FEDERAL STATISTICS

report
of
the
President's
Commission

Volume II
1971

For sale by the Superintendent of Documents, U.S. Government Printing Office
Washington, D.C. 20402 - Price $4 per set of 2 volumes; sold in sets only
Stock Number 4000-0269

Statistics and Data Analysis in the Food and Drug Administration

By
Cuthbert Daniel
Edward R. Tufte
Joseph A. Kadane

TABLE OF CONTENTS

SUMMARY OF REPORT

The Bureau of Drugs has regulatory responsibility for all medicinal drugs and devices developed, produced, and consumed; it develops standards and conducts research on the efficacy, reliability, and safety of drugs; it reviews New Drug Applications, operates an adverse reaction reporting system, and performs many other drug-related functions. The vast scope of these responsibilities is indicated by the following figures:

--In 1970 doctors wrote about two billion prescriptions, including some 225 million for "mind-affecting drugs" – stimulants, sedatives, tranquilizers, and the like. By 1975, there will be three to five billion prescriptions written each year.

--About 1.5 million hospital admissions each year are due to illnesses caused by drugs.

--In recent years, the FDA has had to review over 4,000 drugs for efficacy.

The Bureau of Drugs received a budget in FY 1971 of $17 million – a very small budget for the job to be done. About one penny per prescription written will be spent this year by the

Bureau of Drugs to evaluate the safety and efficacy of old and new drugs, to inspect the manufacturer of drugs, to certify some types of drugs, to assess adverse reactions caused by all drugs, and to perform its other tasks. It is in the context of such a small budget and limited scientific competence that the Bureau's statistical work is done.

With respect to the statistical work of the Bureau of Drugs, we have found:

1. The last one and a half years have seen a considerable improvement. There is still a long way to go.
2. To do its job adequately at present, the Bureau needs about ten Ph.D. level statisticians and data analysts of various types. It now has two to four of those needed.
3. About ten statisticians with less training are needed. Half of these are now at work.
4. The computer operations in support of statistical work are negligible. In the short run, three scientific programmers would make better use of the currently under-used facility.
5. Statisticians should fully participate as peers in the review of New Drug Applications.
6. The FDA and the drug industry should work to assure that material included in New Drug Applications is relevant to the drug under consideration.
7. The New Drug Applications (with the exception of legitimate trade secrets contained in them) and the FDA's evaluation of those applications should be made public and open to all interested parties.
8. Drug surveillance and adverse reactions reporting are very weak within the Bureau and currently produce little useful output. The Bureau needs external help. The Bureau should participate in the letting of large-scale contracts for the monitoring of drug reactions. In planning such proposals, not enough attention has been given to the analysis of the data.

SECTION O - THE WORK OF THIS REPORT

We have gathered information for this report from interviews, including meetings with many officials in the Food and Drug

Administration (FDA), statisticians employed in several drug houses, officers of the Pharmaceutical Manufacturers Association, several drug researchers and statisticians with academic affiliations, and others. We were met with uniform courtesy by these busy men and women. In addition, we have consulted many printed sources, including the *FDA Papers,* some internal FDA materials, the extensive Congressional hearings dealing with drugs and the FDA, and some of the many books, papers, and newspaper articles published on the subject.

This report first describes the responsibilities of the Bureau of Drugs and the role of statistical analysis in assuring that drugs meet the legal requirements of safety and efficacy. We then turn to our specific conclusions and recommendations in six areas related to the statistical work of the Bureau of Drugs.

SECTION I – THE RESPONSIBILITIES OF THE BUREAU OF DRUGS IN THE FOOD AND DRUG ADMINISTRATION

The Bureau of Drugs has the basic regulatory responsibilities for medicinal drugs and devices developed, produced, and consumed in the United States. According to the February 1, 1970 reorganization statement for the Food and Drug Administration, the Bureau of Drugs

> develops standards and medical policy and conducts research on efficacy, reliability, and safety of drugs and devices for man; reviews and evaluates New Drug Applications and claims for investigational drugs; conducts clinical studies on safety and efficacy of drugs and devices; operates an adverse drug reaction reporting system; oversees surveillance and compliance programs on drugs and devices; provides scientific and technical support in drug biology and drug chemistry; assumes responsibility for regulations, model codes, and other standards covering drug industry practices and fosters development of good manufacturing practices; oversees the antibiotic and insulin certification program.[1]

1 *FDA Papers,* May 1970.

The vast scope of these responsibilities is indicated by the following figures for drug usage under FDA responsibility:

--In 1970 consumer expenditures for prescription and non-prescription drugs and devices were about $19 billion.[2]

--In 1970 doctors wrote about two billion prescriptions. Some 225 million of these prescriptions were for "mind-affecting drugs" – stimulants, sedatives, tranquilizers, and the like.[3]

It is estimated that by 1975 doctors will write three billion prescriptions.[4]

--Dr. Charles C. Edwards, the current FDA commissioner, wrote: "According to the 'Drug Utilization Review and Control Report', made by Dr. Donald C. Brodie of the Health Services and Mental Health Administration and issued last April (1970), it has been estimated that the incidence of complication in drug therapy is roughly 10 percent, and that approximately 5 percent of the patients admitted for medical treatment in general hospitals are admitted because of serious drug reactions. It is also estimated that approximately 1.5 million hospital admissions a year are due to illnesses caused by drugs."[5]

--In FY 1969, the FDA (among its many tasks with respect to drugs): reviewed over 2,500 original and supplemental applications for new drugs for human use, certified over 24,000 batches of antibiotics, insulin, and colors, initiated 220 establishment inspections under the Intensified Drug Inspection Program, and expanded testing for bioavailability of drugs and research on mycotoxins, cyclamates, and oral contraceptives.[6]

--In recent years the FDA has started to implement the

---

2 Charles C. Edwards, "Rational Drug Therapeutics," *FDA Papers,* February 1971, p.4.

3 "Growing Use of Minding-Affecting Drugs Stirs Concern," *New York Times,* March 14, 1971, p. 36.

4 Charles C. Edwards, Interview, *U.S. News & World Report,* April 19, 1971, p. 52.

5 Charles C. Edwards, "Rational Drug Therapeutics", *FDA Papers,* February 1971, p. 4.

6 FDA background material.

1962 Kefauver-Harris Amendments to the Food, Drug, and Cosmetic Act. These amendments require that drugs be shown to be efficacious, as well as safe. The 1962 law applied to new drugs coming on the market as well as to drugs that came on the market between 1938 and 1962. The law has thus required the FDA to make some judgment about the efficacy of thousands of drugs. The basic job was given to 30 review panels of physicians and dentists selected by the National Academy of Sciences-National Research Council Drug Efficacy Study Policy Advisory Committee. These panels produced 2,824 reports for 4,349 drug products. For the last two years the FDA has reviewed these reports involving more than 10,000 drug claims (plus the estimated five times as many similar products not studied) in an effort to reach a decision as to which drugs did not meet the efficacy requirements. This difficult job has been carried on along with the usual work of the FDA.

The range of work is shown in the specific responsibilities given to the Offices within the Bureau of Drugs:

Of New Drugs: Evaluates, for safety and efficacy, New Drug Applications (NDA's) for marketing new drugs; evaluates adequacy of proposed labeling for use and warning against misuse; evaluates manufacturing and laboratory methods, facilities, and controls in factories producing new drugs; reviews notices of claimed investigational exemption for new drugs (IND's) and recommends action to restrict or stop further testing; reviews clinical investigators and scientific investigations of investigational new drugs and New Drug Application areas and coordinates follow-up with the Office of Compliance.

Divisions: Anti-Infective Drugs, Cardiopulmonary and Renal Drugs; Dental and Surgical Adjuncts; Metabolism and Endocrine Drugs; Neuropharmacological Drugs; Oncology and Radiopharmaceuticals, and Scientific Investigations.

Of Marketed Drugs: Evaluates safety and efficacy data and proposed labeling in supplements to New Drug Applications; carries out continuing surveillance and medical evalu-

448-559 O - 71 - 6

ation of labeling, clinical experience, and reports required of applicants for all drugs and devices for which new drug approval is in effect; reviews inspections and other findings to determine if new drugs are being marketed in accord with commitments in New Drug Applications; makes recommendations on withdrawal of approval of the NDA; takes final action on antibiotic and insulin samples submitted for certification and on requests for exemptions from antibiotic certification; reviews for safety, reliability, and effectiveness the new and marketed therapeutic and clinical devices and recommends action on significant hazards or potential danger from inadequacy of direction for use or warning and cautionary information; obtains and evaluates reports of adverse drug reactions.

Divisions: Certification Services, Clinical and Medical Devices, Drug Experience, Cardiopulmonary-Renal Drug Surveillance, Metabolic-Endocrine Drug Surveillance, Neuropharmacological Drug Surveillance, and Surgical-Dental Drug Surveillance.

Of Compliance (Drugs): Advises the Bureau Director and other officials on the law, regulations, legal-administrative problems, regulatory problems, and administrative policies concerning regulatory responsibilities for drugs and devices; conducts studies to determine medical policy and support regulatory action; develops compliance and surveillance programs covering regulated industries; develops or coordinates development of regulations and other standards covering industry practices and fosters development of good manufacturing practices; conducts programs to encourage voluntary compliance by industry; on request, supports and guides District offices in handling legal actions and provides headquarters case development, coordination, and assistance in contested cases; develops and coordinates studies on degree of compliance by regulated industries with statutes and regulations enforced by FDA; monitors and evaluates professional journal advertising and promotional and related labeling to determine veracity of claims.

Divisions: Case Guidance (Drugs), Compliance Programs (Drugs), Drug Advertising, Industry Services (Drugs), Medical Review, and Policy and Regulations.

Of Pharmaceutical Sciences: Provides scientific support for drug compliance programs; develops scientific support for drug compliance programs; develops scientific standards and conducts research on composition, quality, and safety of drugs; operates system for continuous appraisal and improvement of current and proposed drug standards and specifications; devises new chemical, physical, and biological methods to analyze drugs in pharmaceutical preparations and in tissues and body fluids; investigates mechanisms of the underlying chemical reactions; explores use of novel instruments and equipment; designs and participates in collaborative studies to establish the reliability of new methods and to validate important discoveries relating to drug examinations; operates the National Center for Drug Analysis (St. Louis) and the National Center for Antibiotics and Insulin Analysis (Washington); cooperates with the Committee of Revision of the U.S. Pharmacopeia (USP) and the National Formulary (NF) to compose and assemble monographs for inclusion in official drug compendia.

Divisions: Drug Biology, Drug Chemistry, National Center for Antibiotics and Insulin Analysis, and National Center for Drug Analysis.[7]

In order to perform all these jobs in FY 1971, the Bureau of Drugs received a budget of $17 million. This figure, though painfully small, represents a significant increase over previous appropriations. By almost any standard, it is a very small budget for the job to be done. One way to put it into perspective: for each prescription filled this year in the United States, about one penny will be spent by the Bureau of Drugs to evaluate the safety and efficacy of new and old drugs, to inspect the manufacture of drugs, to certify some drugs, and to assess the adverse reactions of all drugs.

---

7 *FDA Papers*, May 1970.

It is in this context that we turn to the role of statistics and quantitative analysis in the Bureau of Drugs.

SECTION II – THE ROLE OF STATISTICS AND DATA ANALYSIS IN THE BUREAU OF DRUGS

In this section we seek to

(1) show how some types of statistical analysis, when combined with good medical judgment, are necessary if the Bureau of Drugs is to meet its responsibilities under the law,
(2) point to the particular places in the Bureau of Drugs where particular quantitative and statistical tools would prove useful, and
(3) evaluate the current statistical and quantitative work in the Bureau of Drugs in comparison to the work it needs.

The basic job of the Bureau of Drugs is to assess evidence concerning drugs. Such evidence consists of chemical and pharmacological data, the results of experiments on test animals, and human experience with drugs. The assessment of such evidence requires a diversity of skills although the decisions on the clinical significance of the effects of a drug on humans must rest with clinically trained officials. A large part of the evidence on which the decision-making official relies is quantitative, in the form of counts, measurements, or subjective records from a number of cases – measurements which are often taken by a number of observers in different clinics. Single examples rarely suffice, except when they indicate that more cases should be accumulated. At the other extreme a full enumeration of a population is rarely needed or indeed possible.

The officials making decisions about drugs must usually deal with quantitative evidence gained from experimental designs collecting data which are samples of a larger population. Statistics is the art and science of dealing with samples, constructing experimental designs, and analyzing quantitative data – and its techniques can, upon occasion, contribute to the collection of useful information about drugs and the making of sound decisions concerning the safety and efficacy of drugs.

Experienced statisticians and data analysts can contribute to the solution of such problems as:

--How are data taken in different clinics and in different clinical trials to be combined or contrasted? This difficult problem occurs in a great many New Drug Applications as well as when the drug is on the market.

--What is the most economical way to monitor the effects of total drug consumption in the nation?

--What types of experimental designs will reliably and economically obtain information about the safety and efficacy of a drug?

--What constitutes a fair set of tests for the claimed therapeutic equivalence of two or more drugs manufactured by different companies?

--Does the average life span of this set of 20 mice – which have been treated with drug A at a low dose-level throughout their lives – differ importantly from the life span of this other set, which have had no drug A?

--Does a bias in the selection of patients for treatment or control groups invalidate the results of a particular study?

--Does the method of data analysis bias the results? Would another method show a different conclusion?

--How large a sample is required to detect a side-effect that occurs with a frequency of less than one in a thousand?

--Can drug A be judged better than drug B, the present drug of choice, for treating a particular infectious disease?

--Does the finding of 17 defective bottles of a manufactured product justify withdrawal from market, or a more intense search?

--Do early human trials on a few volunteers give a sufficiently clear picture of safety and efficacy to justify larger trials to look for human variation, for the dependence of response on dosage, and for relatively rare side effects?

--Or consider a more detailed question: 20 test organisms are often used in standard toxicological trials for estimating the "no effect dose." Suppose a dose or doses schedule for some

fixed period is found to affect adversely *none* of the 20 test organisms. What can be said about the true proportion of the whole population that may respond adversely? In this particular example, all that can be said with tolerable security (i.e., with a 95 percent confidence of being right) is that less than 14 percent of the population will be affected. It is assumed that there is *no* error of measurement or judgment in making the study and in concluding that no animal was adversely affected. If there is such error, then the limiting proportion is greater than 14 percent. This example is given space here because it is a near-scandalous fact that nearly all otherwise qualified toxicologists have difficulty believing this relation, even though its logic was clearly understood by Blaise Pascal in 1640.

All these questions require mature clinical judgment as well as statistical backing. Sometimes, of course, the case is so clear that professional statistical aid is not needed. But in complex problems (lacking cookbook solutions), and in close decisions, the statistician will play a key role, easing the clinician's problem by warning him that the data and its analysis do not warrant a claim, or by telling him that there is little doubt that a real gain is present. At times there will be a conflict between statistical and medical judgment. Medical doctors may be convinced by personal observation of a single case. Often they will be right in their judgment and perhaps only a single replication is needed for proof. But generally the way to find out if they are right is to look quantitatively at more cases, in a controlled design, adjusting for patient, diagnostic, and clinic variation. The collaboration between doctors and statisticians will not always be a happy one – but it is a necessary collaboration to guarantee objectivity in the work of the Bureau of Drugs.

It is now time to relate specific statistical and quantitative techniques to specific sections in the Bureau of Drugs. We first present a list of statistical aids in generation, collection, evaluation, and interpretation of data. We then give a table showing the relationship between these statistical tools and the work of the Bureau.

The following statistical tools may be useful in the assessment of the safety and efficacy of drugs:

a. Learning about the properties of large populations by full enumeration. Censuses. For example, the full record of qualified clinicians.
b. Sampling large populations to make projections (inferences) about the whole. For example, how is the percent defective bottles of a drug distributed, both among batches for one producer and among producers?
c. Estimating population properties, and differences between properties of several populations. For example, differences in effects of drugs, dose-dependence by bioassay.
d. Controlling quality (of data, of products, of system operation) by sampling continuous processes. Making decisions using variable data about shifts in the underlying system. For example, quality control charts for purity or strength of insulin batches from several producers.
e. Data analysis and fitting equations to data. Concise summary of complex quantitative information. For example, representing a process in which ten independently operating conditions affect each property of the product.
f. Designing experiments and comparative tests. Reducing bias and increasing precision in multifactor experiments.
g. Managing large files of information with computer. Accrual, storage, editing, retrieval, tabulation. For example, name and chemical composition for all prescription drugs.
h. Developing model-equations to fit processes that require probabilistic description.

Chart 1 shows how these statistical methods are related to the function of the Bureau of Drugs. An asterisk (*) indicates that some statistical work in the particular area is now being carried out. An X indicates that significantly more statistical aid is required to meet the present needs of the Bureau. We find a considerable shortage in almost all areas, although the last year has seen considerable improvement with respect to statistical help. We suspect that two or three years ago there would have been few asterisks – indicating that some statistical work is done in a particular

area – in our table. Thus we believe that the statistical work in the Bureau of Drugs is moving in the right direction, but the responsibilities of the Bureau of Drugs to patients, to physicians, and to the pharmaceutical industry can only be fully met when decisions are based on data of guaranteed quality, and are analyzed by competent clinicians and statisticians.

Given that statistical tools have helped and will continue to help the Bureau of Drugs, we now turn to findings and recommendations with respect to statistical operations in five aspects of the work of the Bureau of Drugs.

BUREAU OF DRUG FUNCTIONS

Statistical Areas	New Drug Applications		Compliance		Pharmaceutical Testing		Post-Marketing Surveys		Clinical Research and Spec. Studies	
Census, Surveys	*		*	X			*			
Sampling			*	X		X		X		
Estimation Bioassay	*	X	*	X		X	*	X	*	X
Quality Control				X		X				
Data Analysis, Fitting Equations	*	X			*	X	*	X	*	X
Design of Experiments	*	X				X			*	X
Data Processing	*	X	*	X	*	X	*	X		
Statistical Modeling								X		

CHART 1

SECTION III – THE STATISTICAL WORK OF THE BUREAU OF DRUGS: FINDINGS, CONCLUSIONS, AND RECOMMENDATIONS

1. New Drug Applications
2. Drug Surveillance and Evaluation
3. Computer Operations

4. External Statistical Support for the Bureau of Drugs
5. The Bureau of Drugs: A Place for Statisticians

1. New Drug Applications

Some 70-80 New Drug Applications (NDA's) are submitted each year to the Bureau of Drugs. The Bureau evaluates the evidence supporting the safety and efficacy of the drug in order to decide if the drug can be placed on the market. The table shows the number of new NDA's and resubmissions received in 1969 and 1970 by the Bureau, as well as the number of Investigational New Drugs (IND's).

FY	New NDA's	Resubmissions	IND's
1969	71	149	835
1970	75	119	1122

The volume of material in each NDA is large, from a few hundred to many thousands of pages. Each records, among other things, the details of one or more clinical trials, sometimes as many as 20. At present it appears that the Bureau does not have sufficient professional staff to do justice to all this work. Some trials are studied carefully, some necessarily only scanned. This is unfair to the applicants, to the medical and pharmacological FDA officials who must make decisions on the NDA's acceptability, and to the general public who are affected by those decisions.

Statisticians in the Bureau of Drugs have started to participate in the evaluation of the NDA. *Their participation should be extended so that statistical judgment is applied at the beginning of the NDA evaluation process.* Statistical thinking has played a major and useful role in the recent FDA guidelines on drug testing (one general set and 29 sets covering specific types of drugs). These guidelines, negotiated between the FDA and the industry, help set ground rules for useful experimental designs. The introduction to the general guidelines states:

> *Statistical expertise is required in planning, design, execution and analysis of clinical investigations and clinical pharmacology in order to ensure the validity of estimates of parameters for safety and efficacy obtained from these*

studies. It is always desirable in planning and conducting such studies to have the active participation of biostatistician(s).

This is followed by a listing of the usual principles of good clinical trial design, most of which are of statistical origin.

The classification of clinical trials into Phases I, II, and III (respectively toxicological trials on small numbers of subjects, perhaps 3-14; efficacy and safety trials on larger closely monitored groups, say 20-200; larger trials for details on dose-dependency, tolerance, side effects, requiring usually from 100 to 1000 subjects) corresponds also to increasing statistical requirements, both in their planning and in their analysis.

Put in rough time order of their appearance, statistical methods would be expected to be decisive in:

1. Specifying target populations and corresponding sampling schemes. (Diseased or healthy population? A placebo or present drug of choice for control? etc.)
2. Setting size of trial (allowances for dropouts).
3. Randomization in allocating treatments to subjects (overall, blocked or otherwise restricted).
4. Monitoring quality control on data accrual and editing.
5. Interim cross-tabulating and analysis.
6. Final data analysis (specified in as much detail as possible beforehand). Numerical analysis of multiply-classified data, never exactly matched on all important factors.

The continued development and acceptance of the guidelines on drug testing may help break an unhealthy pattern of conflict and confusion between the FDA and those applying for a New Drug Application. The vicious circle (described in our interviews with employees of the FDA and the drug houses as well as officials in the Pharmaceutical Manufacturers Association) traps the NDA applicant and the FDA in a huge volume of research reports, clinical data, bibliographies, and other materials all going to make up NDA's consisting, in at least one case, of 180 volumes of material, each three inches thick.

Although our interviewees disagree somewhat on where the responsibility for the problem rests, they all describe the following

pattern: An NDA is submitted and the FDA has 180 days to review it. Since the FDA lacks the scientific resources to evaluate the large number of NDA's in the time allowed, it sometimes holds the NDA for a few months and then informs the manufacturer that the NDA is incomplete. New material must be submitted and then the 180-day clock is started again. The applicants for an NDA, aware of the many ways in which an NDA can be "incomplete," often submit NDA's containing every sort of material possibly relevant to the new drug as well as a good deal of material of no relevance. Such material is frequently submitted because genuine evidence for safety and efficacy is simply lacking. In fact, a major share (perhaps more than half) of NDA's are rejected out of hand for obvious shortcomings. We suspect that some NDA's are submitted merely on the hope that they might get by.

One FDA official described a similar problem with the material submitted for the assessment of the efficacy of a drug: "In one instance, in the objections filed to our implementation of the NAS-NRC recommendations, there was submitted a list of more than 100 so-called scientific studies and reports purporting to show something regarding a particular drug. It took considerable time for our legal and scientific staffs to establish that this evidence was mostly irrelevant and obviously inadequate.

"Such a submission does a disservice to the Food and Drug Administration (even if we prevail), to you who are upholding professional standards, and to the public for whom we should all be spending our time more productively, and finally to the drug industry."[8] Employees of the drug houses (both in our interviews and as reported in a paper by Louis Lasagna) also indicate mixed experiences with the FDA review of NDA's. Lasagna wrote: "One *hears conflicting stories about FDA's handling of NDA's. Some drug house employees state that they have been treated in exemplary fashion by FDA monitors; others complain of stupidity, arrogance, unreasonable demands, and delays lasting for years.*"[9]

8 John Jennings, M.D., Assistant to the Commissioner for Medical Affairs, quoted in the *FDA Papers,* September 1970.

9 Louis Lasagna, "1938-1968; The FDA, the Drug Industry, the Medical Profession, and the Public," in John E. Blake, ed., *Safeguarding the Public; Historical Aspects of Drug Control* (Baltimore: The Johns Hopkins Press, 1970), p. 175.

The generally proposed solutions to this widely recognized problem are to raise the level of scientific competence of the FDA (so that the review of the NDA can be completed in the 180-day time period without having to negotiate for further time), and to make changes in the procedures for handling NDA's. Such solutions are badly needed. We are concerned in this report with the problem of voluminous and irrelevant NDA's because, as statisticians continue to take a more active role in the review of NDA's, they too will find their skills wasted in wading through the masses of marginal material.

As a modest step toward reducing these difficulties, we recommend that the cooperation displayed in developing guidelines for research designs for new drugs be continued by the statisticians of the drug houses and those of the FDA. A special problem, deserving the contemplation of both sets of professional statisticians, is the assessment of multiclinic data. New drugs are often tested on small groups of patients in many different places by many different investigators. The quality of the data generated ranges from fraudulent to superior. Even without bad data, combining the voluminous results of many different small, dissimilar clinical trials is difficult and time consuming. These problems with compatability of data need help from all sources in the FDA and the industry.

The law requires the FDA to protect "trade secrets" in the NDA's from becoming public knowledge. At present the FDA interprets this law to mean that the entire NDA is to be protected. We believe that this interpretation should be re-evaluated in order to allow greater scientific and public oversight of the NDA process. We recommend that the FDA publish the reasons for its decisions on the NDA, along with the relevant studies of the drug submitted in support of the NDA. The publication of the analysis of the NDA and the studies decisive to the analysis would have important advantages over the current system. First, the FDA would be going on record in publishing reasons for its actions. In itself, this could lead to better decisions and strengthen the nonpolitical character of those decisions. Second, it would create a publicly accessible body of common tradition on what constitutes

acceptable evidence for the safety and efficacy of a drug. Third, by making clearer what is specifically required with respect to clinical and statistical evidence, the publication procedure might reduce the submission of overly long NDA's.

We have not looked into the specific mechanics of publication. Once the FDA has written down the reasons for its decision, the obvious steps toward publication include: selecting the appropriate material for publication, checking it for legitimate trade secrets, and distributing the material to a subscription list. Subscribers would include drug companies, medical schools, statisticians, drug authorities in other countries, as well as those in other government agencies, such as National Institutes of Health (NIH). Individual physicians might be interested in the material on drugs of special concern for them. We hope that the publication of this material would widen the scientific audience interested in the evidence offered in support of decisions concerning the safety and efficacy of new drugs, as well as improve the quality of the decisions and the evidence submitted in the NDA. We believe that this proposal could prove to be of major importance and it should be given a fair trial.

2. Drug Surveillance and Evaluation

The purposes of monitoring the experiences with drugs on the market are to detect adverse reactions and to improve the efficacy of drug therapy. Assessing the costs and benefits of particular drugs is an important and difficult enterprise requiring medical, statistical, and computer expertise. In this brief report, we will indicate the scope of the problem of drug surveillance and evaluation. We will then consider FDA's response to the problem, including its plan for developing a national drug monitoring program as well as its current adverse reaction program.

An effective drug reaction monitoring system would provide information for a cost-benefit analysis of drug usage. Such information would include:

"(i) the definition of reasonably precise probabilities for the efficacy and toxicity of alternative treatments available for a given

condition; and (ii) computer-based correlation of patients' characteristics and drug response so that the doctor will be able to tailor drug treatment to the needs of the individual patient with a predictability of response that is not available today."[10]

A decade ago, the first oral contraceptives were made available and shown to be highly effective and widely acceptable. Side effects were noted from the start, but oral contraceptives have become increasingly accepted nonetheless. Many of these side effects are transient and disappear after a time, and many occur in some users and not in others. Other possible side effects are potentially very serious and occur seldom enough that an individual physician cannot hope to make judgments about them on the basis of his own experience. It is this latter class of side effects with which I shall be concerned.

There are two aspects of public policy with respect to such effects which I believe require clear identification.

There is, first, the right of an informed individual to elect to take a reasonable risk in order to achieve some goal which he believes desirable. I make several trips a year from Chicago to Washington. Purely for convenience I travel by air and not by rail. I am one of the few who examined the available information on risk in the two modes of travel and I judge that I am accepting a non-negligible risk in choosing the convenience of air travel, but I do it quite deliberately. On the other hand, I no longer smoke cigarettes, although I was once a regular smoker. The benefits were in that case outweighed for me by a number of serious disadvantages. I am a strong believer in the freedom of the informed individual to decide for himself how to balance benefits and risks which primarily concern him, and I include here the right of the patient, consulting with her physician, to elect to accept some risk in return for the benefits of oral contraceptives. To me, the role of government in this area is clear – it is to be sure both physicians and patients are as well-informed about the state of knowledge on the benefits and

10 Louis Lasagna, "The Pharmaceutical Revolution: Its Impact on Science and Society," *Science*, 66, 5 December 1969, p. 1228.

risks of oral contraceptives as it is reasonably possible to make them.

A second aspect of the public policy is of comparable importance. It is to protect the public from clearly unreasonable risks, both by seeing that the risks are adequately studied and understood and, in some cases, by eliminating them from the environment. In some respects the level of concern by government should be much greater than that of even a prudent individual. For example, an increase in the level of ionizing radiation over the United States due, perhaps, to building nuclear power stations, would increase the risk that I might get leukemia. If the increase were small, I might reasonably ignore it, as an individual. The government, however, must weigh the combined risk to all of us against the benefits which additional nuclear power plants might provide.

In either case it must be emphasized that reasonable judgments are only likely to arise out of weighing of risks and benefits. It is common to hear that "even one avoidable death is too many," or that the "only acceptable level of pesticide residue on food products must be zero." Neither as governments nor as individuals do we in fact behave this way, and the refusal to weigh risks against benefits often results in the blind acceptance of risks which analysis would show to be unreasonably high. I believe, for example, that the almost disastrous errors made in the program of safety-testing for the Salk vaccine in the mid-1950's resulted from an unwillingness to accept the notion that there might be a measurable risk and to seek to evaluate its magnitude. A typical example of a known risk which we accept in return for benefit is the requirement of smallpox vaccination for school children. A small but definite number of children die as a result of being vaccinated. We accept such costs as necessary to achieve a greater benefit.[11]

11 Statement of Paul Meier, Ph.D., Professor of Statistics, Department of Statistics, University of Chicago, in Competitive Problems in the Drugs Industry, Hearings Before the Subcommittee on Monopoly of the Senate Select Committee on Small Business, Part 16 (February-March, 1970) 6548-6549.

CURRENT DRUG MONITORING

As noted earlier, there are a good many adverse reactions to therapeutic drugs. Perhaps 1.5 million hospital admissions yearly are due to illnesses caused by drugs. For certain illnesses drugs obviously have tremendous benefits. In general, however, cost-benefit data for particular drugs are remarkably thin. "Decisions regarding drug treatment must frequently be made without adequate knowledge of the clinical effects of the drugs at issue. The extent of this problem was recently emphasized in a report by the Divison of Medical Sciences of the National Academy of Science-National Research Council (NAS-NRC) following an intensive review of over 3,000 drug formulations marketed between 1938 and 1962. The review panels of this *Drug Efficacy Study* rated about 7 percent of the preparations as 'ineffective,' and with a large proportion (a majority) of the remainder, the information supplied by manufacturers was considered insufficient to fully assess efficacy. (20 percent of all drug claims were rated as 'effective'; 39 percent of all drugs were rated as 'effective.') Individual members of the various panels were also invited to submit their thoughts on the insights gained from their participation in this study, and, according to the report, 'letter after letter expresses concern and surprise about the generally poor quality of the evidence of efficacy of the drugs reviewed."[12]

Consider one of the most widely studied drugs, halothane. In their formal recommendations, the authors of *The National Halothane Study* comment on the shortage of information:

> *We recommend the establishment of a cooperating group of institutions to serve as a panel-laboratory for the acquisition of trustworthy information on new drugs (not merely anesthetics) as they come into use.*
>
> *In the history of medicine, it is doubtful whether any drug was ever more extensively studied both before and after its introduction than halothane. Yet, after halothane had been given to patients perhaps ten million times, it was impossible to give firm, reliable answers to many basic questions*

12 Hershel Jick, *et.al.*, "Comprehensive Drug Surveillance," *Journal of the American Medical Association,* 213 (August 31, 1970), p. 1455.

about its effects. Two such questions were: "How does the death rate after operations under halothane anesthesia compare with death rates when other anesthetics are used?" "Does halothane induce significantly more hepatic dysfunction than other widely used anesthetics?" The National Halothane Study attempted to answer these questions by using existing records. Although 856,500 operations were brought under scrutiny, the answers given are predictably and regrettably short of those desired. For example, the important questions of nonfatal hepatic injury was not taken up by the study. The limitations of knowledge on halothane are certainly not peculiar to it. Limitations at least equally compelling apply to nearly any drug introduced in the past. Had halothane been administered a few scores of thousands of times in the context of any experimental informational-gathering system, similar in kind to a cooperative randomized clinical trial, reliable information might have been acquired for over-all death rates, and possibly for nonfatal hepatic injury as well.[13]

A similar situation holds for the oral contraceptives, potent drugs used by over eight million women. The FDA Advisory Committee on Obstetrics and Gynecology in August, 1969 recommended:

"The Food and Drug Administration assures adequate surveillance of approved contraceptive drugs.

> The inadequacy of surveillance of contraceptive drug use in the United States and other countries is apparent. Voluntary reporting of adverse reactions tends to be capricious and may be misleading....

"Strengthen the surveillance system of the Food and Drug Administration.

> This recommendation from the previous report has not been satisfactorily implemented. A system should be devised so that when adverse reaction reports are received, they are

13 J. P. Bunker, W. H. Forrest, F. Mosteller, and L. D. Vandan, *The National Halothane Study* (Washington, D.C., U.S. GPO, 1969) pp. 417-418 Part VI, "Formal Recommendations" by J. P. Bunker.

448-559 O - 71 - 7

made readily and immediately accessible."[14]

Neither recommendation has been carried out.

Even for widely consumed drugs, post-marketing studies of low-incidence effects are conducted at a slow pace. The "Sartwell Report": a retrospective study of thromboembolism and oral contraceptives begins indicating the time scale involved: "The suspicion that oral contraceptives might predispose women toward vascular occlusive phenomena arose about 1961, largely from the publication of case reports. An *ad hoc* committee in 1963 advised that 'comprehensive and critical' studies to look into the possibility be conducted. Nevertheless, little was done in this direction, despite the great increase in the use of these potent drugs. By the time the Advisory Committee on Obstetrics and Gynecology of the Food and Drug Administration began to prepare its first report on the oral contraceptives in 1965, it was evident that an epidemiologic study was even more urgently needed than in 1963. The present study was begun in November 1965, in direct response to this need."[15] And this study itself was published almost four years later in August 1969. It formed a large share of the evidence that led the Advisory Committee on Obstetrics and Gynecology to conclude that there was an etiologic relation between thromboembolic disorders and the use of oral contraceptives.

CURRENT PRACTICE IN DRUG MONITORING IN THE FDA

(1) The FDA has neither the scientific talent nor the resources to conduct serious long-term studies of drug reactions in human population. It is unlikely that the FDA will, in the next few years, be able to attract the medical, statistical, and computing talent for such studies.

(2) We have reviewed two proposals circulating within the FDA concerning long-run studies of drug and chemical reactions.

---

14 Second Report on the Oral Contraceptives (Washington, D.C. U.S. GPO, August 1, 1969), p. 8.

15 P. E. Startwell, A. T. Masi, F. G. Arthes, G.G. Greene, and H. E. Smith, "Thromboembolism and Oral Contraceptives: An Epidemiological Case-Control Study," in Advisory Committee on Obstetrics and Gynecology, Food and Drug Administration, *Second Report on the Oral Contraceptives* (Washington, U.S. GPO, August 1, 1969), p. 21.

One proposal soliciting contracts for a pilot study of a National Drug Monitoring Program is well thought out except for the actual analysis of the data. Past experience indicates that such data analysis problems are usually difficult. Thus while we feel that the FDA should award contracts for long-run studies of drug reactions, further work is needed in the preparation of the proposals in order to assure adequate analysis of the data. The second proposal, concerning the long-run safety evaluation of environmental chemicals, needs further work with respect to design and analysis of the planned experiments.

(3) The adverse reaction drug monitoring system operating within the FDA obtains about 25,000 reports per year of adverse reactions from drug houses, hospitals, and others. The law requires the manufacturer of a drug to report periodically to FDA on any information it has of a possible "adverse reaction." Federal and private hospitals under contracts also submit accounts of such reactions to FDA. Additionally private physicians occasionally write to FDA about specific cases. These statements are examined by physicians employed by FDA. Presently the private hospital contracts are being phased out and the adverse reaction reports are being put on a computer.

Unfortunately these cannot be used for reaching secure conclusions. Causation cannot usually be inferred from fragmentary reports of the kind received by FDA. There is no sampling plan, and hence no sense in which the results are representative. Little systematic statistical analysis can be done since there is no way to standardize the adverse drug experiences with respect to patient characteristics or favorable reactions to the drug.

In the past, a selection of 40 or 50 adverse reaction reports, chosen by the staff on an intuitive basis, was published monthly by FDA. This system had the twin faults of quite possibly overlooking important adverse reactions, and of discouraging the use of effective drugs that were not responsible for the reaction reported. As a result this practice has been abandoned, and no published output is forthcoming from the adverse drug reaction program of FDA.

The FDA could require, as part of the approval of NDA, a designed prospective study to monitor the drug's impact. This has

already been done in the case of L-Dopa, a particularly effective drug which was given early conditional approval. Such studies permit randomization and thus easy interpretability of results. Furthermore, this method curtails the tendency to collect large amounts of unneeded information. We feel that this approach warrants more emphasis than it has received in the past.

In conclusion, we find that in dealing with problems of adverse drug reactions, the FDA has collected, and proposed to collect, large amounts of information without adequate thought about the use to which the information would be put. Greater involvement of statisticians in decisions about what to do concerning adverse drug reactions could have the desirable effect of deemphasizing large computer systems with "all the information" available, and emphasizing instead economical collection and interpretation of data to answer the important questions.

CONCLUSIONS AND RECOMMENDATIONS

Systematic drug surveillance and evaluation would provide valuable information about the safety and effectiveness of the drugs prescribed in the two billion prescriptions written per year. Information is now available for only a few drugs. The FDA, however, does not have the capability to gather further information – although it is the agency that must judge the safety and effectiveness of drugs. We recommend that the FDA, in collaboration with other government agencies related to health, seek the funds for contracting for large-scale systematic surveillance and evaluation of major therapeutic drugs – including some over-the-counter drugs. Such cost-benefit information concerning drugs will be difficult to obtain and will require the combined efforts of clinicians, epidemiologists, systems analysts, and statisticians. The Bureau of Drugs also needs to make greater use of its statisticians in planning such drug surveillance projects, particularly with respect to data analysis.

3. Computer Operations

There is very little statistical programming and computerized data analysis in the Bureau of Drugs. The under-utilized computer

available is used mainly for administrative purposes and assorted record-keeping. Only one scientific programmer is available in the Bureau of Drugs. There is a single terminal to a large machine in Washington along with inconvenient and slow access to the FDA machine. It is often said that statisticians cannot be expected to work effectively or happily without large scale computer facilities, and this statement is often true. But two qualifications should be made. It is not necessary that every statistician or statistical group have sole control of such a center. Much statistical work, some of it of major status, can be done with little or no large scale computer service.

We have observed inefficient hand calculations by Bureau statisticians of long experience but evidently poor training. We suppose that a considerable fraction of the demand for computer services comes from the frustration and delay caused by such time-wasting work by hand and desk calculators. The lack of programming support has also led to some inefficient efforts at statistical programming by Ph.D. level staff.

The Bureau of Drugs needs much greater support with respect to scientific programming: data-processing facilities of an order of magnitude greater and better than those now existing are predictably needed by the Bureau in the long run.

In view of the costs and poor record of large new computer operations, higher priorities in the Bureau, and the present incomplete plans of the Bureau, we make the following short-run recommendations:

(a) the addition of three scientific programmers in direct support of the statistical work of the Bureau;
(b) a serious effort to make effective use of the currently under-used facilities now available;
(c) efforts to improve routine computational practices now done by hand or machine. We suggest three programmers only as a serious start. It may well be that as the Bureau becomes better able to handle its work, a much larger programming effort will be required.

4. External Statistical Support for the Bureau of Drugs

The social importance of the work of the Bureau is self-evident.

Therefore we must consider what external aid can be marshalled in its support. Our suggestions fall into immediate and long-term categories.

The most obvious immediate aid could come from a group of interested senior statisticians (biometricians and biostatisticians) who would give more visibility to the Bureau's needs for able staff statisticians. Careful recommendation of individuals to whom the Bureau does not have ready access (first-rate graduate students and new Ph.D.'s, experienced statisticians not now known to be considering job changes, etc.) could gather fresh talent. We feel that the current procedures for recruiting new statistical talent into the Bureau need additional support and improvement. Another reason for care in making new appointments is the irreversibility of such appointments. While the civil service rules quite rightly protect appointees from capricious termination, they may work to retain the relatively incompetent worker in the office where his or her presence is extremely costly. We believe this has already happened a good many times in the Bureau. We do not feel it is necessary here to give a listing of all the attributes of an organization that will attract and retain a large group of professional statisticians. It must be clear to the present administrators that prospects for advancement, for professional contacts, and for the time needed to think about major problems, must be guaranteed to able staff members.

Two major long-term problems that need more time than the Bureau now has available are outlined below. They are both of immediate interest, but even tentative solutions seem some distance in the future.

1. *Cost-benefit analysis in drug development and distribution.*

   Little quantitative thinking appears to have been done on weighing the gains against the losses in developing and distributing a new drug. In the development stages a complex but obscure interplay of ethical, legal, and financial considerations undoubtedly takes place behind the study of the drug's safety and efficacy. After distribution, only minor effort, grossly disproportionate to the seriousness of problem, is now made to collect safety (and adverse reaction) information even though

such information is required by law. The result is that the consuming public constitutes the membership of a poorly conducted clinical trial. Surely here is an opportunity for a major contribution by statisticians and clinicians of FDA; but equally surely an external group could be expected to make a useful contribution.

2. *Studies of long-term and rare side effects.*

A valuable start in the area of extensive small animal trials has been made in the memorandum of November 9, 1970, submitted to the Commission of FDA. This proposal recommends a large facility (the "Pine Bluff facility") to house large numbers of test organism under controlled conditions, with, of course, a large staff. Part of its work would be to evaluate the life-long effects of "environmental chemicals" which includes drugs.

No consideration seems to have been given in that memorandum to the study of the simultaneous exposure of individual animals to more than one material. Since all humans are exposed to more than one environmental chemical, these studies will gain in validity if multiple exposures can be arranged. It is not supposed here that this extension will be easy, but only that it should be worked on by several groups, externally as well as within the Bureau of Drugs. While several proposals for nationwide monitoring of adverse drug reactions and rare side effects are now being circulated, it does not appear to us that sufficient statistical effort has been put into these plans. An external Bureau of Drugs biometric advisory committee should, not, of course, control such plans, but it might provide insights that would not otherwise be forthcoming.

Recent symposia on the teratogenicity, carcinogenicity, and mutagenicity of many environmental chemicals including drugs, indicate that these fields are well advanced and are now able to detect many such effects in small animals and in micro-organisms. This implied an increased effort of several orders of magnitude when translated into full scale studies of all suspected new drugs and chemicals. Again, sustained and sympathetic study by statistical groups outside the Bureau but in close touch with its work,

seems necessary.

5. The Bureau of Drugs: A Place for Statisticians

The Bureau of Drugs needs roughly ten statisticians of Ph.D. level or equivalent training. Of the ten needed, two to four are now at work. The number ten is not given casually and is not inflated against future compromise. It is, rather, a minimum that will soon be exceeded if the Bureau is to advance rapidly toward meeting its legal responsibilities. The primary responsibility of the Bureau is to reach sound conclusions on NDA's promptly. About 70 new NDA's are submitted each year. We concurr with Dr. Charles Anello's judgment that these require, on average, one month's attention from a senior statistician. Thus roughly six statistician years are needed for this work alone. We do not have good estimates of the time required for making summary reports on each of the thousand-odd IND's received every year, for review of protocols submitted for new drug studies, and for reviewing the supplements and additions that are frequently made to each NDA. Nor is it possible to produce an exact analysis of the time required for statistical contributions to the large drug-monitoring programs in view. Almost all requests for statistical aid in compliance problems are now simply postponed. It seems modest indeed to judge that four person-years of the time of senior statisticians will be required to do these jobs.

The ranges of specialization, experience, and competence are as wide for statisticians as for members of any other profession. Several of the new recruits should be biometricians or possibly epidemiologists; several should have had more than a year of experience in applied statistics. Most graduating Ph.D. level statisticians are not immediately qualified to do the work of the Bureau of Drugs: the Bureau needs men and women experienced in the analysis of clinical and biometric data. Thus, in recruiting new statisticians, a premium should be placed on relevant experience. The new statisticians are needed for direct work in NDA's, for improving the many sampling programs of the Bureau, for the statistical research in many areas not now at satisfactory levels (example: advanced analysis of complex data from drugs supplied by different manufacturers), for aid in designing the large computer system the Bur-

eau of Drugs will ultimately need, for mature advice to the Commissioner and his Assistants, for developing and supervising the complex systems that will be required for monitoring a national adverse drug reaction system, and finally for studying the largely unsolved problems connected with detecting the effects of long-term, multiple-drug dose-schedules.

About ten statisticians of lesser education and with less technical statistical knowledge are also required for the difficult and important sort of auditing plus the detecting that many IND's and NDA's call for. There are dozens of other pieces of useful work that such statisticians can (and do) perform in the Bureau, but these do not require the training of the group described in the paragraphs above. A majority of this group are already on the job.

Listing of these problems requiring mature biostatistical contributions should not be taken to imply that no start has been made. There have in fact been large improvements in the last year. The Bureau is fortunate in its present biometric head (Dr. Charles Anello). Lengthy discussions with Dr. Anello have not revealed any major differences of opinion as to objectives or means of accomplishing them. Indeed this section of our report may be read as reflecting our thoughts on how best to aid him.