Computer Analysis of Epidemiologic Data on Effect of Drugs on Hospital Patients

DENNIS SLONE, M.B., LEONARD F. GAETANO, B.S., LESLIE LIPWORTH, M.D., SAMUEL SHAPIRO, M.D., GEORGE PARKER LEWIS, M.D., and HERSHEL JICK, M.D.

DRUG EFFECTS in man have been of increasing concern in recent years to both public health and government agencies. As a consequence of a burgeoning technology in the production and development of potent drugs and an ever increasing role by government regulatory agencies, information is needed if we are to make informed decisions regarding their safe use. It has become increasingly apparent that much of the data required must be obtained by prospective epidemiologic techniques. Until recently the administrative framework did not exist for obtaining such information; nor was any system easily available for its storage, analysis, and utilization. The availability of a system capable of search and assembly procedures to deal with vast masses of information has made it possible to conduct widespread surveys(*1*).

Answers to the following important questions are necessary in order to provide a sound basis for the clinical use of drugs.

1. In what percentage of cases do adverse reactions occur with each individual drug?
2. What is the nature of these reactions?
3. Do the beneficial effects of a given drug justify its use in view of its hazards?
4. What specific populations are more likely to develop adverse or beneficial effects from a given drug?
5. Which drug effects are related to interaction with another drug or group of drugs?

To provide these answers, information must be obtained on the total population exposed to each drug, including the dose and the duration of exposure. An estimate of the "true" incidence of adverse reactions in the population under study is required. This requirement calls for the identification of all reactions together with some estimate of whether the implicated drug actually produced the reaction. The consequences of the reaction must also be noted. Such information is of only limited value unless it is related to the usual vital statistics of the exposed population (age, sex, race, and so forth), the diseases from which the patients are suffering, and a full list of all other drugs they are taking.

The authors, with the exception of Dr. Lipworth, are in the division of clinical pharmacology at the Lemuel Shattuck Hospital, Boston; Dr. Slone is associate director, Dr. Jick is director of the division, and Mr. Gaetano is a systems programer. Dr. Slone, Dr. Shapiro, Dr. Lewis, and Dr. Jick are assistant professors of medicine, Tufts University School of Medicine, Boston. Dr. Lipworth is director of the statistical unit, bureau of chronic disease control, Massachusetts Department of Public Health.

The investigation was supported by a training grant from the Public Heatlh Service (HE-5616), a contract from the National Institute of General Medical Sciences, Public Health Service (PH-43-67-1485), and a grant from the Pharmaceutical Manufacturers Association Foundation.

Finally, an estimate of the efficacy of the drug is also required, since the evaluation of any drug depends on the interplay between the disease and the potential risks and benefits. Gross distortion may emerge from incomplete, inadequate, or inaccurate data.

Since July 1966, the division of clinical pharmacology at the Lemuel Shattuck Hospital and Tufts University School of Medicine has operated a comprehensive drug surveillance system. Within a defined population, detailed epidemiologic information, as just outlined, has been collected on all drugs given to all patients. This system is currently operating in selected wards of five Boston hospitals, namely, Lemuel Shattuck Hospital (chronic medical), Peter Bent Brigham Hospital (acute medical), Boston City Hospital (acute medical), Boston Floating Hospital (pediatric), and Boston State Hospital (psychiatric). Thus far we have analyzed the data from the first three hospitals only.

In this surveillance system, trained nurses, referred to as "monitors," collect the information. Their source for this information is primarily the physician responsible for patient's care. In effect, a new member is added to the basic ward team whose primary role is the acquisition of accurate data. This system differs in one important respect from most other systems used previously to survey drugs. The onus for recording the data is removed from the physician himself and is made the direct responsibility of the additional member of the ward team. The reliability of the record is thus improved because the physician is relieved of the burden of extra paperwork—work which is not directly concerned with his primary role of providing medical care.

Each nurse monitor is assigned to one medical ward of about 20 to 30 patients, which is usually under the care of two physicians. When a patient is admitted, the nurse monitor fills out a vital statistics sheet, on which information is entered on the patient's age, sex, birthplace, relevant past history, and family history. The nurse performs simple genetic tests, such as the evaluation of color blindness and the ability to taste phenylthiourea. A blood sample is obtained from each patient, and by means of a series of hematological and biochemical tests each patient is classified into one of several subgroups. Following is a list of the tests which have been performed since January 1968 on each subject from whom a blood sample could be obtained. (The results provide a genetic profile of the patient).

1. Agglutinogen systems
   - ABO and subgroups
   - Rh and subgroups
   - Kell
   - Lewis
   - MNSs
   - Duffy
   - Kidd
   - Xg
   - Secretor status
2. Serum protein systems
   - Haptoglobin phenotype
   - Haptoglobin titer
   - Total serum protein
   - Serum albumin
3. Miscellaneous
   - Color blindness
   - Ability to taste phenyl-thiourea

For each drug ordered by the physician, the nurse monitor fills out a separate starting sheet, which includes the name of the drug, dose, frequency, route of administration, and the specific therapeutic indication. She keeps the sheet until the drug is stopped, whereupon a "drug stopping" section on the sheet is completed. The date and specific reason for stopping the medication are recorded. At the same time an efficacy rating on the drug is obtained from the physician and, where relevant, from the pateient. Among the 16 reasons for which a drug might be stopped, the two most important ones are the occurrence of an adverse reaction and a judgment that the drug was "not effective." Any side effect which has occurred during administration of the drug, even if the drug is continued, is noted.

Information on the occurrence of any suspected adverse reaction is obtained within 24 hours after the event. (An adverse reaction is defined as an unwanted or unintended result of having received any drug.) Once reported, the alleged reaction is investigated by members of the clinical pharmacology group. The likelihood of the reaction's being due to the drug is assessed as "definite," "probable," "doubtful," or "unknown," and the nature of the reaction is characterized in greater detail.

This system allows for the introduction of controlled drug trials, and a series of randomly

assigned and double blind comparisons between different hypnotics, as well as between different analgesics, have been performed within the framework of the overall drug surveillance program. In brief, when a physician has no strong preference for a specific compound for a patient, he orders—for example, a sleeping medication—by category. The nurse monitor is then responsible for seeing that the appropriately coded supply of medication is made available and given to the patient. This drug order is monitored just as for any other patient and is incorporated into the overall data bank. At the end of a particular study, a code is introduced into a computer program, and the data are analyzed according to previously set questions.

In controlling the collection and processing of such large amounts of data, several difficulties have arisen which have provided an exercise in transfer of data to the computer. The purpose of our paper is to detail the techniques used to handle these data and to describe the computer programs we have thus far developed.

Input Phase of Computer Processing

Before analysis by computers, the data sheets follow a process of keypunching, storage on magnetic tape, and submission to computer programs. The sheets are thoroughly checked to ensure accuracy and completeness. After verification, they are sent to a keypunching section for transcription to punched cards. With a few exceptions the forms which the nurse monitor uses have been designed to be self-coding. For example, among side effects, the nurse monitor might check a box labeled "nausea." For the keypuncher's benefit, this box is labeled "111" (the correct computer code for nausea). Transcription errors are thus minimized.

Keypunching of the sheets results in the following three sets of data:

Vital statistics file. Cards in the vital statistics file link each patient's hospital number with his age, sex, race, and so forth, dates of admission and discharge, discharge diagnosis, and laboratory data on biochemical genetic markers.

Drug administration file. Cards in the drug administration file link each patient's hospital number with coded drug numbers, drug dosage and frequency, dates on which drugs were started and stopped, indications for starting and stopping, and occurrences and kinds of adverse effects.

About once a month the accumulated card decks for the three files are added to the corresponding magnetic tape files by means of a standard computer card-to-tape software package and a standard software package. The cumulative files of vital statistics, drug administrations, and adverse reactions are inputs to the computer programs.

Current Computer Programs

Four programs currently operative are a drug profile program, a patient profile program, a "true" incidence adverse reaction program, and a drug interaction program.

Drug profile program. The drug profile program tabulates for each drug the number of patients exposed (denominator), the number of patients with adverse reactions (numerator), the reasons for starting and stopping the drug, and the efficacy ratings of the drug given by physicians and—where relevant—by patients. In addition, the distributions of certain epidemiologic variables, such as sex, age, race, and blood group, in the total sample exposed to the drug, as well as the corresponding distributions in certain subsamples, are tabulated and compared. These subsamples consist of patients with adverse reactions, patients with "good" ratings for drug efficacy, and finally those with "poor" ratings for drug efficacy.

Patient profile program. In the patient profile program, a medical information sheet is produced on each patient, which lists in particular the vital statistics, some selected clinical and laboratory data, the discharge diagnoses (up to four per patient), and all drugs received by the patient during each hospital admission, together with any side effects that may have occurred.

"True" incidence program. In the "true" incidence program, data on adverse reactions to each drug are tabulated. Reactions thought to be "doubtfully" drug related are eliminated after thorough clinical investigation. The "definite" and "probable" reactions are expressed as a ratio of all investigated reactions. The assumption is then made that the same ratio obtains in those patients who have not been inves-

Figure 1. Output for drug profile program, with codeine as an example

07/23/68

DRUG NO. 903 CODEINE

PATIENTS WITH SIDE EFFECTS

11149	13126	14318	16537	16786
16890	16940	17029	17563	17735
17746	17757	17899	17990	18481
18588	18671	19075	19313	19691
20085	20217	90299	93229	

PEOPLE EXPOSED= 355 NUMBER WITH SIDE EFF= 24

PERCENT WITH SIDE EFFECT= 6.8

REASONS FOR STARTING

PAIN	307
HEADAC	8
COUGH	15
PREOP	3
OTHER	6
VARIAB	16

REASONS FOR STOPPING

SIDE EFF	NOT EFF	OTHER	VARIAB
12	22	229	92

PERCENT NOT EFFECTIVE= 6.2

DOSAGE DATA	VARIAB	OTHER	30.0 MG	1.0 GRAIN	32.0 MG	60.0 MG	15.0 MG
AFFECTED	7	0	14	1	0	1	0
TOTAL-AFF	51	11	190	4	20	34	12

EFFICACY RATING

	GOOD	FAIR	POOR	UNDEC	VAR
DOCTOR	136	35	35	75	74
PATIEN	92	29	27	137	69

PERCENTAGE CALCULATION

	GOOD/TOTAL	GOOD/DECID	POOR/TOTAL	POOR/DECID
DOCTOR	38.3	66.0	9.9	17.0
PATIENT	26.0	62.2	7.6	18.2

EPIDEMIOLOGY

NO. MALES= 144 FEMALES= 210 WHITES= 323 NON-WHITES= 31

	UNAFFECTED	W/SIDE EFFCT	ST ERR	
AVRAGE AGE	55.0	61.3	3.2	SIGNIF
AVRAGE WEIGHT	143.9	143.5	9.0	NOT SIGNIF

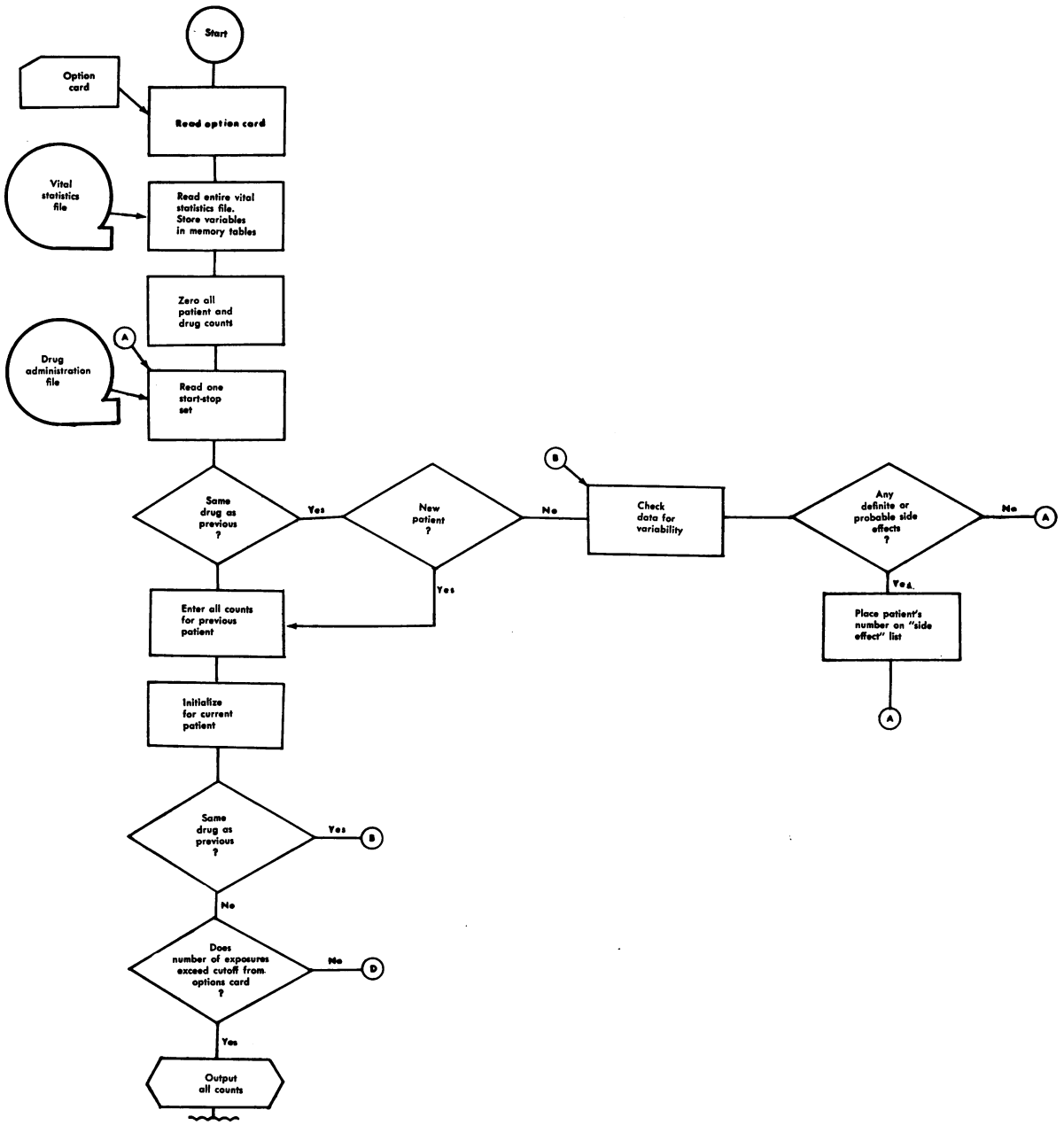
Figure 1. Output for drug profile program, with codeine as an example—Continued

```
                          MALE                FEMALE
PCT. W/SIDE EFFCT          5.6                  7.6  CHI=        0.3  NOT SIGNIF AT .05

                         WHITE             NON- WHITE
PCT. W/SIDE EFFCT          7.4                  0.0  CHI=        1.4  NOT SIGNIF AT .05

BLOOD GROUP          A          B         AB          O
SIDE EFF             6          5          0         11
TOTALS             118         39         13        145
PERCNT             5.1       12.8        0.0        7.6

CHI SQU=     3.76   NOT SIGNIF 3 DEG OF FREEDOM
```

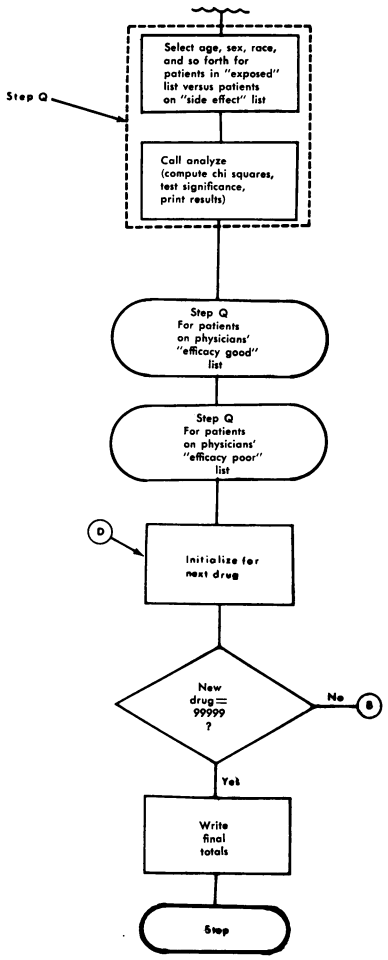
```
                 DR NOT  GOOD          DOC RT  GOOD
AVRAGE AGE          56.3                 53.9              1.6  NOT SIGNIF
AVRAGE WEIGHT      144.5                142.8              4.9  NOT SIGNIF

                          MALE                FEMALE
PCT. DOC RT  GOOD         28.5                 44.8  CHI=        8.9       SIGNIF AT .05

                         WHITE             NON- WHITE
PCT. DOC RT  GOOD         37.2                 48.4  CHI=        1.1  NOT SIGNIF AT .05

BLOOD GROUP          A          B         AB          O
GOOD                51         12          5         53
TOTALS             118         39         13        145
PERCNT            43.2       30.8       38.5       36.6

CHI SQU=     2.33   NOT SIGNIF 3 DEG OF FREEDOM
```

```
                 DR NOT  POOR          DOC RT  POOR
AVRAGE AGE          55.6                 54.0              2.7  NOT SIGNIF
AVRAGE WEIGHT      144.4                139.3              8.1  NOT SIGNIF

                          MALE                FEMALE
PCT. DOC RT  POOR         12.5                 8.1  CHI=        1.4  NOT SIGNIF AT .05

                         WHITE             NON- WHITE
PCT. DOC RT  POOR         10.2                 6.5  CHI=        0.1  NOT SIGNIF AT .05

BLOOD GROUP          A          B         AB          O
POOR                 6          6          3         16
TOTALS             118         39         13        145
PERCNT             5.1       15.4       23.1       11.0

CHI SQU=     7.16   NOT SIGNIF 3 DEG OF FREEDOM
```

Figure 2. Logic diagram for



drug profile program



tigated or in whom a cause-effect relationship was either uncertain or unknown. The "true" incidence of adverse reactions, thus, is as follows:

$$\frac{\text{Patients with "definite"+"probable" reactions}}{\text{patients with "definite"+"probable"+"doubtful"}} \times \frac{\text{total patients with alleged reactions}}{\text{total patients exposed}} \times 100.$$

The same adjustments are applied to a breakdown of "true" incidence, first by whether the reactions induced the physician to stop the drug ("drug discontinued") or to continue it ("drug continued") and then, within each of these categories, by whether the reactions were transient (lasted less than 1 day) or were prolonged; were major (either caused severe clinical disease, threatened life, or resulted in prolonged hospital treatment), or were minor. Finally, under the headings "transient-major," "transient-minor," "prolonged-major," and "prolonged-minor," the exact reactions which occurred are tabulated. Thus a physician can obtain an estimate of the predicted risk attendant upon use of each drug.

Drug interaction program. The drug interaction program is intended to identify drugs which are associated more frequently than would occur by chance with (*a*) all adverse reactions attributed to any specific drug, (*b*) "good" physicians' ratings for any specific drug, and (*c*) specific adverse reactions (for example, nausea or rash).

Drug Profile Program

Input

Vital statistics file, sequenced by patient's number

Drug administration file, sequenced according to drug number and, within drug number, by patient's number

Single pass program. Current running time—about 12 minutes for 3,000 patients, 800 drugs

Output (fig. 1)

For each drug, a page is output giving the following information:

1. Drug code number and name
2. Number of patients receiving one or more doses
3. Identification numbers and the number and percent of patients having a side effect, "doubtfuls" having been removed
4. Breakdown of reasons for starting the drug
5. Breakdown of reasons for stopping the drug, together with the percentage of patients for whom the drug was stopped because it was ineffective
6. Breakdown of dosage data, the population being divided into those patients with and without a reaction
7. Breakdown of the physicians' and the patients' efficacy ratings for the drug. The "good" and "poor" ratings are expressed as (*a*) a percentage of all the ratings given, including those in the category of "variable" and "undecided," and (*b*) a percentage of the ratings representing a decision, namely $\frac{\text{"good" or "poor"}}{\text{"good"+"fair"+"poor"}}$. Patients who receive any given drug more than once may have different starting and stopping

Figure 3. Output for patient profile program

```
    THE CODE 9 ALONE OR IN A SERIES = TO UNKNOWN

HOSP. NC.   16940                        H          SOC. SEC. NO.

HOSPITAL  L.S.H.   WARD  7S

MALE       CAUCAS       BORN IN U.S.      DATE OF BIRTH 12-30-88   AGE   78

HGT.   65 IN.  WGT.   999 LBS.

RADIOTHERAPY  YES

DATE OF ADMISSION   9-28-67

DATE OF DISCH. FROM STUDY  10-13-67      NO. DAYS IN STUDY   15

DATE OF HOSP. DISCH. 10-13-67      NO. DAYS IN HOSP.   15

RESULT OF PRESENT ADMISSION HOME

MEDICAL DATA

               BLOOD GROUP B         R H FACTOR  POS.

                   HGB.   14.0            DIFF. WBC.
                   HCT.  43                P-74    B- 0
                   WBC.   12000            L-18    E- 1
                                           M- 6
                   BUN.  18

               URINE
                   S.G. 1.018     PROT. 0.0      SED. 1
                   PH.   6.0      SUG.  0

DISCHARGE DIAGNOSIS   1. 163  MALIGNANT NEOPLASM OF LUNG,USP
                      2.
                      3.
                      4.
```

DRUG	FORM	DOSE	UNIT	FRQ	RT.	INS	IND.	START DT
CODEINE	SOL	30.0	MG	6.0	PO.	PRN	PAIN	9-28-67
METHADONE	SOL	5.0	MG	6.0	PO.	PRN	PAIN	9-28-67
ASPIRIN	SOL	600.0	MG	6.0	PO.	PRN	PAIN	9-28-67
DIGOXIN	SOL	0.3	MG	1.0	PO.	STD	C.H.F.	9-28-67
ISUPREL	LIQ	5.0	DROP	4.0	INH	STD	BRONCH	9-28-67
MILK OF MAGN	LIQ	30.0	CC	1.0	PO.	PRN	DIARRH	9-28-67
CYTOXAN	LIQ	930.0	MG	1.0	IV.	STA	METTUM	9-29-67
KCL	SOL	20.0	MEQ	2.0	PO.	STD	ELEDIS	9-29-67
ROBITUSSIN	LIQ	10.0	CC	3.0	PO.	STD	COUGH	9-29-67
QUIBRON	SOL	1.0	TABLET	4.0	PO.	STD	BRONCH	10- 2-67
MUCOMIST	LIQ	2.0	CC	4.0	INH	STD	BRONCH	10- 2-67
NEMBUTAL	SOL	100.0	MG	1.0	PO.	PRN	INSOMN	10- 3-67
LIBRIUM	SOL	10.0	MG	3.0	PO.	STD	ANXIET	10- 5-67
METHADONE	SOL	10.0	MG	6.0	PO.	PRN	PAIN	10-10-67
GELUSIL	LIQ	30.0	CC	***	PO.	PRN	GASTRI	10-10-67
PHENERGAN	SOL	25.0	MG	6.0	PO.	PRN	PAIN	10-11-67

indications, dosage schedules, and drug efficacy ratings. These differences are then counted as "variable" in each of these categories

8. Breakdown of the patients into the following groups:

a. Male and female, white and nonwhite, with incidence of side effects in each group, along with chi square values and their statistical significance

b. Average ages of the two populations and associated standard error

c. Average weights of the two populations and associated standard error

d. Number of patients in blood groups A, B, AB, and O with and without side effects, along with associated chi square value and its significance

9. Same procedure as for 8, but this time the population is divided into those patients whose physicians rated the efficacy of the drug as "good" and into the remainder

10. Same as for 8 except that population is divided into the patients whose physicians rated the efficacy of the drug as "poor" and into the remainder

Logic (fig. 2)

Start after initialization. The entire vital statistics file is read, and variables such as age, sex, race, and so forth are stored in the memory tables for each patient, along with his number. All counts are zeroed.

A. One item from the drug administration file is read. (An "item" is the start to stop administration of one drug to one patient. A patient may have several items for the same drug with varying doses.) If this drug is the same as the one previously recorded, check the patient's number. If the patient's number is the same as his previous number, go to B.

B. Check all the patient's data for variability, and if there was a "definite" or "probable" side effect from this drug, enter the patient's number in the side effect list and go to A. When the patient's number changes, enter the counts for the previous patient in the various lists; initialize for the current patient. Again check the drug versus the previous drugs prescribed for the patient. If the same, repeat. If the drug is different, test the number of exposures to see whether it exceeds the print cutoff figure. If not, go to D. Otherwise, print all summary totals for the previous drug.

Then, for each patient on the lists of patients exposed, of patients affected (those with adverse reactions), of patients whose physicians rated efficacy of drug "good," and of patients whose physicans rated efficacy of drug "poor," select the age, race, sex, and weight and perform standard error and chi square calculations. As an example, a 4 by 2 cell for blood groups A, B, AB, and O for those on the "affected" list versus those on the "exposed but not affected" list is generated and printed; a chi square is also computed and printed and its significance tested and printed.

Figure 3. Output for patient profile program—Continued

STOP DT	STOP REAS	TOXICITY	MD EFF	PAT EFF	TOT DOSE	NO.EP
10-10-67	NO EFF	CONSTI	POOR	POOR	630.0	21
10-10-67	CHANGE		POOR	POOR	105.0	21
10-13-67	DS W/M		GOOD	GOOD	12600.0	21
10-13-67	DS W/M		GOOD	UNKWN	3.8	15
10-11-67	EFF.		GOOD	GOOD	260.0	52
10-13-67	DS W/O		UNKWN	GOOD	180.0	6
10-13-67	STAT.		FAIR	UNKWN	1860.0	2
10- 4-67	EFF.		GOOD	UNKWN	200.0	10
10-13-67	DS W/M		GOOD	GOOD	420.0	42
10-13-67	DS W/M		UNKWN	GOOD	44.0	44
10-11-67	EFF.		GOOD	GOOD	72.0	36
10-13-67	DS W/M		FAIR	GOOD	1200.0	12
10-13-67	DS W/M		GOOD	GOOD	240.0	24
10-13-67	DS W/M	NAUSEA	GOOD	GOOD	110.0	11
10-13-67	DS W/O		GOOD	GOOD	60.0	2
10-13-67	DS W/O		UNKWN	GOOD	150.0	6

D. Initialize all counts to zero. If the next drug is 99999, stop. If not, go to B.

Q. In procedure Q, data on patients are analyzed and printed for each of the categories with side effects—those on the physicians' "efficacy good" list and those on the physicians' "efficacy poor" list.

Patient Profile Program

Input

Vital statistics file, sequenced by patient's number and then by admission date

Drug administration file, sequenced according to patient's number and then by admission date

Single pass program. Current running time—1.2 seconds per patient

Output (fig. 3)

For each admission of a patient, a page of print gives the patient's number, name, social security number, hospital, ward, race, place and date of birth, age, height, weight and surface area, marital status, number of pregnancies, selected clinical and laboratory data, and discharge diagnoses. The bottom half of the printout lists all the drugs administered to the patient during one hospital admission. For each drug, the printout gives the name, form of the drug, dose and unit, frequency of administration, route and instruction, starting indication and reason for stopping, dates for starting and stopping, side effect—if any, efficacy ratings of the drug by physician and patients, and the total dosage and number of administrations. For example, the printout in figure 3 concerns a patient who received codeine in solid form, 30 mg., six times a day by mouth, by a p.r.n. (as circumstances may require) order, prescribed for pain from September 28 to October 10, 1967. The drug was stopped because it was judged to be ineffective. Constipation was noted as an adverse effect. Both physician and patient rated drug's efficacy as "poor." The total dose received was 630.0 mg.; the drug was given 21 times.

Logic

Start. Read option cards and initialize program.

A. Read one set of data from the vital statistics file, check patient's number for all nines, which indicates end of job. If all nines, output the variables, such as name, age, and so forth. This information constitutes the top half of the output page.

B. Check the drug buffer. If empty, read one set of data from the drug administration file. Compare the patient's ID (identification number) on the set from the vital statistics file with the number on the set from the drug administration file. If patient's ID in the vital statistics file is the greater, purge the buffer and go to B (that is, skipping drug cards until the right patient is found). If the two IDs are equal, check whether the dates in the drug administration file fall within this hospital admission. If yes, print out, with the spinoff tape if requested, the drug data and purge the buffer; then repeat. If no, go to A. If the patient's ID in the vital statistics file is neither greater nor equal to the ID in the drug administration file, it is less than the number in the drug administration file; therefore go to A (that is, if there are no more drug cards for this patient, get the next patient's cards, leaving the drug buffer full).

True Incidence Program

Input

Data card, with total number of patients exposed from the drug profile program

Adverse reaction file, sequenced by drug number

Single pass program. Current running time—0.25 seconds per drug

Output (fig. 4)

For each drug, the output consists of the following: total number of patients exposed; "true incidence" of adverse reactions; under the two categories of "drug continued" and "drug discontinued," the respective percentages of reactions that were "transient-major," "transient-minor," "prolonged-major," and "prolonged-minor," along with the specific side effects in these categories; and finally, the mortality rate for the drug. All percentages are adjusted as explained in the Logic section. The percentages expressed in this program represent the overall estimated risks with the use of the

Figure 4. Output for "true incidence" program

```
          DRUG-- CODEINE                                                    07/26/68

TOTAL EXPOSED  =    355

                                                                TRANSIENT=LESS THAN 1 DAY
                                                                PROLONGED=MORE THAN 1 DAY

'TRUE' INCIDENCE ADVERSE REACTIONS=     5.4 PCT.

                                                                DEFINITE+PROBABLE    ALLGD ADV REACT
                                                  'TRUE' INCIDENCE= ---------------- . ---------------- . 100
                                                                DEF+PROB+DOUBTFUL    TOT INDIV EXPSD

                                                  ALL PERCENTAGES REPRESENT ADJUSTED OVERALL INCIDENCE


              DRUG DISCONTINUED  3.3                                DRUG CONTINUED  2.2

TRANSIENT-MINOR  0.5                          TRANSIENT-MINOR  0.0

                    ITCHIN  0.5                                          0.0

TRANSIENT-MAJOR  0.0                          TRANSIENT-MAJOR  0.5

                            0.0                                    OTHER  0.5

PROLONGED-MINOR  1.6                          PROLONGED-MINOR  1.6

                    CONSTI  1.6                                    CONSTI  1.6

PROLONGED-MAJOR  1.1                          PROLONGED-MAJOR  0.0

                    NAUSEA  0.5                                          0.0
                    HALLUC  0.5                                          0.0



MORTALITY                   0.0


SECONDARY REACTIONS
VERTIG     1.
VOMITI     1.
HPOTNS     1.
```

drug. Both denominator and numerator figures are used.

Logic

After initializing the totals, and so forth, to zero, read one data card containing the desired drug number and the total number of patients exposed to that drug from the drug profile program.

A. From the "adverse reaction" file, read one set of data yielding the current drug number (CDN). If CDN is not the same as the desired number and is less than it, go back to A; if it is greater, check the current patient's ID. If the number is the same as the previous ID, go back to A, if not, add 1 to the total reactions count. Then check the "investigator's impression" field. If the investigator's impression was equal to doubtful—If=to doubtful—that is, if he did not believe the reaction was due to the drug, go to the beginning of A. If not equal to doubtful and if equal to unknown, that is, if the investigator did not know or the datum was otherwise unknown, add 1 to the unknown total. Otherwise, add 1 to the definite total. Then check whether the drug was continued. If yes, enter the subsequent totals in the continued column; if no, in the discontinued column; then go to the beginning of A.

C. Compute the correction factor which equals

$$\frac{\text{Patients with "definite" + "probable" reactions}}{\text{patients with "definite" + "probable" + "doubtful"}} \times \frac{\text{total patients with alleged reactions}}{\text{total patients exposed}} \times 100.$$

Adjust all percentages with this correction, output the results, and go to the start.

Drug Interaction Program

Input

The input is obtained by first running the patient profile program and obtaining an intermediate or spinoff tape. For each patient, the tape contains a record of all drugs he received during each hospital admission. Since the interaction program is multipass, this condensed tape provides an efficient input.

Multiple pass program. Current running time—about 3.0 minutes per pass.

Output

Each page of output presents the name of the drug (subject drug to which a side effect was attributed, followed by a list of drugs) the associated drugs that were given along with the subject drug. For each associated drug, the number and percent of people receiving both drugs and experiencing no side effect from the subject drug and, similarly, the number and percent experiencing a side effect are recorded. Significant associations with p values at the 0.01, 0.001, and 0.0001 levels for each associated drug are printed.

The total number of patients in the two categories "side effect" and "no side effect" is given at the end of the page. Similar output is obtained for physicians' ratings of "good" drug efficacy versus the remaining patients and for drugs given when specified adverse reactions did, and did not, occur.

Logic

Start. Read one data card giving the number and name of subject drug.

A. Read one input record from the patient profile program spinoff tape. (One logical item contains all the drugs a patient received within a given hospital admission.) If the patient's number is the same as his previous one, check the side switch. If the switch is on, go to A; if off, go to A1 (see next paragraph). If the patient's number is different from the previous one, check the side switch. If it is on, turn it off and go to B. If it is off, for each drug in the list formed in step A2 (see paragraph after next), add 1 to the corresponding element in the overall drug matrix in the "no side effect" column and 1 to the "without" total and go to B.

A1. Check the logical items for the presence of subject drug. If no (not present), go to A. If yes (present), did the patient have a side effect from the subject drug? If yes, set the side switch and proceed through the logical item, counting 1 in the "side effect" column of the overall drug matrix for each drug the patient received before onset of the reaction (these are the associated drugs); add 1 to the total "with" count and go to A. If no (no side effect from subject drug), check to see if current hospitalization is the patient's first admission.

A2. If current hospitalization is the patient's first, make a list of all drugs he has received, then go to A. If not his first (a list has already been made), go to A.

B. If the patient's ID is not all nines, go to A1. If all nines, rewind the data tape; print the title with the number and name of subject drug for each column in the overall drug matrix; compute for each drug versus the total for the column the percent of patients with and without side effects. For each row in the matrix whose sum is greater than 25 exposures, perform a chi square calculation on the total number of patients with and without side effects. If the difference is significant at the 0.01 level, output a line, noting also whether the difference is at the 0.001 and 0.0001 levels. Go to the start.

Results and Discussion

As of June 12, 1968, a total of 2,514 patients had been entered and discharged from the drug surveillance system and 26,102 drug exposures had been monitored. Thus, on the average, each patient had received 10.4 drugs. Seven hundred and seventy-eight patients (30.9 percent) experienced adverse effects which, after evaluation, were judged to be "probably" or "definitely" due to drugs.

More than 700 drugs were monitored, and it is to be anticipated—especially for drugs which are not given very often—that considerable time must elapse before sufficient data on any one drug will reflect the population universe from which the sample was drawn.

Ultimately the value of such a surveillance system depends on its usefulness to the public and to the medical community. It is therefore encouraging to find that information of practical value has already been obtained. Borda and associates (*2*) found, after thorough investigation, that in 69 percent of all patients with possible side effects, the suspicion of a drug reaction was probably valid. In 22 percent of the patients with possible side effects, drug interaction was thought to have played a role. This result supports the hypothesis that multiple drug treatment increases the likelihood of side effects. In only 20 percent of the patients in whom drug interaction was suspected, were these reactions

considered minor; in 80 percent they were considered either major or moderate. This result represents an appreciable incidence of morbidity from drug therapy. Thus the risk of hospitalized medical patients in selected wards of experiencing drug reactions which are more than trivial is about 25 percent—80 percent of 30.9 percent.

[A reaction is considered moderate or major by virtue of one or more of the following criteria: (*a*) clinical severity and morbidity—a purely subjective judgment, (*b*) occurrence of serious sequelae, (*c*) a threat to life, (*d*) resultant increase in hospitalization, and (*e*) death. For example, transient nausea that occurred after a patient took a sleeping tablet would be considered a minor reaction. Massive gastrointestinal hemorrhage with resultant shock—all attributed to heparin and necessitating blood transfusion and other measures—would obviously be a major reaction.]

The foregoing remarks represent general conclusions on the adverse effects of drugs. More specific observations about certain drugs have also emerged. Among 97 consecutive patients given heparin (*3*), bleeding complications occurred in 18 of 56 women (32 percent) and in 6 of 41 men (15 percent). The average age of the 24 patients with reactions was 66 years, compared with 55 years for the remainder. The incidence of toxicity in women over 60 was 50 percent; in men over 60, 19 percent; in women under 60, 14 percent; and in men under 60, 10 percent ($P<0.03$). The efficacy of the drug was the same regardless of age or sex. These results indicate the need for caution in giving heparin to elderly women.

Of 441 consecutive patients who received digoxin, it was judged to have been effective in 91.7 percent and poor in 2.8 percent. The overall incidence of toxicity attributed to the drug was 18.4 percent, and the "true" incidence of toxicity was calculated at 16.8 percent. The mean body weight of the patients with adverse reactions was 159.8 pounds and of the remainder, 145.9 pounds. This difference was significant (S.E. 1.3, $P<0.05$). The mean age of the patients with adverse reactions was 66.2 years and of the remainder, 64.4 years (S.E. 16); the difference was not significant. Toxicity due to digoxin was found to be associated with meperidine, morphine, heparin, hydrochlorothiazide ($P<0.05$), furosemide, aminophylline ($P<0.01$), and prochlorperazine ($P< 0.001$). No racial or sexual differences were apparent, and the mean daily dosage did not affect the incidence of adverse reactions. The efficacy of digoxin was similiar in all groups.

These results show that digoxin is both dangerous and highly effective; that heavier patients (presumably because they are edematous and receive diuretics) are at greater risk than lighter patients; that the drug is not contraindicated specifically because of age; that factors other than dosage can also be responsible for toxic effects; and that certain drugs given at the same time seem to be associated with an increased risk of toxicity. The latter point needs to be carefully interpreted since associations do not necessarily imply cause and effect relationships. Prochlorperazine, for example, is undoubtedly associated with an increased incidence of toxicity because it is used in the treatment of nausea attributed to digoxin. Other associations could be due to chance and require confirmation.

Data on drugs such as heparin and digoxin accumulate fairly rapidly so that we were able to carry out initial epidemiologic analyses less than 2 years after the drug surveillance system was started. Suggestive associations are also emerging for many other drugs, but the data base is still too small.

Do efficacy ratings have measurable value? Partly to test this question and partly to formulate a new method for assessing the efficacy of hypnotics, a randomized, double blind study comparing pentobarbital 100 mg., chloral betaine 750 mg., diphenhydramine 50 mg., and placebo (calcium lactate) was introduced into the monitored ward. Essentially, the physician who had no preference was asked to order the "hypnotic study drug" instead of any specific hypnotic agent; the drug was then monitored in the routine way. Three indices, namely, efficacy ratings given by physicians and by patients ("good," "fair," "poor," or "don't know") and the number of times each of the drugs was stopped because it was considered ineffective were analyzed after the study had been completed and the code broken. There were 100 orders for the placebo and 50 for each of the

drugs. The results showed that the method is sensitive in discriminating between active hypnotics and a placebo. Apparently judgments by the physician (and where appropriate, by the patient) in evaluating drugs in routine clinical medicine have real meaning and can be used in this drug surveillance system for quantitative analyses when population samples are comparable.

Development of a computer facility to handle the system outlined here is complicated. The major difficulty is the design, development, and refinement of programs. Once this task has been completed, increasingly large increments of data can be handled by the same programs without a corresponding increase in complexity.

At present, only medical patients in hospital wards are monitored. The same methods, if necessary with appropriate modifications, could be used to survey ambulatory, obstetrical, and other populations.

Summary

In recent years the increase in the number of pharmacologically active agents has highlighted the need for epidemiologic information on their efficacy and toxicity. In the absence of reliable denominator and numerator data, however, the scope of the problem cannot be determined. A comprehensive surveillance program designed to obtain such information on hospitalized medical patients was introduced into selected wards of five Boston, Mass., hospitals in July 1966 by the division of clinical pharmacology of the Lemuel Shattuck Hospital and Tufts University School of Medicine.

Extensive information on the vital statistics of individual patients and complete records of all medications they received, together with estimates of efficacy and toxicity, are collected in order to describe subpopulations which react either adversely, beneficially, or negatively to different drugs. The conventional description of individual patients is further extended by use of a biochemical genetic profile. Data are analyzed with the aid of a computer facility.

The surveillance system permits weighing the adverse and beneficial effects for any given drug. Drug interaction can be studied, and specific drug effects associated with genetic characteristics can be demonstrated. The system has provided useful, practical information. It is hoped such information will assist in the development of more rational therapeutics.

REFERENCES

(*1*) Slone, D., et al.: Drug surveillance utilising nurse monitors. An epidemiologic approach. Lancet 7469: 901–903, Oct. 22, 1966.

(*2*) Borda, I., Slone, D., and Jick, H.: Assessment of adverse reactions within a drug surveillance program. JAMA 205: 645–647, Aug. 26, 1968.

(*3*) Jick, H., Slone, D., Borda, I., and Shapiro, S.: The influence of age and sex on the effects of heparin. New Eng J Med 279: 284–286, Aug. 8, 1968.