1
Summary

1. Summary

INTRODUCTION

To market a drug, the manufacturer must provide evidence of its efficacy and safety to the U.S. Food and Drug Administration (FDA). Once these premarketing requirements are met and the drug has been released, FDA can remove a drug from the market—after giving due notice and an opportunity for a hearing—because of new evidence on the drug's efficacy and safety, the discovery that the drug was approved on the basis of any untrue statement of a material fact, or the failure of the drug to meet manufacturing standards. In cases where a drug may be an "imminent hazard to the public health, " FDA can suspend the drug's approval immediately, giving prompt notice of the action and offering the opportunity for an expedited hearing.

In premarketing testing, the numbers and types of patients used to demonstrate a drug's efficacy and safety are limited compared with the numbers and types of patients who will eventually be prescribed the drug after it is marketed. The initial decision to approve a drug for use, however, must be made on the basis of the available knowledge.

Although postmarketing surveillance cannot provide knowledge of the safety or efficacy of drugs at the time of their introduction on the market, various kinds of postmarketing surveillance have been proposed over the past decade to monitor and aid in modifying the use of drugs. The principal focus of postmarketing surveillance proposals has been on the safe use of prescription drugs, even though the range of issues has encompassed both efficacy and safety considerations, e.g., concern over refinements in use as well as better definition of drug risks.

Current interest in prescription drug evaluation and monitoring is focused on the premarketing approval process and the length of time it takes for a drug to be approved by FDA; postmarketing surveillance appears to have waned as a policy issue. Thus, policy formulation and implementation for the premarketing approval process is being pursued without parallel efforts for the postmarketing period.

However, postmarketing surveillance deserves attention as a policy issue for both short- and long-term objectives. Regarding short-term action, if current testing requirements for the premarketing approval process are reduced, pharmaceutical manufacturers could be required to maintain their drug evaluation responsibilities by increasing postmarketing surveillance. Regarding long-term action, postmarketing surveillance remains a policy issue irrespective of current interest in the premarketing approval process: it is only after marketing that a drug's full therapeutic and harmful potentials can be determined.

One way to shorten the premarketing period of the drug approval process would be by reinterpreting the regulations for assessing safety and efficacy. This report provides theoretical and experiential criteria for evaluating how such changes may affect the ability of current guidelines to detect a drug's harmful and beneficial effects. It also discusses the kinds of qualitative changes in the evidence required for drug approval that FDA is implementing. Finally, the report identifies options relating to FDA's postmarketing surveillance. These options could be implemented regardless of whether there is a change in current premarketing drug approval requirements.

THE DRUG APPROVAL PROCESS

A drug's sponsor must provide: 1) "adequate tests by all methods reasonably applicable to show whether or not such drug is safe for use under the conditions prescribed, recommended, or suggested;" and 2) "substantial evidence that the drug will have the effect it purports or is represented to have" (2 I U. S. C., sec. 355 (d)). This statutory language has led in practice to FDA's establishing a premarketing phase of drug testing that consists of two parts: 1) the investigational new drug (IND) application process, and 2) the filing of a new drug application (NDA).

The IND application describes the investigators' qualifications and the planned clinical trials, the chemical composition of the drug, and data on the pharmacology and toxicology of the new drug collected in animal studies and in prior human studies, if any, such as those conducted in other countries.

The clinical investigations in the IND process are divided into three phases (24):

- *Phase I: Clinical Pharmacology* is that phase in which a drug is first used on humans to confirm dose ranges and pharmacologic effect. The number of subjects in phase I varies depending on the drug, but is usually in the range of 20 to 80 (excluding control patients). Pharmacodynamic and metabolic studies, in whichever stage of investigation they are performed, are considered to be phase I clinical pharmacologic studies.
- *Phase II: Clinical Investigation* consists of controlled clinical trials to demonstrate a drug's effectiveness and relative safety. These are performed on closely monitored patients of limited number, usually 100 to 200 patients, with equal numbers of control patients.
- *Phase 111: Clinical Trials* are expanded controlled and uncontrolled trials to gather additional evidence of a drug's effectiveness for specific indications and to more precisely define its adverse effects. Phase 111 studies observe a total of 500 to 3,000 patients in more natural settings—in clinics, outpatient hospital facilities, and private practice. Phase 111 usually consists of more than two controlled trials.

After completion of the testing required under the IND application, the sponsor may file an NDA. At least two well-controlled studies establishing each indication for which the drug is intended are required. More than one indication can be established in a single study. (These requirements are under review; see chs. 3 and 6.)

All INDs are classified by chemical type and therapeutic potential, so that those drugs considered by FDA to be of particular therapeutic importance can receive priority review. The highest classification is given to drugs that are new molecular entities (type 1) and that may represent important therapeutic gains (type A)—type 1A drugs.

Several mechanisms are available to FDA to obtain information about drugs once they have been approved for marketing. Once the NDA has been approved, the sponsor is required to monitor information and submit reports about the drug. Other information on adverse drug reactions (ADRs) is monitored by FDA in a number of ways:

- the Spontaneous Reaction Reporting Program, in which information on ADRs is sent to FDA by physicians, pharmacists, and hospitals;
- a monthly review of the medical literature on ADRs (reports and letters to the editors of medical journals, etc.);
- intensive surveillance and epidemiologic studies of ADRs in selected hospitalized and ambulatory populations;
- several specialized registries that collect and analyze possible ADRs;
- in-house monitoring and research studies of such data bases as those of the Medicaid Medical Information Systems of some States and those of commercial sources of drug use data; and
- the World Health Organization, which exchanges reports with FDA, each summarizing the ADRs added to their systems in the previous year.

This postmarketing information is useful for two purposes. First, it may provide the grounds for FDA to remove a drug from the market, when such action is appropriate. Second, it is used by FDA to ensure that limits are placed on advertising and promotional claims and that the drug's labeling is appropriate.

FDA may request further studies when there are questions about a drug that were not sufficiently answered by the phase 111 studies, but which dc~ not warrant delaying the release of what promises to be a useful new product (24). Although FDA has no explicit authority to require such studies, these "phase IV" studies are almost always performed, as the alternative would be nonapproval of the drug.

HISTORY AND OBJECTIVES OF POSTMARKETING SURVEILLANCE

As a result of 1974 hearings before the Senate Committee on Labor and Human Resources' Subcommittee on Health, the Department of Health, Education, and Welfare formed a Review Panel on New Drug Regulation. The panel issued its report in May 1977 (16).

A bill was subsequently introduced in the Senate in early 1978 to revise the drug provisions of the Food, Drug, and Cosmetic Act. A revised bill, S. 1075, the Drug Regulation Reform Act of 1979, passed the Senate in September 1979. A similar bill, H.R. 4258, was not acted on by the House of Representatives. Included in the Senate bill were the following specifications: 1) drug sponsors could be required to conduct postmarketing surveillance of a drug for up to 5 years; 2) a prescription drug could have its distribution limited if the drug could not otherwise be found to be safe and effective; 3) the standard for a drug's immediate removal from the market would be changed from the drug being an "imminent hazard to the public health" to the less stringent standard of "unreasonable risk of illness or injury to any segment of the population;" and 4) establishment of a "National Center for Drug Science. "

During this period, in a speech to the Pharmaceutical Manufacturers Association, Senator Edward Kenned$_{y}$ (D-Mass.) suggested that a better system was needed for monitoring the use and effects of prescription drugs after they were marketed. As a result, the Joint Commission on Prescription Drug Use was established in 1976, funded largely by the drug industry, with the mandate to design a postmarketing surveillance system to detect, quantify, and describe the anticipated and unanticipated effects of marketed drugs, and to recommend a means by which information on the epidemiology of prescription drug use in the United States could be distributed regularly to interested parties. The Joint Commission issued its report in January 1980 (42), but by this time, interest in postmarketing surveillance had waned, and the commission's report and recommendations were little noticed.

In 1976, the year in which the Joint Commission was formed, an interagency agreement was signed between FDA and the Experimental Technology Incentives Program (ETIP) of the National Bureau of Standards in the Department of Commerce. The purpose of ETIP was to provide incentives or reduce barriers to technological innovation through changes in the regulatory process. ETIP's agreement with FDA was to jointly fund a program to determine if improvement in postmarketing surveillance could help reduce the regulatory requirements of the premarketin~ period, principally those of phase III of the IND process and those of the NDA process. The specific experiment was to develop postmarketing surveillance systems and a method of managing and evaluating the reform (11). The project concentrated on collecting the information required to design these systems (12). By 1982, FDA had assumed most of the funding, as ETIP was to be phased out that year.

A Commission on the Federal Drug Approval Process was convened in mid-1981 to examine how FDA's procedures for the approval of new drugs could be expedited without compromising

public safety and to make recommendations on the development of cost-effective postmarketing surveillance to guarantee the quick withdrawal from the market of drugs that cause significant adverse effects. The commission had its genesis in a joint hearing held in April 1981 by the House Science and Technology Committee's Subcommittee on Natural Resources, Agriculture Research, and Environment and its Subcommittee on Investigations and Oversight. The first meeting was held in July 1981. The commission completed its work and announced its general findings in the spring of 1982, and its printed report was to be released in late 1982.

FDA is examining specific ways to speed up the drug approval process. It is reviewing past phase 111 trials to see if longer trials or those with large samples have contributed useful information beyond that obtained in phase 11 and early phase III testing. Past postmarketing studies that FDA required are also being reviewed to see if they provided the information that they were designed to obtain. Data on FDA approval time are being reviewed to see what other factors may slow the approval process. And, as a pilot test, an FDA committee is reviewing the pharmacologic and clinical data on selected drugs at the end of phase 11 testing, and will make recommendations about the best time for gathering additional information (e.g., phase III v. the postmarketing period) (11).

In March 1982, the FDA Commissioner began a related organization by merging the Bureau of Drugs with the Bureau of Biologics, and replacing the Director of the New Drug Evaluation Division. The merged bureaus have since been designated the National Center for Drugs and Biologics.

Finally, in a related development, the Senate passed by a voice vote, in the first session of the 97th Congress, the Patent Term Restoration Act of 1981 (S. 255). The bill would restore to the term of a patent the time lost in complying with the Government's premarketing testing and review requirements, up to a maximum of 7 years, Patented products eligible for extension would not be limited to human drugs, but would include "human drugs and biological, antibiotic drugs, animal drugs and biological, food additives, color additives, pesticides, other chemical substances, medical devices, and any other product subject to Federal premarket requirements" (72). In September 1982, the House of Representatives voted on the bill under suspension of its rules. Under such conditions, a two-thirds vote was required for passage, and although the bill received a majority of the votes, it fell just short of the two-thirds majority needed.

METHODS OF SURVEILLANCE

The primary objective of postmarketing studies is to develop information about drug effects under customary conditions of drug use. Initial clues about a drug's potential effects come from the experimental studies carried out with both animals and humans in the premarketing period. Spontaneous or voluntary reporting (e. g., in letters to the editors of medical journals) is the oldest, and to date, the most productive source of new information about a drug's possible effects once a drug is marketed. Other types of studies are used to examine in more detail the possible effects of a drug. In general, these other types of studies use either cohort or case-control methods.

Thus, four types of studies are generally used to identify drug effects: 1) controlled clinical trials, 2) spontaneous or voluntary reporting, 3) cohort studies, and 4) case-control studies (23,50,61,77).

Controlled clinical trials match treatment and control groups as closely as possible, minimize bias through such methods as randomization and "double-blinding," and directly monitor patients for the duration of the study. Controlled clinical trials are considered the most definitive method for evaluating a drug's efficacy and safety, but they are often costly or impractical in specific situations, for example, when a drug's effects are

rare, or appear only after long-term use or a long latency period.

Voluntary reporting by physicians and other health providers, hospitals, and consumers may act to alert FDA and pharmaceutical firms to possible adverse effects of drugs, so that the inference of an association between a drug and an observed health condition may be further studied by cumulative, careful reporting, and confirmed or disconfirmed by more vigorous methods. Underreporting may be a serious deficiency of this method. A drug may also be erroneously associated with an adverse effect until the suspected association fails to show up in repeated, statistically validated studies.

Cohort studies follow a defined group of patients (the cohort) for a period of time. In this method, patients are not randomly assigned to groups, and there is no blinding. Cohort studies are usually prospective and observe the cohort from the beginning of drug use. A group of patients taking the drug of interest is assembled and followed to see, for example, if adverse reactions occur. A second group of patients (the controls) with the same medical condition, who are not taking the drug and who may be receiving alternative treatment, but who are otherwise matched as closely as possible with the cohort, may also be studied in parallel. The control group is used to identify the frequency of occurrence of any condition observed in the drug-exposed group which is due to causes other than the drug (i. e., the "background incidence" of the condition). In this method, patients can be directly monitored to ensure they take the drug appropriately, and to observe the drug's effects; *or* monitoring can be less controlled. With less control, a larger cohort can be followed, but bias is thus increased.

Case-control studies identify patients with the adverse effects to be studied (the cases), and compare them with a sample (the controls), drawn from the same cohort that gave rise to the cases. Controls are matched as closely as possible with the cases, except with regard to the drug's suspected adverse effect, to examine whether exposure to the drug is the cause. Patients with conditions suspected of being associated with a certain drug would have their medical records reviewed or be interviewed concerning the use of that drug. The histories of the controls would also be studied for information about drug use in the general population. By comparing the proportion of drug users among the cases with the proportion of drug users in the general population, it is possible to infer the relative frequency with which adverse reactions occur in users of certain drugs as compared with nonusers. A sufficient number of appropriate cases must be identified and accurate histories of exposure to drugs must be obtained.

Controlled clinical trials and prospective cohort studies can be used to determine a drug's beneficial as well as adverse effects. Case-control studies are usually used to trace adverse effects back to prior drug use. Voluntary reporting can uncover additional uses of drugs as well as their adverse effects, but reporting of adverse effects is much more common.

The ability of a particular surveillance method to detect a drug's effect depends on two factors: 1) the time that transpires between use of that drug and the occurrence of the drug's effect (the latency period), and 2) how often the effect occurs (its frequency). There are many other determining factors, such as accuracy of observation, and accuracy and completeness of medical records, but these factors present more of a problem in the design of a study's details.

Controlled clinical trials, because of their relatively short duration, will detect only acute or subacute effects. Long-term cohort studies can detect delayed effects, but the data bases necessary for such long-term, large studies are still sparse. Voluntary reporting is usually the way in which long-term effects are first identified. Long-term effects are usually confirmed through retrospective case-control studies, but such studies' reliance on historical data such as medical records can limit their accuracy.

The chance that a particular study will discover a drug effect also depends on the study's sample size and the frequency of the drug effect, For example, in a cohort study, if a drug causes blindness in 1 out of every 100 users (1/100), how many users must be observed to find one case of blind-

98-828 O - 82 - 3

ness? If there are 1 million users of the drug, there would be 10,000 users blinded. But in a small sample of only 100 users, the probability of finding one or more cases of blindness would only be 63 percent. If the sample were 200 users, the probability of finding one or more cases would increase to 86 percent. With a sample of 500, the probability would be 99 percent that at least one case of blindness would be found in the observed users.

To state it another way, what number of users would have to be observed to be 95 percent sure of finding one or more cases of blindness when they occur at a frequency of 1 in 100 users? The answer is 300 users, and the general rule is that the number of users in the sample must be three times the reciprocal of the frequency; e.g., for a frequency of 1 in 1,000, the sample would have to be 3,000 to be 95 percent sure of observing at least one case.

Except for some effects that are unique to a specific drug, many drug effects (e. g., stroke, bleeding, skin rashes) are indistinguishable from conditions due to other causes. The "background incidence" of a condition must be known before purported drug effects observed in a study can rightly be attributed to a drug.

Larger sample sizes are needed to determine a drug's effect as the background incidence of a condition increases and as the frequency of a drug's contribution to a condition decreases. For example, given a background incidence of 1/100, as the incidence of a drug's added effect decreases from 1/100 to 1/10,000, the sample size would have to increase from 1,600 to 11 million to remain 95 percent sure of observing at least one case of the added effect. The relationship between background and added incidence is also revealed in considering sample sizes at the extremes. For a known background incidence of 1/1,000 and an added incidence of 1/100, the sample size needed to observe at least one case of the added effect is only 500. But when the background incidence is 1/10 and the added incidence is only 1/10,000, the sample size must be 98 million. These illustrations merely indicate what sample size is required to observe an effect when background incidence is known.

Controlled clinical trials are used primarily for evaluating drug efficacy, not safety, because they are carried out on hundreds, or, at the most, a few thousand drug users. Their use for evaluating drugs already on the market is also limited by their high cost and logistical problems. In fact, the use of controlled clinical trials for determining efficacy alone is already constrained by these two factors (9,46).

These limitations of controlled clinical trials in evaluating the safety of marketed drugs have led to relying on cohort and case-control methods for postmarketing studies. While these latter methods can only indicate an association between a drug and observed conditions, not that the relation is causal (49,77), the cumulative experience of multiple cohort and case-control studies showing consistent associations between a drug and such an effect can lead to a high degree of confidence that the relationship is causal. The most prominent examples of drug studies showing consistent associations are those on oral contraceptives and the risks of cardiovascular disease; similar examples of nondrug studies are those on the risks of smoking.

ISSUES AND OPTIONS

Issue 1:

Revising premarketing tests and shortening the drug approval process.

The efficacy and safety tests in animals and humans specified in FDA regulations for premarketing approval are based on broad statutory language. Efforts to shorten the drug approval process have focused not on the statutory language but on the regulations issued by FDA to implement the law. Thus, the focus here is on oversight issues, not on legislative changes.

Proposals to curtail or eliminate phase 111 premarketing tests, or shift them to the postmarketing period, can be evaluated both theoretically and experientially.

Theoretically, phase III testing is significantly more sensitive than phase II testing. Adverse effects with an incidence of 1 /100 or more are more likely than not to be detected in the 100 to 200 patients given a drug in phase II. But the theoretical sensitivity of detection rises in phase III to 1/500 with 500 patients and to 1/1,000 with 1,000 to 3,000 patients (see ch. 4, table 5).

These observations are relevant to the detection of adverse reactions, but they are not so relevant to the detection of therapeutic effects. Since a drug that helps only 1 in 100 patients would not be very effective, efficacy should be established in phase II. Phase III is intended to gather additional evidence on a drug's effectiveness for specific indications.

If phase 111 testing were curtailed or eliminated, there is also the question of whether premarketing evaluations would test sufficient numbers of patients to reasonably ensure a drug's safety or give substantial evidence of its efficacy. Even under current regulations, the use of a drug on human subjects is very limited before the drug is released for market: 20 to 80 patients in phase I; 100 to 200 patients in phase II; and 500 to 3,000 patients in phase III—a range of only 620 to 3,280 patients per drug (excluding controls).

In addition to theoretical criteria, experiential criteria could be applied in considering proposals to curtail or eliminate phase 111 tests. The diminished power to observe adverse drug effects that such changes theoretically entail may not in fact be found, judging on the basis of actual experience in phase 111 testing, or if it is, it may only concern infrequent, minor effects. Agreement of the experiential data with the differences theoretically expected would strengthen the hypothesis that curtailing phase 111 would lower the capacity of current premarketing tests to identify adverse reactions. If the experiential data fail to detect the theoretical differences, then a better case can be made for curtailing phase 111, with or without transfer of some of its testing to the postmarketing period.

Current interpretations of the statutory requirements for "adequate tests" of safety and "substantial evidence" of efficacy emphasize methodology, as reflected in the requirement that each indication for which a drug is intended be supported by at least two well-controlled clinical trials. But FDA can alter the criteria by which it approves drugs. For example, propanolol, the first beta-blocking drug approved for use in the United States, was approved by an advisory committee on the basis of all the evidence presented to FDA, even though no one study was found to be adequate and well controlled (21). And in late 1981, timolol, another beta-blocker, was approved, on the basis of evidence from a foreign study, for use in preventing death and recurrent heart attacks in patients who have survived initial heart attacks (26).

The approval of propanolol and timolol illustrates that FDA can grant exceptions to its usual requirement of two well-controlled U.S.-based clinical trials. In such cases, expert judgment relies on qualitative, not quantitative, criteria in approving a drug, and such an approach falls outside the theoretical and experiential guidelines outlined above. If FDA is to rely increasingly on such qualitative criteria through increased use of advisory committees, it will be necessary for FDA to develop general guidelines to aid the advisory committees in their deliberations. Otherwise, in a case-by-case analysis, evidence of the same quality may lead to approval for one drug and nonapproval for another.

Issue 2:

Improving postmarketing surveillance and its role in the drug approval process.

Even if phase 111 testing were not curtailed or eliminated, FDA's powers in the postmarketing period could be strengthened to enhance its surveillance role.

Postmarketing surveillance "systems" that have been advocated are not systems in the formal sense, but a series of related activities oriented toward several purposes, with the regulatory approval process being only one. Three activities are most frequently mentioned. First is the building of a resource base through training of additional experts and improving epidemiologic tools such as methods for cohort and case-control studies. Second, unless a drug effect has a sufficient fre-

quency of occurrence (usually identified as 1/1,000) and for delayed effects of, for example, greater than 1 year, strengthened voluntary reporting is the most realistic method of identifying possible adverse drug reactions. Once such reactions are suspected, clinical trials, case-control, and cohort studies could be used to determine whether an association with drug use in fact exists. Third is the development of an efficient method for monitoring selected drugs after their release into the market. The most frequently mentioned mechanism is formation of prospective cohorts of drug users.

These aforementioned components of a postmarketing surveillance "system" and FDA's role in supporting and using them are oversight issues.

There are also several legislative options that could strengthen FDA's powers in the postmarketing period. The following legislative options are presented for congressional consideration.

Option 1: Give FDA the power to require postmarketing studies.

A variation of this option is for FDA to use its existing regulatory powers over advertising and promotional practices to "certify" an industry-sponsored postmarketing study.

Option 2: Give FDA the power to restrict the distribution, dispensing, and administration of a drug.

A variation of this option is for FDA to use its existing regulatory powers to develop a parallel approval process for the use of a limited group of drugs during phase III testing, such as for drugs of unusual need and promise.

Option 3: Change the standard for a drug's removal from the market from "imminent hazard to the public health" to "unreasonable risk of illness to any segment of the population" or some other less stringent standard.