Postmarketing Surveillance and Adverse Drug Reactions
Current Perspectives and Future Needs

Timothy Brewer, MD, MPH
Graham A. Colditz, MD, DrPH

With the use of any medication comes the possibility of unintended consequences. These events, when harmful, often are referred to as adverse drug reactions (ADRs). While the nature of the intended benefit from using the medication is known, ADRs can include both predictable and unpredictable events. Premarketing trials frequently do not have sufficient power to reliably detect important ADRs, which may occur at rates of 1 in 10,000 or fewer drug exposures. Premarketing trials also lack the follow-up necessary to detect ADRs widely separated in time from the original use of the drug or delayed consequences associated with long-term drug administration. These trials often do not include special populations such as pregnant women or children who may be at risk for unique ADRs or for an increased frequency of ADRs compared with the general population. Taken together, these limitations of premarketing clinical trials mean that, in the United States, the Food and Drug Administration (FDA) approval of a new drug does not exclude the possibility of rare but serious ADRs or common, delayed ADRs.

Safety is not an absolute concept. The seriousness of the underlying illness and the availability of alternative effective treatments will alter what are considered tolerable ADRs. For example, the toxic effects of many available chemotherapeutic agents would be unacceptable in drugs marketed for uncomplicated urinary tract infections.

In addition to ADRs, medication use may be associated with unintended consequences that are beneficial as well as detrimental. Postmarketing studies of hormonal therapy in postmenopausal women have shown a reduction in deaths from cardiovascular disease compared with nonusers, and oral contraceptive users have a lower risk of ovarian cancer than nonusers. In this article, however, we focus on methods used to uncover adverse effects.

Adverse drug reactions can be divided into 2 categories: events that otherwise occur rarely in the population and events that represent an increased frequency over a relatively common rate in the general population. These 2 categories of ADRs may be further subdivided into 3 groups based on the occurrence of the event relative to the use of the drug: those that occur shortly after initiation of drug use, those that occur with long-term use, and those that occur remotely after the drug has been discontinued. Both the frequency of the event, rare or relatively common, and the timing of the event relative to drug use influence the likelihood of detecting the ADR with different surveillance methods.

Author Affiliations: Channing Laboratory, Department of Medicine, Brigham and Women’s Hospital and Harvard Medical School (Drs Brewer and Colditz), and Technology Assessment Group, Harvard School of Public Health (Dr Colditz), Boston, Mass.

Corresponding Author and Reprints: Timothy Brewer, MD, MPH, Channing Laboratory, 181 Longwood Ave, Boston, MA 02115.

See also pp 830 and 841.
A number of methods have been used to identify previously unknown detrimental outcomes that may be attributable to the use of medications. These methods include premarketing clinical trials, postapproval spontaneous case reports, aggregate population-based data sources, computerized collections of data from organized medical care programs, and postmarketing studies. Combining data from similar sources, such as clinical trials, also has been suggested as a means of detecting ADRs. These methods vary in their utility for detecting unintended outcomes and for linking the outcomes with previous medication use. We examine the use of different methods to identify and confirm ADRs.

**CASE REPORTS**

More serious ADRs have been noted first in case reports than any other detection method. In a comparison of postmarketing cohort studies with spontaneous reporting for detecting ADRs, Rossi and colleagues found that none of 3 phase 4 studies detected new ADRs, while spontaneous reports of new ADRs were received for 2 of the 3 drugs. Case reports require only the suspicion that an adverse event may be related to the prior use of a drug and some mechanism for alerting others.

To improve the detection of previously unknown serious ADRs and knowledge about regulatory actions taken in response to ADR reports, the FDA introduced the MEDWATCH program in June 1993. Health care professionals are encouraged to report serious events suspected to be caused by medications, medical devices, special nutritional products, and other products regulated by the FDA. Serious events are those that lead to death, hospitalization, significant or permanent disability, or congenital anomaly or require medical or surgical intervention to prevent 1 of these events. Physicians may report ADRs by telephone, fax, or mail or through the FDA’s MEDWATCH Internet site (Table). Approximately 1 year after the introduction of MEDWATCH, the number and quality of ADR reports to the FDA increased. This increase, however, was attributed to improved reporting by pharmacists. Physician reports declined slightly during this period.

Despite the importance of physician reports for detecting ADRs, serious adverse events that may represent ADRs are underreported by physicians to either manufacturers or the FDA. MEDWATCH is 1 in a series of initiatives to increase and improve physician reporting of suspected ADRs. Educational programs, including direct mailings and presentations combined with a streamlined reporting process and feedback to physicians, have been shown to improve the number and quality of ADR reports. Payment of reporting fees increases ADR reporting rates, although reporting rates declined significantly after reimbursements were stopped.

The utility of case reports as a screening tool for ADRs is influenced by the frequency of the adverse outcome in the underlying population and the temporal relationship with the drug exposure. Unusual or rare events that occur during initial or long-term drug use are more likely to be detected by case reports than increases in common events or events that occur remotely in time from the medication use. Temafloxacin, a fluoroquinolone antibiotic, was withdrawn within 6 months of its introduction in 1992 because of the association between its use and hemolytic anemia in otherwise healthy individuals. Spontaneous reporting rapidly identified this ADR because it was rare in the general population and occurred within 1 week of drug use.

The association between valvular heart disease in younger women and the use of the appetite suppressants phentermine and fenfluramine took longer to identify with spontaneous reporting probably because the development of the ADR required a longer period of use. Even in this instance, however, its detection was aided by the fact that the ADR was an otherwise rare disease in a population with ongoing or recent exposure to the drug(s).

In announcing MEDWATCH, then-Commissioner Kessler wrote that the lack of spontaneous reports linking silicone breast implants with autoimmunelike disorders delayed the detection of this problem even though implants had been in use for approximately 30 years. Autoimmunelike symptoms are relatively common in women without implants, the increase in risk with exposure, if any, is likely to be small, and symptoms occur years after the initial exposure. Adverse drug reactions meeting this description are unlikely to be reliably detected by any spontaneous reporting system.

Additional limitations of spontaneous reporting include both erroneous reports and the fact that prescribing patterns and reporting rates are not linked. Comparisons of physicians’ reports of ADRs with expert reviewers’ opinions or with standardized assessment methods have demonstrated poor agreement between physicians and the other methods in assigning causality of the ADR to the medication.

In 1 study of almost 30 000 general practitioners in the United...
Kingdom, Inman and Pearce\textsuperscript{23} found that 10% of practitioners wrote approximately 40% of the prescriptions for recently released drugs. Furthermore, the more likely a physician was to prescribe a new drug, the less likely he or she was to submit an ADR report.\textsuperscript{23}

Patients are another potential source for case reports of suspected ADRs. As with physician reports, the quality of patient reporting has been raised as a concern. In a study that relied on reporting forms and telephone questioning, patients were less likely to attribute “events” to the prescribed medication than an expert panel that reviewed the event forms.\textsuperscript{24} Though patient reports were less sensitive than physician reports, large-scale reporting of events from patients might be valuable for earlier detection of symptomatic reactions to new drugs.\textsuperscript{24}

**SURVEILLANCE SYSTEMS**

The creation of computerized prescription and laboratory databases has greatly enhanced the ability of institutions and organizations to screen for known ADRs.\textsuperscript{25} Changes in medication orders, orders for antidote medications such as antihistamines or opiate antagonists, drug levels, and laboratory information such as *Clostridium difficile* toxin titers have all been used as screens. Screening adverse event monitors have been more effective in documenting ADRs than simplified voluntary reporting or educational programs.\textsuperscript{25,26} Hospital-based systems can greatly increase the reporting of known ADRs, but their value for identifying new, unknown ADRs remains unclear.\textsuperscript{20} Only ADRs that occur during hospitalization are recognized. Adverse drug reactions that occur after discontinuation of the offending medication may be missed by these systems. Since these systems rely on algorithms to detect ADRs, events unrelated to the algorithms go unnoticed.

Many hospital systems do not have a sufficient sample size to reasonably detect unknown ADRs. Government and private-insurer patient databases are another option for evaluating the nature and frequency of ADRs. Advantages of these data systems include their size and low study costs.\textsuperscript{8} Depending on the database used, investigators looking for ADRs may have the ability to link hospitalizations, outpatient visits, and prescription use. For example, linked vaccination records and hospitalizations were used to assess the risk of convulsions after diphtheria-tetanus-pertussis vaccination and febrile convulsions or idiopathic thrombocytopenia purpura after measles-mumps-rubella vaccination in the United Kingdom.\textsuperscript{27}

Population-based surveillance systems potentially may be used to recognize ADRs in which the adverse event also occurs in unexposed populations though at a reduced frequency, that occur after long-term use, or that occur remotely from the drug exposure. The latter requires databases that have been maintained for years and in which events can be linked with either current or previous medication use. The value of these databases in identifying new ADRs remains to be determined, but they should be explored for adverse events that spontaneous reports are less likely to detect.

One large-scale surveillance system currently used to identify adverse events is the Vaccine Adverse Event Reporting System (VAERS). VAERS is a unified national system managed jointly by the FDA and the Centers for Disease Control and Prevention (CDC).\textsuperscript{28} VAERS receives reports from the public as well as physicians, manufacturers, and public health clinics. VAERS data have been used to describe previously unreported vaccine adverse effects.\textsuperscript{28} Besides VAERS, which is a passive surveillance system, the CDC has initiated an active surveillance study of vaccine ADRs using 4 health maintenance organizations.\textsuperscript{29} Whether this active surveillance system will enhance the recognition of vaccine-related adverse events is unknown.

Using a large linked database minimizes potential errors such as underreporting or recall bias.\textsuperscript{30} Potential weaknesses with linked record systems include the accuracy of the linkages between record systems, the reliance on possibly inaccurate or incomplete records, and the time frame covered by the records. Validating reported diagnoses can be done to minimize inaccuracies in linked records,\textsuperscript{29} but it does not affect potential bias caused by using incomplete records. Despite these potential limitations, evidence suggests that rigorously established record linkage systems can provide estimates of ADRs. In a prospective epidemiologic study of coronary heart disease, computerized linkage alone was as effective as direct contact with patients in identifying ADRs.\textsuperscript{31}

**POSTMARKETING COHORT STUDIES**

As noted above, postmarketing cohort studies to detect unknown ADRs have been considered disappointing.\textsuperscript{9,11} Spontaneous reports will likely remain the most efficient way to detect rare adverse events that occur temporally with drug use. The value of postmarketing cohort studies may be to elucidate adverse events that are relatively common in exposed and unexposed populations but occur with increased frequency among individuals exposed to the drug. Epidemiologic cohort studies allow for the assessment of risk factors and the control of potential confounders to a greater extent than spontaneous reports. Large cohorts, not established solely for ADR detection, offer a rich data source of disease risk factors and can add surveillance for ADR at low marginal cost. With such epidemiologic cohort studies, investigators have examined risk factors for breast cancer from postmenopausal hormone use\textsuperscript{32} and identified an association between oral ulcers and the use of nonsteroidal anti-inflammatory drugs.\textsuperscript{33}

When studies use different study populations, different definitions for exposure to the drug, or different definitions for the adverse event, results may vary between no risk and an increased risk of an adverse event with drug exposure. Because individuals in cohorts are not randomized to drug use or no drug use, confounding also may affect the results. The power to detect small increases in risk with drug use will depend on the size of the cohort, with large cohorts needed to reliably demonstrate increased risks of 2- or 3-fold. Despite these potential limitations, cohort studies are an important adjunct to sponta-
neous reporting to determine if adverse events that occur in exposed and unexposed populations happen with increased frequency with drug exposure.

META-ANALYSIS

By evaluating information from multiple sources, including premarketing and postmarketing trials, observational studies, and case reports, the FDA synthesizes available research data to determine if a drug is safe. Meta-analysis, the quantitative analysis of 2 or more independent studies for the purpose of determining an overall effect and of describing reasons for variation in study results, is another potential tool for identifying ADRs and assessing drug safety. Meta-analysis already has been proposed as a method for determining effectiveness of interventions and therapies. In contrast to the published experience of using meta-analysis to evaluate effectiveness, the use of meta-analysis to assess safety remains limited to date. One example of the value of meta-analysis for demonstrating ADRs is the increased mortality associated with the routine use of intravenous lidocaine prophylaxis in patients with acute myocardial infarctions. Though 6 studies individually had too few deaths to conclude that intravenous lidocaine prophylaxis was associated with increased mortality compared with no lidocaine therapy, the summary results demonstrated a significant excess mortality among the lidocaine group.

Suggested roles for meta-analytic techniques include the establishment of associations between drugs and adverse events, estimation of the frequency of ADRs, and identification of subgroups at increased risk for ADRs. Meta-analysis has been used to increase the statistical power for comparing outcomes or assessing outcomes in subgroups. Therefore, it is reasonable to believe that these techniques also may be useful for the evaluation of medication use and adverse events when individual trials are not large enough to demonstrate a clear association or for estimating a relation with increased precision.

CAUSALITY

The development of a symptom or detrimental outcome while taking a medication does not establish the drug as the cause of the injury. Likewise, the development of an event or disease remotely in time from the use of a drug does not exonerate the original therapy from being the source of the problem. Determining which adverse events are caused by drugs with reasonable certainty is an essential, though difficult part of documenting ADRs. Paradoxically, surveillance systems with good adverse event reporting rates increase the probability of receiving spurious ADR reports when the incidence of the complaint in the overall population is not too rare.

Otherwise rare adverse events that occur temporally with the initial use of a drug can be reasonably deduced to be ADRs on the basis of spontaneous reports only. The association between temafloxacin and hemolytic anemia cited above is 1 example. Likewise, adverse events that occur with drug rechallenge also are assumed to be ADRs on the basis of these data and spontaneous reports. The association between vaccination and hair loss is an example. In the case of rare adverse events that occur remotely after drug use or adverse events that are relatively common in the unexposed population, however, spontaneous reports may only be sufficient to raise the concern of a possible association with drug use. In these situations spontaneous reports are a signal that an ADR may exist, and additional studies are needed to sufficiently conclude causality.

For adverse events that are not rare and occur temporally with initial use of a drug, case-control studies have been the most effective method for assigning causality of adverse outcomes to a therapy that are otherwise unpredictable based on known toxicology studies, the structure or function of the medication, or use history of similar agents. Even when the risk for a possible adverse effect is predictable based on nonclinical information—such as mosulactam and a potential increased risk for bleeding—case-control studies have been important in confirming causality. In contrast, national voluntary reporting systems, postmarketing surveillance schemes, and hospital surveillance systems have contributed less in these situations to concluding that the cause of the adverse event was an ADR.

To assess adverse events that occur remotely after drug exposure or that happen in exposed and unexposed populations, case-control studies, cohort studies, clinical trials, linked computer databases, and meta-analyses can be used. Limitations of these methods include power considerations and study design. Cohorts may be too small to reliably detect increased risks of 2- or 3-fold for some exposures. Case-control studies, cohort studies, and meta-analysis may be subject to bias such as exclusion (selection) bias, which may give spurious results. Early case-control studies suggested a relationship between reserpine use and breast cancer that was not confirmed in later studies. One explanation for the conflicting results was that, by excluding individuals with a history of cardiovascular disease from the control group but not the cases, the original study results were influenced by exclusion bias. Case-control studies may be affected by misclassification or recall bias, and the results of cohort studies may be influenced by confounding. In addition to these potential sources of error, meta-analysis results may be affected by unique sources of error such as publication bias. To be helpful in assigning causality of ADRs, these methods must be used appropriately with careful consideration given to potential sources of error. For computerized databases, validation of the data is important for avoiding erroneous results. Results ideally should be confirmed with separate data before concluding causality.

FUTURE NEEDS

Despite important progress in evaluating ADRs, there still is no reliable method for identifying potential ADRs that occur widely separated in time from the original use of a drug, occur with measurable frequency in the unexposed population, and have no predictable relationship to the major effects of the drug. These ADRs are not reliably de-
tected with spontaneous reporting systems such as MedWatch. For example, the identification of clear-cell adenocarcinoma in young women exposed to stilbestrol in utero was aided by the otherwise rare occurrence of the disease in this age group. Primary infertility among women, a much more common adverse outcome from in utero stilbestrol exposure, was not detected until years after the first reports of increased risk for adenocarcinoma led to the creation of stilbestrol-exposed and stilbestrol-unexposed cohorts for follow-up. Because primary infertility is a relatively common problem in young women, occurring in 14% of control women in 1 study, it is possible that the increased risk of 33% among stilbestrol-exposed women would have gone undetected in the absence of previous concerns for additional ADRs raised by the recognition of the risk for vaginal adenocarcinoma.

The identification of ADRs that occur after prolonged administration of drugs for chronic diseases also remains difficult. One suggestion for identifying unintended effects of medications administered long-term is to compare disease or mortality rates with markers for population usage of the drug under concern. Recognition of a rise and fall in asthma death rates in children in the United Kingdom, which coincided with the use of potent nebulizers, or the lack of excess bladder cancer among high users of saccharin-containing products are examples of how disease statistics might be used to identify or rule out possible ADRs. However, as Stolley noted, few databases with the information necessary to conduct these studies are available. Even when such databases exist, the potential for bias or confounding needs to be considered when interpreting any results.

Systematic data exploration in otherwise similar populations of medication users and nonusers of sufficient sample size should be undertaken to look for ADRs. Clinical trials, medical records, and computerized databases are potential sources for data exploration. Meta-analysis provides 1 type of analytic tool for exploring questions not posed in original studies. One proposal is the use of European population databases to look for cases of agranulocytosis, Stevens-Johnson syndrome, and toxic epidermal necrolysis due to new drugs. Though data exploration can be used to evaluate rare adverse events, its most valuable role is likely to be identifying ADRs missed by spontaneous reporting systems. Data exploration may lead to the identification of spurious associations, so potential associations would need to be confirmed or rejected with additional studies of (ideally) other populations. In addition to potential erroneous associations attributable to bias or confounding, many of these databases are likely to be inpatient or outpatient medical records. Their use for pharmacoepidemiology raises important issues about confidentiality. Because it would not be feasible to obtain individual consent to look at each record in a large database, mechanisms need to be in place to ensure the privacy of individual records when conducting these studies.

Associations between adverse events and drug exposure evaluated by using randomized, controlled clinical trials are least likely to be affected by confounding. However, the cost and logistics of trials of sufficient power to confirm ADRs not recognized during premarking studies prohibit use of randomized trials as a realistic option in many cases. Furthermore, the degree of certainty afforded by randomized trials often is not needed to assume causality or risk, and in the case of potentially life-threatening risk, it might be imprudent to wait for confirmatory trials to be conducted. The evidence necessary for the FDA to undertake regulatory action is often less than that derived from a clinical trial and can be fairly limited in some cases.

With many drugs and diseases, where does one begin to look for previously undescribed ADRs? One option might be to begin looking at the most commonly used drugs and the most significant diseases or outcomes. From a public health perspective, it is more important to detect an increase in mortality from pulmonary emboli in oral contraceptive users than it would be to discover the same relative risk for increased mortality from a much less commonly prescribed drug. A second option would be to look for ADRs that might be predicted based on the profile of adverse effects for the medication. A change in the risk for cardiovascular disease with carbamazepine use is an example of an important clinical outcome, a medication commonly given long-term, and a potential mechanism to predict a possible relationship.

To minimize chance associations, relationships detected by data exploration need to be examined with independent analyses before causality is attributed to a medication. Case-control studies, case series, and, where data exist, cohort studies are important for assessing potential associations between adverse events and drugs. Though cohort studies less commonly identify or provide initial confirmation of ADRs, these types of studies can assess both multiple potential ADRs and associations with less potential bias than spontaneous reports and case-control methods. However, cohort studies may still be subject to confounding by indication. When multiple clinical trials contain information on the outcome of interest, meta-analysis is an additional tool for assessing possible associations between adverse events and medication use. Population databases provide another data source for evaluating the potential ADRs. Even with these new methods, though, concern on the part of physicians and patients will remain fundamental to the identification of new ADRs.

Funding/Support: Dr Brewer is supported by Clinical Investigator Award 1 KOB A101444-01A1 from the National Institute of Allergy and Infectious Diseases.

REFERENCES