls that reduce external validity, and contribute to the levels of adverse drug events associated with the release of a new drug to the wider patient population.

An examination of these controls, and the effects they can have on patient safety, underscore the importance of knowing about how the clinical trials of a drug are undertaken, rather than relying only on the recorded outcomes.

As the majority of new drugs are likely to be prescribed to older patients who have one or more comorbidities in addition to that targeted by a new drug, and as the RCTs of those drugs typically under-represent the elderly and exclude patients with multiple comorbidities, timely assessment of drug safety signals is essential.

It is unlikely that regulatory jurisdictions will undertake a reassessment of safety issues for drugs that are already approved. Instead, reliance has been placed on adverse drug event reporting systems. Such systems have a very low reporting rate, and most adverse drug events remain unreported, to the eventual cost to patients and healthcare systems.

This makes it essential for near real-time systems that can pick up safety signals as they occur, so that modifications to the product information (or removal of the drug) can be implemented.

This paper will consider the impact on patient safety of procedural controls that can be exercised in the randomised controlled trial study design, comment on patient safety issues with current methods for post-marketing identification of adverse drug events, and consider methods which attempt to give better real-time identification of patient safety signals for newly released drugs.

Introduction

The list of drugs with regulatory approval subsequently implicated in adverse drug events (ADEs) in large numbers of patients continues to grow [1-3]. ADEs are now one of the major causes of morbidity in developed countries [4-7].

Analyses of ADEs often focus on events surrounding the patient, the prescribing staff, the pharmacy staff, the administering staff, and the choice of drug, and indeed many ADEs are associated with such procedural considerations. These considerations often start with the assumption that the drug *per se* has been approved for safe use in the marketplace in accordance with the approved directions. However, the root cause of a patient's ADE may lie in the design of the randomised controlled trials (RCTs) that supported the approval of that drug for marketing.

The elephant in the room, as far as newly approved drugs and patient safety is concerned, could be what has happened in the pre-approval clinical trial processes, and the extent to which the RCT study design can be unintentionally (or intentionally) manipulated in ways that avoid finding evidence of patient safety signals.

As Feinstein has argued, RCTs are "designed and analysed according to strategic policies about what questions the trial is intended to ask, what answers are to be obtained, what is to be done with the data, and who is to be convinced by the results [8]". From the point of view of

a manufacturer, an RCT is a tactical exercise to support approval to market a new drug.

Impact of evidence-based medicine (EBM) on the status of the RCT

The EBM focus on assigning relative values to evidence sources has prioritised evidence produced via RCTs as 'best evidence'. RCTs are now the only acceptable source of evidence in applications to most drug regulatory authorities for approval to market a new drug. For example, in the U.S., where 50% of the world's new drugs are launched [9-10], the Food and Drug Administration (FDA) is required by Federal legislation to accept only evidence from RCTs to support applications for approval to market a new drug. This prioritisation of RCTs has influenced the perception that evidence derived via an RCT is 'above reproach' in relation to patient safety.

Controls used in implementations of the RCT study design

Because the RCT study design is an explicitly scientific activity which focuses on the effects of a single intervention on modulating a single outcome, a range of controls are used to minimise the effects of other factors that could affect the outcome to be measured. These include controls that can be applied to select the patients who can participate in the clinical trial, controls that can be applied to the scope of the research questions that are the focus of the trial, and controls that can be applied to the statistical analyses used to assess the outcomes.

Effects on patient safety of comorbidity and co-medication exclusion protocols in RCTs

One of the major control issues which can affect post-approval patient safety risk is the application in RCTs of exclusion protocols relating to permitted comorbidities and co-medications for participants. These exclusion protocols can certainly protect clinical trial participants from undue risk, but they can also restrict the possibility of finding important evidence of potential risk.

When patients with the morbidity targeted by a new drug also have comorbidities and/or take co-medications that are specifically excluded from the RCTs of that new drug, then the evidence from those trials will not give sufficient information on the effects (and safety) of use of the new drug by those patients. When the majority of patients in a receiving population have excluded comorbidities and/or take excluded co-medications, this can translate to substantial post-marketing morbidity and mortality. However, drug licensing bodies such as the U.S. FDA do not require evidence that a clinical trial population is representative of routine clinical practice [11].

Data from a recent Australian arthritis population study analysing linked de-identified hospital separations and medications dispensing data showed that 68% of patients over 65 years of age who suffered from arthritis also suffered from cardiovascular or cerebrovascular diseases and/or took medications that were restricted in the RCTs of the arthritis drug rofecoxib [12]. Prior to its voluntary safety withdrawal in September 2004, it has been estimated that rofecoxib was implicated in more than a 1,000 ADEs in Australian patients, of which about 30% resulted in deaths [13]. An estimate of the number of excess cases of heart attack or sudden cardiac death in U.S. patients taking rofecoxib ranged from 88,000 to 139,000, of whom 30%-40% probably died [14]. The patients in rofecoxib RCTs were not typical of the patients presenting in the everyday clinical setting.

Effects on patient safety of age-based exclusion protocols in RCTs

Application of age-based participant exclusion protocols in RCTs

is of particular significance for older age groups. Avorn [7] notes that persons aged over 65 years form about 14% of the population in most industrialised countries, yet they consume nearly one-third of all drugs. The likelihood of experiencing an ADE increases as an individual patient's drug regimen becomes more complex, and this complexity is much more frequently found in elderly patients, who experience an increasing number of comorbid conditions as they age.

A new drug can be marketed as providing an effective therapy for a specific indication experienced by an elderly patient; however, that specific indication is frequently experienced by that elderly patient in a radically different context from that addressed in the clinical trials of the new drug. As Nelson [15], Wagemakers & Wolffers [16] and others [5] note, RCTs typically under-represent the elderly, and exclude patients with multiple comorbidities, yet these are the patients who are more likely to be prescribed drugs, and experience the highest rate of adverse events. Furthermore, older age patients are the group that is most frequently prescribed new drugs immediately they are approved. Solomon and Avorn observe that once a new drug is out in the pharmacy, 'unexpected' adverse events are common [5]. However these ADEs are not really 'unexpected'; rather, they are the outcome of a clinical trial program that did not include the range of patients who would be likely to be prescribed the new drug. Use of inappropriate medications by elderly patients remains an important patient safety and public health concern [17].

Effects on patient safety of non-inclusion of racial/ethnic groups in RCTs

Responses to drugs can be altered by a wide variety of individual characteristics that affect the pharmacokinetics of an individual drug, or class of drugs. Physiological processes such as absorption, metabolism, distribution or excretion of an individual drug or class of drugs can vary by age, sex, and, as Evelyn et al. note, they can also vary by racial group [18]. For example, Caucasians are more likely than Asians to have abnormally low levels of the liver metabolising enzyme cytochrome p450 2D6; African Americans do not respond well to several classes of antihypertensives. Evelyn et al. advise that attention needs to be paid to potential racial and ethnic differences when enrolling patients in clinical trials, in order to ensure patient safety for individuals in these racial and ethnic groups who may be prescribed the trialled drugs in the future [18].

Mason et al. [19] consider that excluding patients of ethnic minority groups from clinical trials is unethical, introduces substantial bias, and will mean that clinical trial findings are based on unrepresentative populations. They recommend increased awareness and monitoring of recruitment and retention of ethnic minority groups in clinical trials, and that analysis of data by ethnicity of subjects should be done consistently. Hussain-Gambles [20] agrees that ethnic minority groups are frequently under-represented in clinical trials and comments that this affects the generalisability of clinical trial findings. Both Evelyn et al. [18] and Hussain-Gambles recommend that strategies should be developed to increase awareness and monitoring of recruitment and retention of ethnic minority groups in clinical trials.

Effects on patient safety of pre-randomisation exclusion protocols in RCTs

Rothwell draws attention to the use of pre-randomisation run-in periods before a substantive RCT commences. These run-in periods can be implemented in order to identify and exclude patients who show poor compliance, or patients who show signs that the treatment to be trialled is ineffective [11]. The selection of participating centres

from hospitals or specialist care, as opposed to primary care, is also a pre-RCT selection issue, as patients selected from these settings will have a different level of the relevant morbidity from those managed in primary care, and can also have a different level of disease management [11].

Effects on patient safety of selectivity of research questions in RCTs

When selectivity in establishing participant exclusion protocols is added to selectivity in the formulation of research hypotheses, the possibility of finding important evidence of potential patient safety risk is further reduced. The definition of the research problem, and the selection of variables to be evaluated, can both be made from a position of pre-specified interest in anticipated results [8]. If the methodology used in a clinical trial pays no attention to a particular issue, then this can contribute to the absence of safety signals concerning that issue [21].

Again, using rofecoxib as an illustration, it is interesting to see that cardiovascular (CVD) endpoints were adjudicated in early clinical trials that were submitted to the FDA, but adjudication of these endpoints in later trials was modified when it appeared that the drug could be associated with thrombotic morbidity [22-23]. This modification of endpoints was added to an increased focus on controlling participant enrolment in RCTs. This combination acted in synergy to reduce the possibility of finding adverse evidence concerning rofecoxib and CVD risk [22-23].

Psaty and Furberg consider that inattention to cardiovascular events in trials of COX-2-selective NSAIDs contributed to minimising the possibility of uncovering evidence of cardiovascular harm, and as a result, only a small number of adverse cardiovascular events were recorded in studies that were not designed to assess cardiovascular outcomes [21]. In testimony to a 2005 U.S. Senate Committee on drug safety, Psaty argued that "the failure to pose a question often precludes the possibility of obtaining an answer" [24].

Effects on patient safety of using composite endpoints in RCTs

Designers of RCTs can combine clinical events (sometimes, events of very different severity) in their primary outcome measure. This can produce a measure of the overall effect of the treatment (and affords greater statistical power), but is problematic when used to guide treatment [11]. Kotaska notes that it is "easier to show a statistical difference in a combined end point rather than a single end point. Yet combined end points can be misleading" [25].

For example, in reviewing the 2002 Heart Protection Study involving the drug simvastatin, Conradi and Taylor note that the composite endpoint was the "first major vascular event", which included a range of fatal and non-fatal cardiac events or strokes, as well as amputations (precipitated by vascular occlusion), or admission to hospital for unstable angina [26]. They consider that "this arbitrary account of events and procedures does not apply to the risks of an individual patient, and cannot be used to explain the potential benefit of a proposed treatment" [26].

Another illustration is provided in the VIGOR trial data. The numbers of patients in the naproxen and rofecoxib arms were equivalent in terms of patient years at risk. The number of patients experiencing any thrombotic cardiovascular composite endpoint (cardiovascular death, myocardial infarction [MI] or cerebrovascular accident) in the naproxen arm was 18, and in the rofecoxib arm that figure was

35. However, whilst only 4 of those 18 adverse cardiovascular events in the naproxen arm were MIs, 20 of the 35 adverse cardiovascular events in the rofecoxib arm were MIs [27]. However, the published report of the VIGOR RCT recorded that the overall rates of death from cardiovascular causes were similar for the rofecoxib and naproxen arms [28].

Jüni et al. undertook a meta-analysis of RCTs of rofecoxib, and they conclude that Merck's use of composite cardiovascular endpoints in reports concerning rofecoxib "will have diluted any increase in risk of myocardial infarctions" [29]. Psaty and Furberg argue that use of composite endpoints "will tend to drive the relative risk toward the null and enhance the chances of finding a non-inferiority" status for a new drug undergoing clinical trial [30].

Effects on patient safety of use of surrogate endpoints in RCTs

Another potential safety risk is the acceptability of the use of surrogate endpoints in RCTs as substitute markers for clinical endpoints. These were introduced to offset the problem that testing the capacity of a new drug to prevent or modify an undesired long-term adverse clinical outcome could take a long time, because the adverse clinical outcome itself may take a long time to develop. It is considered not feasible to expect patients in active treatment groups to continue for the requisite time for the possibility of long-term effects becoming established, and it is judged improper to expect placebo groups to remain effectively untreated for an extended period. In other situations, for example, HIV research, it is considered essential to expedite the possibility of finding a means of treating a potentially serious morbidity state.

However, use of surrogate endpoints can have an adverse effect on patient safety, as the surrogate endpoint becomes the focus of research. The effect on patient safety through the use of a surrogate endpoint is illustrated in the clinical trials of rosiglitazone. The long-term clinical endpoint was the reduction of CVD in patients with Type II diabetes. More than 65% of deaths in patients with Type II diabetes are from cardiovascular causes, and one of the principal contributors to the onset of CVD in these patients is the prevalence of persistent high blood sugar levels. Reduction of blood sugar levels was thus regarded as an acceptable surrogate for the capacity of the new drug to reduce CVD, and this endpoint was to be measured by reduction in glycated haemoglobin levels [31].

However the initial RCTs of rosiglitazone were not adequately powered to determine the effects of rosiglitazone on the microvascular and macrovascular complications of diabetes, including the clinical endpoint of CVD. It was not until the numbers of MIs increased in patients taking rosiglitazone that evidence was slowly accumulated that the drug itself directly caused MIs. The time lapse in establishing that this was the case was protracted. The FDA Office of Surveillance and Epidemiology has calculated that there were between 41,000 to 205,000 major cardiovascular events in Type II diabetics which were potentially attributable to rosiglitazone therapy between its approval in 1999 and 2006 [32].

Effects on patient safety of the dosage levels used in RCTs

Another issue which can skew the evidence in RCTs of new drugs is the use of supratherapeutic dosing of highly selected clinical trial populations in short-haul clinical trials. Short-term administration of a dose known to be considerably in excess of that necessary to provide symptomatic relief is sometimes used as a surrogate means of estimating the boundaries of safety, tolerability and efficacy of a new drug, as well as the risk of the new drug over time.

Dosage guidance that is informed by the new drug manufacturer's exercise to replicate real-world tolerability, efficacy, and short- and long-term safety does not reliably translate into appropriate dosage guidance for an unscreened real-world clinical population. Adverse events related to drug tolerability, efficacy and patient safety will often show up when the drug is released into the community where patients with compromised health status are prescribed the new drug.

The frequent use of contraindicated dosage levels of comparator drugs also makes it difficult to estimate the comparative safety of a new drug. The strategy of under-dosing or overdosing of comparator drugs is implemented in clinical trials of a new drug in order to demonstrate its superiority in comparison with the comparator drugs. Dosage levels of comparator drugs were a significant issue in clinical trials of rofecoxib.

Rofecoxib was compared with a range of other drugs in various clinical trials, but these comparator drugs were often administered at levels not recommended by the manufacturers, as recorded in the relevant Product Information (PI). Sometimes the comparator drug dose was administered at a level that was higher than the manufacturer's recommended dose, and sometimes it was administered at a lower level. The apparent intent in adopting this strategy was that the performance of rofecoxib would be better in terms of the relevant symptom to be assessed, and better in terms of fewer unwanted side-effects.

Effects on patient safety of the pooling of RCTs

Pooling data from a number of clinical trials and presenting them as if they were a planned multi-site single RCT is also a way of presenting evidence that can make it difficult to estimate the comparative safety of a new drug. A study by Langman et al. [33] was presented as a "planned combined analysis" of eight trials, when these eight trials had different endpoints, were of different durations, used different dosage levels of the trialled drug in different trials, used different comparators in different trials, and different dosage levels of the those comparators in different trials. Statistical analyses using pooled data from a range of clinical trials have the potential to mask the possibility of identifying patient safety signals.

Effects on patient safety of selectivity in statistical techniques used in RCTs

Other important concerns with the clinical status of evidence from RCTs are related to the emphasis on numerical measurability, and the use of the EBM-preferred statistical model in the design and analysis of RCTs.

Feinstein [8] observes that there are actually two aspects of statistical significance: quantitative significance and stochastic significance. Quantitative significance refers to the magnitude of an observed distinction within clinical trial results, and relates to the *clinical importance* of the observed difference; stochastic significance is a mathematical calculation that denotes the probability that an observed distinction within trial results might arise by chance alone.

The $p < 0.05$ stochastic boundary is the accepted indicator of the significance of an outcome from a clinical trial; however when assessing this statistically derived measure of significance, it is important to remember that, in essence, statistical significance is *not* a measure of clinical significance [8].

There is no standard for comparing quantitative (clinical) significance, because the importance of a quantitative clinical difference will vary, depending on what is the subject of the clinical trial. In some trials, small quantitative differences can have significant clinical importance; in other trials, even large quantitative differences may not have significant clinical importance. For example, the accepted measure of statistical significance ($p < 0.05$) suggests that a result is significant if it has only a 1 in 20 likelihood of being due to chance. However the clinical significance of that finding is far greater if that chance translates as death or serious morbidity, compared with, say, a side-effect such as nausea.

Cochrane [34] also comments that it is often possible to achieve a result that is statistically significant but which may be clinically unimportant, and he notes that these estimates of significance are very dependent on the numbers enrolled in the clinical trials. Feinstein [8] notes that "for the stochastic aspects of significance, nothing will be significant if the group size is too small, and anything can be significant if the group is big enough". The presence of the $p < 0.05$ boundary for significance is a statistical artefact.

Using stochastic probability techniques to calculate the significance of RCT results allows mathematical constructs to decide what is or is not clinically important, and what may be clinically unsafe may be masked in the numbers.

Kerridge et al. note the patient safety issues relating to within-trial averaging, and they argue that "the overall results of a trial represent an average effect, and even within a trial population, some will experience a greater than average improvement in outcomes, while others may suffer harm" [35]. Kotaska confirms this in his observation that the mean outcome for all participants means that clinically important safety factors relating to individuals can be lost [25].

Overall, then, the range of controls and protocols that can be used in the design, conduct, and analysis of RCTs do not necessarily render the evidence derived via an RCT as 'above reproach' in relation to patient safety.

Who is responsible for identifying patient safety issues associated with new drugs?

The process of analysing patient safety issues in the RCTs submitted in support of an application for approval of a new drug is the statutory responsibility of national drug regulatory authorities. These same authorities are responsible for setting up systems to record details of post-marketing adverse drug events.

However assessing pre-approval and post-marketing drug safety is costly and time-consuming, and both of these processes can be subject to political pressures.

Effects on patient safety of funding agreements between manufacturers and drug regulatory authorities

Funding agreements for regulatory agencies have been implemented by various national regulatory authorities as a means of cost containment (e.g. the U.S. Prescription Drug User Fee Act, or the Australian Therapeutic Goods Amendment Act). These agreements require manufacturers of new drugs to pay the 'independent' national regulatory agency for the new drug approvals process, and, as a quid pro quo, these agreements allow manufacturers to set statutorily agreed performance requirements relating to the time taken to complete the approvals process.

Effects on patient safety of the U.S. prescription drug user fee Act

To illustrate the influence of the joint government and manufacturer

agreements, the current version of the U.S. Prescription Drug User Fee Act (PDUFA) requires that the FDA is to complete the assessment process for 90% of applications for new Priority Drugs in 6 months, and 90% of applications for new Standard Drugs are to be completed in 10 months [36]. It is the manufacturer who makes the case that a new drug is an advance over currently available therapy, which then places that drug in the Priority Drug group.

The FDA new drug approval process is now amongst the shortest in the world, and the average time it takes for the FDA to approve a new drug has dropped by 40% since the first implementation of PDUFA in 1992. This has contributed to raising the U.S. share of the world's new drug launches from 8% to 50% [9]. From the patient safety perspective, this could be considered as representing an increased worldwide exposure to drug safety issues generated within the U.S. regulatory framework.

Carpenter et al. [37] have explored the impact of accelerated PDUFA deadlines on the quality of decision-making in drug safety assessments, and they find that pre-deadline approvals are associated with three to five times the rate of safety-based withdrawal from the global market for all classes of drugs. The shortened times of the FDA new drug approvals process involves approvals being made early in the clinical development process, reducing the time-window for recognising adverse drug events [38].

For example, Dr M. Villalba, in her Medical Officer Review of the New Drug Application NDA 21-042 (rofecoxib), expressed concern in her official summary about a group of six-week trials that were submitted by the manufacturer as part of the supporting data in the application for approval to market rofecoxib. She noted that the data in these six-week trials suggested that thromboembolic events were more frequent in patients receiving rofecoxib than placebo [39]. Villalba noted that "with the available data, it is impossible to answer with complete certainty whether the risk of cardiovascular and thromboembolic events is increased in patients on rofecoxib" and she recommended larger studies to answer this and other safety comparisons" [39].

Rofecoxib had been nominated by the manufacturer as a Priority Drug, and the review process was required to be completed in six months. The drug was approved in this time-frame, and it went on to be the cause of many thousands of cases of cardio- and cerebrovascular morbidity and mortality until it was withdrawn. Rofecoxib was approved when not enough was known about its action in precipitating adverse thromboembolic processes. The antibiotic drug telithromycin is another example of a drug with an expedited release onto the market in 2004 without appropriate safety testing. Even in the short term, it caused significant morbidity and mortality before it was withdrawn. Users consistently experienced symptoms ranging from impaired vision to liver toxicity, yet these symptoms were not explored fully in the pre-release stage [40].

Ironically, the PDUFA arrangements were initiated partly as a result of consumer dissatisfaction with the time taken to get new drugs (especially those for HIV) on the market. Yet by 2007, even before the rofecoxib case had come to a head, there was concern that the FDA regulatory body was becoming accountable to the pharmaceutical industry it regulated [38]. Now it appears to be the case that new drugs are being approved in too short a time-frame, and before the 2007 renewal of the PDUFA agreement, there were many concerned scientists and FDA staff who lobbied to have the FDA funded by government appropriation, with no controlling involvement by industry.

In interviews conducted in 2003 by the U.S. Department of Health and Human Services, 58% of FDA staff reported that the 6-month Priority Drug approval time-frame does not allow for in-depth, science-based reviews, and that this was contributing to the likelihood of adverse drug events remaining unrecognised until after the new drug is approved [41]. This effectively shifts issues relating to adverse drug events to national adverse event reporting systems.

Effects on patient safety of relying on adverse event databases

Relying on checking adverse event reporting databases to provide information on risks to patient safety is problematic. Reports to such databases are inherently subject to Type I and Type II errors, as there is no accurate information on the actual number of events in the exposed population or the actual number of exposed individuals [42]. Adverse event reports are also subject to major under-reporting. In 2004, Trontell, of the FDA Office of Surveillance and Epidemiology, reported that the U.S. FDA MedWatch / AERS (Adverse Event Reporting System) may receive as little as 1% of all adverse events which occur in the population [43].

Troy, former chief counsel at the FDA, commented in 2007 that "the FDA generally assumes that only 1 in 10 adverse events is reported [44]". Other drug regulatory jurisdictions suffer from the same problems. In practice, reported ADEs are regarded as being outliers until the rate and the number of reported adverse events force another look at the way in which the relevant RCT was designed.

In regulatory jurisdictions where reporting of potential adverse drug events is voluntary, even if there are electronic options available for reporting, this 'outlier' issue will apply. The value of the mandatory aspect of the process vs. the voluntary aspect is shown in the managed manual reporting regime in the U.K. As part of the U.K. Prescription-Event Monitoring (PEM) system [45], managed by the independent Drug Safety Research Unit (DSRU), all U.K. GPs are required to complete a form reporting on a broad range of medical events that have been recorded in a patient's notes during a specific time-period since the patient started treatment with a new drug. This manual reporting process is undertaken for the first 10,000 users of a new drug. Removing the need for the prescribing doctor to give an opinion about whether an event might have been caused by the new drug provides the opportunity to identify adverse events that may not have been suspected as being due to the drug that is under post-marketing surveillance.

Reporting in this manner neutralises the problems with Type I and Type II errors which occur in the FDA AERS. Data components are individually examined, and large cohorts are processed electronically. The DSRU states that the Prescription-Event Monitoring is the only method of post-marketing surveillance in Europe that can evaluate the safety of new medicines soon after launch [45].

Effects on patient safety of manufacturers' responses to risks identified after the approval of a new drug

Should a national regulatory authority consider that the accumulation of adverse event reports indicates patient safety concerns, the manufacturer can be required to make changes to the product labelling. However, the wording to be used, and the placement of the labelling changes, remain the prerogative of the manufacturer. The process is one not welcomed by manufacturers, as they consider that such changes will affect sales, and there have been egregious cases where procrastination has meant that details of risks to patient safety were not communicated to prescribers and patients for extended

periods [46-47]. In one case, two manufacturers refused to comply with a labelling change required by a national regulator: one manufacturer took the issue to the Federal Court; the other petitioned the relevant Federal Minister [48].

In the U.S., the Food and Drug Administration Amendments Act 2007 (FDAAA 2007) places an upper limit of 80 days on the process of settling safety labelling issues required after post-marketing studies [49]. Assuming that this process is complied with, there is at least some expectation that U.S. patients and their physicians will be made aware of safety issues in a more helpful time-frame.

Overall, the current post-approval identification of, and action on, patient safety signals associated with new drugs do not well serve the consumers of these drugs, and it is essential that processes that allow closer to real-time assessment of ADEs can be developed.

Before focusing on these processes, we include a brief discussion on a side issue on the proposed responsibility of practitioners to assess reports of RCT evidence.

Should practitioners assess new drug RCTs themselves?

EBM praxis proposed initially that practitioners themselves should locate evidence from published RCTs of treatments that were possibly applicable to their patients, and then assess these RCTs for compliance with an ideal RCT study design and for suitability for their patients. The process was represented as "a simple, logical process for reasoning and decision making" [35]. But time proved critical on many fronts [50], and the emphasis shifted to group development of clinical practice guidelines based on RCTs. The outcome is multiple clinical guidelines with multiple requirements that are not necessarily compatible.

Many RCTs are published multiple times in different journals using different orders of author names, and different titles for the trials. It is possible for practitioner assessors of new drugs to think that certain positive results are supported by many RCTs when in fact they are duplicates. The true levels of ADEs are masked. In an editorial on the fair conduct and fair reporting of clinical trials, Rennie argues that covert reporting of the same data in clinical trials artificially skews the balance of opinion in favour of a new drug, giving an impression of wide support for the efficacy of an intervention [51].

Meta-analyses of RCTs are also proposed as an aid to practitioners to assess the evidence from RCTs. The meta-analysis process applies specific statistical operations to harness the potentially greater power of larger numbers through combining studies to arrive at an overall assessment of the effectiveness of a particular intervention. However, small or poorly conducted clinical trials produce different effect estimates than those trials that are methodologically rigorous. The importance of this exaggeration of benefits factor is that when data from these trials are included in meta-analyses, they contribute to a skewed result [52].

An examination of 11 randomly selected meta-analyses of 127 RCTs, involving over 10,000 patients, showed that when the results of low- and high-quality RCTs were pooled in the relevant meta-analyses, there was a significant 30-50% exaggeration of the efficacy of the trialled interventions across the 11 meta-analyses as a result of the inclusion of the low-quality RCTs [53]. The value to clinicians of evidence from meta-analyses is only as useful as the methodology that underpins each RCT.

Surveillance methods which attempt to give closer to real-time identification of patient safety signals for newly released drugs

RCTs are primarily about assessing how a new drug works for the population in the clinical trials, and not about which patients in the wider community can benefit, and even in the case of well-conducted clinical trials, it is expected that there will be unaddressed patient safety risks for members of the wider population for whom the new drug is prescribed.

Adverse event monitoring systems relying on the reporting of ADEs that occur in the wider population exposed to a new drug will suffer from a time-lag until there are sufficient data that indicate there is a growing problem. This time-lag will be greater when the reporting of ADEs is voluntary, as noted earlier.

The search for improved identification of ADEs has focused on what systems can be put in place that can prospectively identify risk. The key to closing the gap between the release of a drug, and the identification of ADEs associated with its use is access to data that can link patients with medications and with diagnoses. This requires the development of very large linked systems, and of techniques that will support the analysis of the data available.

Examples of the capacity of large integrated systems to detect early safety signals associated with new drugs are seen in the EU-ADR Project [54], the HMORN (CERT) Project [55-56] and the Sentinel System [57], each discussed below. The aim of these projects is to exploit information from electronic healthcare record databases to develop an integrated system for the early detection of drug safety signals.

Also discussed is an Australian initiative designed to offset some of the patient safety issues associated with the use of RCT controls. The model system works by prospectively identifying patient groups potentially at risk of an ADE if prescribed a newly released drug. The system uses linked de-identified national medications dispensing and hospital separations diagnoses data to establish the morbidity profiles of patients who have the morbidity (or morbidities) that are the target of a new drug. These are then compared with the patient profiles of the participants selected to participate in the clinical trials of that new drug in order to assess safety risk.

The EU-ADR Project

The EU-ADR (Exploring and Understanding Adverse Drug Reactions) is a proof-of-concept project drawing on retrospective anonymised aggregated demographic, outpatient, inpatient, and medication prescription data from eight databases in four European countries (Denmark, Italy, the Netherlands and the U.K.). All eight systems are record-linkage systems in which drug dispensing data are linked to a registry of hospital separations diagnoses (as well as other registries). Data were pooled via a distributed network approach by generation of common input data, followed by local aggregation through custom-built software (Jerboa®). The distributed study population was 19.6 million individuals, which corresponded to 60 million person-years of follow-up.

The retrospective proof-of-concept exercise using EU-ADR linked data showed a consistent association of a specified drug class (NSAIDs) with an increased risk of a specified morbidity (upper gastrointestinal bleeding) in all eight databases. When the system is used in the future for real-time early detection of drug safety signals, it is proposed that the safety signals will be substantiated by semantic mining of the

relevant literature, and computational analysis of pharmacological and biological information on drugs, molecular targets, and pathways.

The (HMORN) (CERT) Project

The objectives of the U.S. Health Maintenance Organization Research Network (HMORN) Center for Education and Research (CERT) project were to evaluate the utility of automated healthcare claims data for near-real-time adverse event surveillance, and to identify methodological issues related to this. Nine participating health plans provided data for over 8 million members, with 13 million person-years of follow-up over a 6-year period.

The project assessed the ability to detect ADEs using retrospective data from nine health plans, using a maximised sequential probability ratio test (maxSPRT). The maximised sequential probability ratio test was used because to realise the full potential of prospective adverse event surveillance, accumulated drug exposure and adverse event experience should be evaluated as they accumulate in a way that avoids problems associated with repeated statistical tests on the same data.

The intent of sequential analysis is to quickly and efficiently detect signals of excess risk using data that are routinely collected by most public and private health insurers in the U.S., and the highly summarised data used in the analysis process preserves patient confidentiality. The investigators noted that positive safety signal detection does not imply a causal relationship, and all signals need to be evaluated for clinical importance.

The U.S. Sentinel System

The FDA announced the Sentinel System initiative in May 2008 [51]. This private-public system will eventually be able to search the electronic health data of a minimum of 100 million patients to determine the existence and/or severity of risk factors. Users of the Sentinel System will include government and private sector organisations and academia. The current infrastructure is supported by the FDA, but it is anticipated that Sentinel will be a distributed system (as is used in the EU-ADR project) managed by a consortium of interested parties operating as a public–private partnership.

Mini-Sentinel, the pilot program currently underway, has initially focused on developing the ability of participating health plans and other private organisations to create data files in a standard format and to maintain control of those files. These private organisations perform most analyses of their own data by running computer programs distributed by a coordinating centre, and they provide consistent summarised results for review by the FDA.

It is anticipated that the FDA will soon begin to actively monitor the data, seeking answers to specific questions about the performance of medical products [51].

Using clinical trial and linked administrative health data to prospectively identify potential patient safety risk

Whitstock et al. [12] trialled a model process that used clinical trial data and linked administrative health data to support a prospective assessment of patient groups who could potentially be at risk of an adverse drug event if they are prescribed a newly released drug in the context of their age, gender, comorbidities and/or co-medications.

Using clinical trial data required under U.S. law to be made publicly available, it is possible to develop a profile of the participants in the RCTs of a new drug. Information includes details on study designs, primary and secondary outcome measures, participating patient demographics, eligibility criteria (including exclusion protocols), withdrawals and exclusions from the final analysis, tables of values for primary and secondary outcome measures, and statistical analyses.

Using linked Australian de-identified administrative health data on patient hospital separation diagnoses and patient medications dispensed, it is possible to develop a morbidity profile of the patients who suffer from the morbidity (or morbidities) targeted by that new drug. The clinical trial information and the linked morbidity and medication data are compared to assess which patient groups could potentially be at risk of an adverse drug event associated with use of the new drug.

An advisory can be circulated to physicians to recommend care prescribing of that drug to patients with comorbidities or taking co-medications that have been excluded in the RCTs of that new drug, at least until more is known.

As the majority of new drugs are likely to be prescribed to older patients who have one or more comorbidities in addition to that targeted by a new drug, and as the RCTs of those drugs typically under-represent the elderly and exclude patients with multiple comorbidities [5, 15-16], undertaking an exercise such as this could reduce the numbers of ADE in patients who have health profiles that would have excluded them from the RCTs of a new drug.

Conclusion

The RCT study design may be the prioritised source of best evidence, but the procedural controls that can be exercised within that study design mean that applying this evidence can pose risks to patient safety. Identification of these risks must be made as soon as possible. Currently available adverse event reporting systems do not have the capacity to support that needed timely identification. Real-time or prospective identification of patient safety risks is essential in an environment where more and more new drugs are being launched.

References

1. U.S. Food and Drug Administration (2011) FDA Drug safety podcasts.

2. U.S. Food and Drug Administration (2011) Postmarket drug safety information. Index to drug specific information.

3. ClassAction.org (2011) Defective drug injuries and side effects. Dangerous drug class action lawsuits, recalls, and safety alerts.

4. Miller GC, Britt HC, Valenti L (2006) Adverse drug events in general practice patients in Australia. Med J Aust 184: 321-324.

5. Solomon DH, Avorn J (2005) Coxibs, science, and the public trust. Arch Intern Med 165: 58-160.

6. McGavock H (2004) Prescription-related illness: a scandalous pandemic. J Eval Clin Pract 10: 491-497.

7. Avorn, J (1997) Including elderly people in clinical trials: better information could improve the effectiveness and safety of drug use. BMJ 315: 1033-1034.

8. Feinstein AR (1985) Clinical epidemiology: the architecture of clinical research. W. B. Saunders Co., Philadelphia.

9. Bradley WG, Halle M (2007) Prescriptions for drug safety: reforming U.S. prescription drug regulations to protect consumers. Washington DC, Center for American Progress.

10. Psaty BM, Charo RA (2007) FDA responds to Institute of Medicine drug safety recommendations - in part. JAMA 297: 1917-1920.

11. Rothwell PM (2005) Treating individuals 1: External validity of randomised trials: 'To whom do the results of this trial apply?' Lancet 365: 82-93.

12. Whitstock MT, Pearce CM, Ridout, SC, Eckermann EJ (2011) Using clinical trial data and linked administrative health data to reduce the risk of adverse events associated with the uptake of newly released drugs by older Australians: a model process. BMC Public Health 11: 361

13. Seccombe M (2005) Arthritis drug's side effects linked to 300 deaths. Sydney Morning Herald.

14. Graham DJ (2004) Testimony of David J. Graham, Washington, DC, U.S. Senate Finance Committee.

15. Nelson MR (2004) Research in the aged. Aust Fam Physician 33: 773.

16. Wagemakers A, Wolffers I (2002) Wel slikken, niet testen = Take without testing: Ouderen worden vaak uitgesloten van geneesmiddelenonderzoek. Medisch Contact 57: 1144-1146. Translation provided by Expertisecentrum LEEFtijd = Expertise Centre for Age in Society, Utrecht, Netherlands.

17. Simon SR, Chan KA, Soumerai SB, Wagner AK, Andrade SE, et al. (2005) Potentially inappropriate medication use by elderly persons in U.S. Health Maintenance Organizations, 2000–2001. J Am Geriatr Soc 53: 227-232.

18. Evelyn B, Toigo T, Banks D, Pohl K, Gray B, et al. (2001) Participation of racial/ethnic groups in clinical trials and race-related labeling: a review of new molecular entities approved 1995-1999. J Natl Med Assoc 93: 18-24.

19. Mason S, Hussain-Gambles M, Leese B, Atkin K, Brown J (2003) Representation of South Asian people in randomised clinical trials: analysis of trials' data. BMJ 326: 1244-1245.

20. Hussain-Gambles M (2003) Ethnic minority under-representation in clinical trials: Whose responsibility is it anyway? J Health Organ Manag 17: 138-143.

21. Psaty BM, Furberg CD (2005) COX-2 inhibitors: lessons in drug safety. NEJM 352 : 1133-1135.

22. Thomas WJ (2006) The Vioxx story: Would it have ended differently in the European Union? Am J Law Med 32: 365-380.

23. Singh G (2004) Testimony of Gurkirpal Singh, Washington, DC, U.S. Senate Finance Committee.

24. Psaty B M (2005) Testimony of Bruce M. Psaty, Washington, DC, U.S. Senate.

25. Kotaska A (2004) Inappropriate use of randomised trials to evaluate complex phenomena: case study of vaginal breech delivery. BMJ 329: 1039-1042.

26. Conradi P, Taylor D (2005) The Heart Protection Study and bias in the interpretation of RCTs. J Eval Clin Pract 11: 405-407.

27. U.S. Food and Drug Administration, Center for Drug Evaluation and Research, Division of Cardio-Renal Drug Products, and Targum SL (2001) Cardiovascular safety review [of] rofecoxib.

28. Bombardier C, Laine L, Reicin A, Shapiro D, Burgos-Vargas R, et al. (2000) Comparison of upper gastrointestinal toxicity of rofecoxib and naproxen in patients with rheumatoid arthritis. NEJM 343: 1520-1528.

29. Jüni P, Nartey L, Reichenbach S, Sterchi R, Dieppe PA, et al. (2004) Risk of cardiovascular events and rofecoxib: cumulative meta-analysis. Lancet 364 : 2021-2029.

30. Psaty BM, Furberg CD (2007) The record on rosiglitazone and the risk of myocardial infarction. NEJM 357: 67-69.

31. Nissen SE, Wolski K (2007) Effect of rosiglitazone on the risk of myocardial infarction and death from cardiovascular causes. NEJM 356: 2457-2471.

32. Nissen SE and Wolski K (2010) Rosiglatazone revisited: an updated meta-analysis of risk for myocardial infarction and cardiovascular mortality. Arch Intern Med 170: 1191-1201.

33. Langman MJ, Jensen DM, Watson DJ, Harper SE, Zhao P-L, et al. (1999) Adverse upper gastrointestinal effects of rofecoxib compared with NSAIDs. JAMA 282: 1929-1933.

34. Cochrane AL (1972) Effectiveness and efficiency: random reflections on health services. Abingdon, Berkshire: Nuffield Provincial Hospitals Trust.

35. Kerridge I, Lowe M, Henry D (1998) Ethics and evidence based medicine. BMJ 316: 1151-1153.

36. U.S. Food and Drug Administration, Department of Human Services (2010) FY 2010 performance report to the President and Congress for the Prescription Drug User Fee Act.

37. Carpenter D, Bowers J, Grimmer J, Moffitt S, Nall C, et al. (2007) Deadline effects in regulatory drug review: a methodological and empirical analysis. The Robert Wood Johnson Foundation Working Paper WP-35. Princeton, New Jersey.

38. Avorn J (2007) Paying for drug approvals: who's using whom? NEJM 356: 1697-1700.

39. Villalba ML (1999) Vioxx: Excerpts from primary review of NDA 21-042 – Osteoarthritis.

40. Graham D (2006) Telithromycin and acute liver failure. NEJM 355: 2260-2261.

41. U.S. Department of Health & Human Services, Office of [the] Inspector General (2003) FDA's review process for new drug applications: a management review. Washington DC.

42. Strom BL (2004) Potential for conflict of interest in the evaluation of suspected adverse drug reactions: a counterpoint. JAMA 292: 2643-2646.

43. Trontell A (2004) Expecting the unexpected: drug safety, pharmacovigilance, and the prepared mind. NEJM 351: 1385-1387.

44. Drazen JM, Rainey J, Begg H, Butler AS (2007) Adverse drug event reporting: the roles of consumers and health-care professionals: workshop summary. Ch. 2 Current adverse event reporting. Ch 3. Active surveillance systems.

45. Drug Safety Research Unit [U.K.] (2011) Prescription-event monitoring.

46. United States Court of Appeals, Third Circuit (2008) IN RE: Merck & Co. Case Numbers 07-2431, 07-2432.

47. Kelly J (2005) Fallout from Merck losing its first Vioxx lawsuit. Medscape Medical News.

48. Carlisle W (2005) What went wrong with Vioxx? ABC Radio National, Background Briefing.

49. U.S. Congress, House of Representatives (2007) H.R. 3580: Food and Drug Administration Amendments Act of 2007, Public Law No: 110-185.

50. Whitstock MT (2003) Seeking evidence from medical research consumers as part of the medical research process could improve the uptake of research evidence. J Eval Clin Pract 9: 213–224.

51. Rennie D (1999) Fair conduct and fair reporting of clinical trials. JAMA 282: 1766-1768.

52. Puffer S, Torgerson DJ, Watson DJ (2005) Cluster randomized controlled trials. J Eval Clin Pract 11: 479-483.

53. Moher D, Pham B, Jones A, Cook DJ, Jadad AR, et al. (1998) Does quality of reports of randomised trials affect estimates of intervention efficacy reported in meta-analyses? Lancet 352: 609-613.

54. Coloma PM, Schuemie MJ, Trifiro G, Gini R, Herings R, et al. (2011) Combining electronic healthcare databases in Europe to allow for large-scale drug safety monitoring: the EU-ADR Project. Pharmacoepidemiol Drug Saf 20: 1-11.

55. Brown JS, Kulldorff M, Chan KA, Davis RL, Graham D, et al. (2007) Early detection of adverse drug events within population-based health networks: application of sequential testing methods. Pharmacoepidemiol Drug Saf 16: 1275-1284.

56. Brown JS, Kulldorff M, Petronis KR, Reynolds R, Chan KA, et al. (2009) Early adverse drug event signal detection within population-based health networks using sequential methods: key methodologic considerations. Pharmacoepidemiol Drug Saf 18: 226-234.

57. Behrman RE, Benner JS, Brown JS, McClellan M, Woodcock J, et al. (2011) Developing the Sentinel System: a national resource for evidence development. NEJM 364: 498-499.