Postmarketing Surveillance Of Adverse Drug Reactions: Patient Self-Monitoring

Seymour Fisher, Ph.D., and Stephen G. Bryant, Pharm.D.

Abstract: Background: This report summarizes our experience with a new approach to postmarketing drug surveillance using a pharmacy-based patient self-monitoring strategy, developed in collaboration with Eckerd Drug Company, the American Association of Retired Persons Pharmacy Service, and other pharmacies nationwide.

Methods: Patients presenting prescriptions to collaborating pharmacies for a targeted drug or a standard drug used as a control received an entry form asking them to register and then call a toll-free telephone number to report possible drug reactions. When contacted by patients, study staff conducted a standardized, computer-driven interview. Two brief mail questionnaires were also employed.

Results: Original validation data gathered from 1984 through 1986 indicated that the most commonly expected adverse drug reactions caused by antibiotic and tricyclic antidepressants reported by 162 self-monitoring patients closely matched those elicited from a comparable control sample of 1109 patients who were independently interviewed by our staff. Results from subsequent studies are also described.

Conclusions: We believe this method has great promise for providing not only a cost-effective, complementary, early alerting mechanism for detecting adverse drug reactions, but also the additional possibility for discovering unsuspected therapeutic benefits of newly marketed drugs. (J Am Board Fam Pract 1992; 5:17-25.)

This report describes the evolution of a new pharmacy-based approach to postmarketing drug surveillance and serves to alert physicians to the possibility that their patients may be participating in a nationwide test of the method. The strategy, which relies on patient self-monitoring, was conceived to complement existing record-linkage and voluntary physician-reporting systems, although the method also has the unique capability of detecting possible unsuspected therapeutic indications, as well as potential adverse drug reactions.

Validation studies on selected oral antibiotics and tricyclic antidepressants were first conducted from late 1984 through 1986 at the University of Texas Medical Branch pharmacies, and in 1987 a statewide test was initiated using selected Texas community pharmacies. More recently, we have extended the method nationwide in collaboration with the Eckerd Drug Company and the American Association of Retired Persons (AARP) Pharmacy Service.

Existing Approaches
The US Food and Drug Administration (FDA) relies primarily on a combination of mandatory and voluntary reporting mechanisms for postmarketing surveillance of adverse drug reactions (ADRs). Pharmaceutical manufacturers are required by law to submit to the FDA reports of all suspected domestic ADRs. In addition, the FDA encourages a spontaneous reporting system from physicians. Most of the possible ADR reports originate directly from pharmaceutical manufacturers (90 percent), while the remaining are received from physicians, other health-care providers, and consumers.

A 1980 report by the Joint Commission on Prescription Drug Use concluded that sole reliance on systems of voluntary reporting by physicians to manufacturers, journals, and government agencies was clearly unsatisfactory. The FDA spontaneous reporting system was found by Rossi...
and Knapp in 1984 to be somewhat inferior to physician case reports and letters to the editor published in the medical literature. Although the FDA spontaneous reporting system has been substantially improved in recent years, Strom and Tugwell noted that even in 1990 there remained major liabilities inherent in the FDA first-line system. Clearly, voluntary physician reporting and publishing contribute valuably to a general alerting mechanism, but there is widespread agreement that as a comprehensive system, it suffers from serious flaws: underdetection of low-incidence ADRs, overinclusion of false-positives, and low-reliability incidence estimates for the true-positives. Others have emphasized also that such flaws are especially true for ambulatory patients, many of whom cannot be monitored adequately after initiating treatment. Furthermore, a 1990 external evaluation of the FDA drug review process by the US General Accounting Office concluded that “the serious postapproval risks identified in studying their frequency and seriousness involved a wide variety of adverse reactions.”

In addition to the voluntary reporting system, a number of different approaches to formal systems have been explored. Most depend upon physician judgments in one form or another, including record-linkage methods, such as utilization of Medicaid drug-event data or data relating pharmacy prescriptions to hospital record information. In one evaluation of selected formal monitoring systems, manufacturer-initiated phase IV studies failed to identify a number of significant new ADRs from two drugs later requiring labeling changes. Moreover, while record-linkage analyses are gaining in popularity, Shapiro recently ignited sharp controversy by criticizing the validity of conclusions drawn from automated record-linkage data: e.g., he asserted that most analyses lack accurate information on duration and timing of the outcome, as well as the onset, of target drug exposure, causing uncertainty whether target drug exposure antedated the outcome—information that “can be obtained only directly from the patient”; he also noted that “with rare exceptions, adequate information on confounding can be obtained only from the patients.” Immediately following publication of this paper, several rebuttal commentaries and letters to the editor were offered along with a concluding response from Shapiro.

While the primary concern of the FDA must necessarily focus on the detection of very serious ADRs, this narrow view has discouraged pharmacoepidemiologists from taking any systematic interest in the possible importance of less serious ADRs that can be harbingers of subsequent serious drug toxicity. Some years ago, when a paper by an FDA representative emphasized the need to monitor mainly serious ADRs, a researcher wondered whether more careful attention to early reports of peripheral neuropathy from thalidomide patients might have helped to avert the teratogenicity disaster.

In 1984 Fisher, et al. developed the self-monitoring method under a grant from the National Institute of Mental Health, first reported in 1986. It is noteworthy that, although most suspected ADRs (along with some new benefits) in nonhospitalized populations stem initially from patients’ spontaneous reports to their physicians or to associated medical staff, there had been no previous attempts to develop any formal patient-initiated surveillance approaches. Recently, however, Campbell and Howe successfully used a somewhat similar pharmacy-based, patient-initiated approach in Edinburgh to increase the level of ADR reports directly to physicians.

**Preliminary Studies**

As the system has evolved, patients filling a prescription for a target drug or a comparable standard drug used as a control are now presented with or mailed an announcement of the study along with their medication. The enclosure has a face page indicating the joint collaboration between the particular pharmacy chain or mail-order service and the University of Texas School of Medicine at Galveston. The inside pages provide the details of the study and include a postage-paid entry form.

It has been commonplace unfortunately for ADR-oriented interviews in clinical practice to range from a casual inquiry about how the patient has been feeling since starting treatment to an elaborate review of body or organ systems without ever mentioning any target drugs being monitored. All our subjects have been aware of the drug being monitored, and all our telephone interviews have followed a precise, standardized,
The interview is computer-driven so that the questions appearing on the interviewer's video screen are almost completely determined by the patient's prior response. Another advantage of the computerized interview (simultaneously audio-tape recorded) is that the patient's responses are entered directly into the database at the time of the interview (and later checked for accuracy and completeness), thereby substantially reducing the risk of human error in data transcription. A typical interview with a first-time caller (patient or caregiver) has averaged 16 minutes for patients younger than 60 years old, 19 minutes for patients aged 60 to 69 years, and 18 minutes for patients older than 69 years; repeat calls have been generally shorter because some basic information was obtained earlier. Details of the telephone interview have been previously reported.41

Validation data for the original method in 1984 were obtained from a carefully controlled study using the University of Texas Medical Branch outpatient pharmacy. The major results43 indicated that the most commonly expected ADRs caused by antibiotic and tricyclic antidepressants reported by 162 self-monitoring patients closely matched those elicited from a comparable control sample of 1109 patients who were independently interviewed by our staff. Table 1 shows that an analysis of adverse clinical events reported during patient-initiated interviews with 118 patients taking tricyclic antidepressants and 252 patients taking antibiotics could detect many of the known, common ADRs caused by tricyclic antidepressants and some of the common expected gastrointestinal ADRs† caused by antibiotics, even when

---

Figure 1. Sample inside pages of an announcement distributed by pharmacists to appropriate patients.
Table 1. Percentage of Known Possible Adverse Drug Reactions Based on Total Number of Spontaneously Reported Adverse Clinical Events from 162 Self-Monitoring Patients (50 receiving tricyclic antidepressants and 112 receiving antibiotics).

<table>
<thead>
<tr>
<th>Drug Groups</th>
<th>Tricyclic n = 118*</th>
<th>Antibiotic n = 252*</th>
<th>χ²†</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Anticholinergic</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dry mouth (TCA)</td>
<td>16.95</td>
<td>2.78</td>
<td>23.9</td>
</tr>
<tr>
<td>Blurred vision (TCA)</td>
<td>8.47</td>
<td>1.19</td>
<td>12.6</td>
</tr>
<tr>
<td>Constipation (TCA)</td>
<td>5.93</td>
<td>1.59</td>
<td>5.3</td>
</tr>
<tr>
<td><strong>Neurologic</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sedation (TCA)</td>
<td>11.87</td>
<td>2.38</td>
<td>14.1</td>
</tr>
<tr>
<td>Shakiness or tremor (TCA)</td>
<td>3.39</td>
<td>1.59</td>
<td>1.2</td>
</tr>
<tr>
<td>Head pain or ache (TCA)</td>
<td>4.24</td>
<td>4.76</td>
<td>0.1</td>
</tr>
<tr>
<td><strong>Gastrointestinal</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Diarrhea (TCA)</td>
<td>9.32</td>
<td>7.94</td>
<td>0.2</td>
</tr>
<tr>
<td>Fainting (TCA)</td>
<td>0.00</td>
<td>0.40</td>
<td>0.5</td>
</tr>
<tr>
<td><strong>Oral</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sore tongue (AB)</td>
<td>0.00</td>
<td>0.40</td>
<td>0.5</td>
</tr>
<tr>
<td><strong>Skin</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Itching (AB)</td>
<td>4.24</td>
<td>5.95</td>
<td>0.5</td>
</tr>
<tr>
<td>Rash (AB)</td>
<td>0.85</td>
<td>2.38</td>
<td>1.0</td>
</tr>
<tr>
<td>Hives (AB)</td>
<td>0.85</td>
<td>4.76</td>
<td>3.6</td>
</tr>
</tbody>
</table>

TCA = common tricyclic ADRs; AB = common antibiotic ADRs.

*Because the appropriate total number of participating patients was unknown, the total number of reported adverse clinical events was used as the denominator for calculating percentages. These values should not be interpreted as incidence estimates.

†χ² probabilities: 10.83 = 0.001, 6.63 = 0.01, 3.84 = 0.05, 2.71 = 0.10.

one drug group was compared with another group receiving a very different class of drugs. Chi-square was used as an indicator of the extent to which the two percentages differed (each 2 x 2 chi-square table consisted of the two drug groups and the reported presence or absence of a particular adverse clinical event or set of adverse clinical events). Note also that the percentage comparisons for most of the adverse clinical events were in the expected direction even when the chi-square estimates were small.

In one analysis, we were able to compare one drug with other drugs prescribed for relatively similar indications, we found a previously unreported increased rate of doxycycline-related gastrointestinal distress relative to other oral antibiotics: 6.4 percent of patients taking doxycycline reported nausea or vomiting compared with 0.7 percent of patients receiving tetracycline, 0.4 percent of patients on penicillin therapy, and 1.5 percent of patients taking ampicillin. Whether this marked increase was truly attributable to doxycycline alone or to other factors (e.g., different types of patients for whom doxycycline was prescribed), the salient point was that the number of spontaneously elicited adverse clinical events from a self-monitoring sample of only 78 patients was sufficient to detect the same relation as that seen in a validation sample of 457 control patients surveyed with staff-initiated telephone interviews.

In another analysis, with patients aged 16 to 83 years, the expected positive relation between age and potential ADRs was detected only with patient-initiated spontaneous reporting of adverse clinical events, whereas the more standard staff-initiated survey approach data misleadingly suggested fewer ADRs in older patients than in younger patients. Further, we have recently reported that patients can correctly attribute probable ADRs to their medication and that older patients appear to be capable of discriminating
probable ADRs as well as or possibly better than younger patients.\(^{36,47}\) Attribution accuracy, however, irrespective of age was generally poorer in self-monitoring patients than in patients who were independently surveyed by staff—probably because the patient-initiated method generated a tendency to report mainly those adverse clinical events suspected of being drug-related. While this tendency leads to many incorrect attribution judgments offered by patients and further implies that some important ADRs can go unreported, it has the virtue of eliminating many statistically unwanted reports (false-positives or “noise”) from the database.

During the last 4 years, various pilot studies had to be conducted to modify different relevant factors before we had adequate evidence of a relatively efficient system that could be used on a nationwide basis. By extensive trial and error, we were able to solve numerous problems associated with (1) determining the optimal procedures to increase the motivation of participating pharmacists and patients, (2) developing and refining a useful computer-assisted interview and adjacent mail questionnaires, (3) developing a comprehensive coding system to allow appropriate description and quantification of the reported clinical events, (4) generating the explicit guidelines for our technicians to use in recording valid information obtained from the patient-initiated interviews, and (5) finding the most effective procedures for tracking patients who had entered a study but had not made contact.

In March 1991 we completed a study comparing trazodone with desipramine in nonhospitalized depressed patients using AARP Pharmacy Service customers who were offered modest compensation for their participation. Approximately 35 percent of those receiving the announcement volunteered, and about 15 percent of all valid patients—including the elderly aged more than 80 years, either personally or with the aid of a caregiver—reported one or more valid health changes during the period of self-monitoring (4 percent of the calls came from caregivers). Thirty-eight percent of the 1897 valid patients were at least 70 years old (mean and median were 66 years); and although we had no patients older than 90 years, 7 percent were more than 80 years old (2 percent were older than 85 years). Reports from 389 patient-initiated interviews showed that patients receiving trazodone had significantly fewer anticholinergic side effects than patients taking desipramine, and much of trazodone’s increased sedative action was reported by patients as “sleeping better.” Of special interest is that many of the trazodone-desipramine differences were beginning to emerge in the first 50 interviews.

**Data Analyses**

A valid adverse clinical event or beneficial clinical event was defined as any “new or unusual” adverse or beneficial clinical event that unequivocally began only after the patient started the designated medication (whereas not all valid adverse clinical events were true ADRs, all ADRs must obviously first be valid adverse clinical events; similarly, all beneficial drug reactions must first be valid beneficial clinical events).

As each telephone interview was completed, our computer continuously scanned each reported adverse clinical event or beneficial clinical event, comparing its relative frequency in the two drug groups; the system was highly efficient in detecting an alert at any stage of an ongoing surveillance. Whenever a difference was detected between the two drug groups, the computer identified that adverse clinical event by showing the two incidence estimates in monthly printouts. The threshold level for defining an incidence “difference” depended, in part, on the nature of the adverse clinical event: to ensure that we were not missing any possible difference, for most adverse clinical events we used a liberal chi-square of 1.65, where a \( P < 0.20 \) alerted us to follow the adverse clinical event.

For some key events (e.g., death, hospitalization), all individual adverse clinical events were cumulatively listed along with the two incidence estimates and the corresponding chi-square value. Another subset of adverse clinical events was designated as “serious”: e.g., heart or blood pressure changes; yellow skin or eyes; unconsciousness or fainting; seizure; blood in stools, sputum, or urine; right- or left-sided weakness; numbness; difficulty breathing; and confusion, disorientation, or delirium. We took particular notice when the patient had never previously experienced one of these adverse clinical events or if the event was
described as being distinctly different in quality or quantity from prior episodes.

**Early Detection of Possible ADRs**

When an adverse clinical event or set of adverse clinical events was first identified by our computer, the difference between the two incidence estimates could be relatively unreliable. As more reports of the adverse clinical event were accumulated, the magnitude of the sample difference usually changed. Some adverse clinical events appropriately fell by the wayside: if the significance level of an adverse clinical event rose above 0.20, the adverse clinical event no longer appeared on the printout. Where true differences existed, even if the early incidence estimates remained stable, the increased sample size would lead to lower probability values for the chi-square test, with concomitantly increased confidence in having identified a possible ADR.

A significantly greater relative frequency of a symptom in the targeted drug group not only served an important alerting function but also conveyed some probability of target-drug attribution. Limiting analyses to those events for which relative frequencies differed at some predefined level reduced the number of false-positives for one of the drugs (i.e., adverse clinical events that could have generally high “spontaneous” base rates, as seen in both drug groups). Whenever a difference was detected, we then analyzed all current illnesses and other medications to rule out possible confounding influences. Although the use of a comparison drug still did not assure that the observed incidence differences must be attributed to drug differences, it did help markedly to eliminate the influence of numerous other confounding factors that might have misleadingly suggested either new ADRs or increased incidence of expected ADRs. Depending on how comparable the standard and new drug patients were known or assumed to be, disconfirmation was suggested when unusual or high-frequency reactions reported by the target drug group were observed with equal frequency in the standard drug group (the reactions could, of course, be true ADRs common to both drugs).

For example, one of the most frequent adverse clinical events reported by self-monitoring patients receiving doxycycline therapy was dizziness or lightheadedness (Table 1), which could be the portent of a potentially serious ADR. Patients taking other antibiotics (including tetracycline), however, reported this adverse clinical event with the same relative frequency, leading to the correct inference that dizziness or lightheadedness was not specifically a doxycycline-induced ADR (consistent with the official labeling).

**Detecting Possible New Therapeutic Uses**

There are currently no systematic methods being used to discover new therapeutic indications after a drug has been marketed. Occasionally, even though premarketing clinical trials had only been conducted for one indication, the drug's known pharmacologic properties can suggest another indication for the manufacturer to explore at a later time (e.g., imipramine for enuresis). In the majority of cases, however, finding clues for new indications has relied on serendipity—a concept that we have recognized as being a necessary feature of all science, but one to be avoided as a reliable means of discovery. Strom, et al.48,49 presented a persuasive argument in favor of more intensive and systematic monitoring of drug benefits, and similar recommendations have been made by others.50,51 Surely there is a better way to discover the multiple therapeutic uses for tomorrow's iproniazid, minoxidil, tretinoin—and aspirin—without waiting for clinicians and pharmaceutical companies to run across a suggestive bit of data inadvertently.

One major reason for this methodological hole in drug development has been the nature of the FDA new drug application process. Phase I and II premarketing studies usually have had overall small total samples. To demonstrate drug causality convincingly, premarketing controlled clinical trials (phase III) in support of a new drug application must be highly artificial, uncorrupted by the ordinary conditions of general practice; it also has been relatively common for the total number of patients studied to be fewer than 3000, and it should be fully appreciated that, except for specifically designated nontropic drugs, most clinical trials have excluded patients older than 65 years! Because patients with concurrent illnesses other than the one(s) being investigated for labeling have been excluded, unsuspected benefits generally cannot be seen until the drug has been approved for marketing. Even then, however, as only a subset (probably rather small) of patients
being treated with the new drug will also happen to have another specific condition for which the drug may be beneficial, the unsuspected relation may well go unnoticed or unreported in the medical literature.

Because we began to collect data on beneficial clinical events only recently, we still lack validity. Nevertheless, these data have indicated that patients are ready to report what they believe to be unexpected health benefits: from 389 patients reporting adverse clinical events or beneficial clinical events, 22 universal benefits and 56 personal benefits were described. Many of these reports were obviously related consequences of the drug's primary therapeutic effects (e.g., patients receiving desipramine or trazodone reported "more energy," "increased strength," "alert when awakened") and were of no special clinical significance. Other reports, however, have suggested that the self-monitoring system has potential for discovery (e.g., alleviation of irritable bowel syndrome, reduction of seizure frequency, asthma relief). Clearly, many more reports of unexpected benefits will be required before we can tell whether patient self-monitoring can contribute to the detection of possible new therapeutic uses for a given target drug.

Advantages and Disadvantages
We have emphasized that this patient-initiated self-monitoring approach must be viewed as complementary to other postmarketing surveillance approaches because of the following limitations:

1. The system focuses mainly on reports of symptoms (plus occasional signs) from non-hospitalized patients and would only rarely yield true diagnoses or final objective outcomes. As a corollary, long-delayed insidious ADRs (e.g., interstitial nephritis, endometrial cancer) may never be detected within the system, nor will any ADRs that are only detectable by a clinical laboratory procedure (e.g., low sperm count, leukocytosis).

2. Drug-induced hospitalizations and deaths are unlikely to be detected as efficiently by this system as by other existing postmarketing surveillance methods.

3. Because self-monitoring patients on average tend to report symptoms that are perceived as relatively severe possible ADRs, some genuine ADRs with serious implications could go unreported because they were perceived to be too mild (e.g., bruising easily) or are not recognized as being possibly drug related.

Nevertheless, despite these limitations, our research to date suggests that this patient self-monitoring strategy could be a valuable addition to presently available approaches. The following are some of the major advantages:

1. The patient-initiated method is simple, efficient, and possibly quite cost-effective. For every 1000 valid patients recruited, only a small percentage (to date, approximately 20 percent) require a full interview. Under the right conditions, it takes about 12 months to monitor 3000 to 5000 patients on a new drug and approximately the same number on a control drug at a cost of less than $50 per patient.

2. By always using a standard drug comparison group along with each targeted new drug group (a methodological advance first advocated by Inman17), partial ADR confirmation or disconfirmation is possible, as opposed to the mere reporting of symptoms associated with target drug use. Additionally, the data can also suggest therapeutic advantages when the incidence estimate for one or more known or expected ADRs reported by the target drug group is significantly lower than that observed in the standard drug group.

3. By requesting patients to monitor unexpected health benefits, the method offers a systematic approach toward discovering new therapeutic uses for the target drug or even the control drug.

4. The method yields more accurate—albeit still attenuated—incidence estimates by multiple prompts for identifying new events (numerator) and by providing a fairly exact count of drug exposures (denominator).

Summary
It is clear that a comprehensive postmarketing surveillance system for monitoring pharmacotherapy will ultimately demand the development and application of multiple monitoring methods. Patient self-monitoring has been sadly neglected
at a time when innovative approaches are needed. We believe the concept of patient-initiated surveillance has great promise for providing not only a cost-effective, complementary early alerting mechanism for detecting ADRs, but also the additional possibility of discovering unsuspected therapeutic benefits for newly marketed drugs.

We are indebted to Charles R. Sullivan, R.Ph., Director of Pharmacy Development, Eckerd Drug Company, and Nancy J. Olins, M.A., Director of Program Development, AARP Pharmacy Service, for their invaluable collaboration.

References