Electronic Maintenance of Case Registers

WILLIAM PHILLIPS, Jr., KURT GORWITZ, M.A., and ANITA K. BAHN, Sc.D.

THE HANDLING of data with high-speed electronic equipment such as computers with a "memory" core has come to be referred to as "automatic data processing." Such methods are well accepted in industry and the larger research institutions, but are not yet fully exploited in public health and medical research. Vital statistics and other administrative health data are often required with sufficient frequency and detail to justify these newer methods even when the volume of data is small. Examples are birth and death statistics, hospital admissions and discharges, and rates, percentage distributions, life tables, and other computations based on these data. In addition, the efficiency and versatility of electronic computers enable the statistician to conduct many types of analyses of basic health data that would not have been practical before.

One type of computer application is to case registers of chronic disease. By a chronic disease we mean "one which persists over a long period of time, sometimes even for most of the individual's life. . . . It is the changeability and variation, not stability, that is in fact the dominant characteristic of most long-lived conditions" (*1*). In a register we attempt to study the dynamic characteristics of a chronic disease by systematically following identified cases over long periods of time. This followup, dynamic, or longitudinal aspect distinguishes the disease register from the disease index. The index is merely a file of diagnosed cases which have been identified at some time and do not reflect the current status of a disease.

A variety of chronic disease registers have been established for cancer, tuberculosis, and rheumatic fever (*2–8*). Registers differ according to (*a*) principal purpose—medical service, health supervision, research; (*b*) location—in the clinic, hospital, health department, biometrics office; (*c*) type of initial and followup data collected and manner of collection; and (*d*) extent to which all diseased individuals in a community or defined population have been identified. We shall discuss certain common features of these registers and illustrate the use of the computer for a psychiatric case register.

Problems

The two principal maintenance problems of registers are the matching of incoming case reports against previous records, to determine if this is a new or a previously reported case, and record updating. Matching by patient name or by clinic or hospital number has been either clerical, aided by visible record cards, or mechanical, through the sorting and collating of punchcards. The updating of records has been achieved by clerical posting to a register card or adding to a chronological file of punchcards or reports of actions for each individual. These maintenance problems are not insurmountable if the volume of cases and actions is not too large. If, however, additions to the register

Mr. Phillips is digital computer programer, and Dr. Bahn is chief, Outpatient Studies Section, Biometrics Branch, National Institute of Mental Health, Public Health Service. Mr. Gorwitz is chief statistician, Maryland State Department of Mental Hygiene, Baltimore, Md. *This paper was presented at the annual meeting of the American Statistical Association in New York City, December 28, 1961.*

greatly exceed removals due to death or for other reasons, the size of the register may quickly get out of hand.

A more difficult task is that of abstracting and summarizing the continuous records in order to answer questions such as: What is the health status of all registered individuals as of a specified time? What is the average length of time between recurrent episodes? What are the current age and residential characteristics of patients classified by previous register history? Such questions require examination of a number of recorded events for each individual, necessitating either clerical review of hand-posted or machine-listed records or complicated collating and sorting of punchcards. Furthermore, while the records are being used for analysis, they are not available for record updating. These problems have contributed to the limited research use of registers to date.

Maryland Psychiatric Case Register

How can the computer aid in the maintenance of case registers? Let us describe the Maryland Psychiatric Case Register as an illustration.

Background. Reports on admissions and discharges of Maryland residents for all of the inpatient and outpatient psychiatric facilities in the State and the District of Columbia are routinely collected, punched, and tabulated for administrative purposes. The type and number of facilities which report to the Maryland Psychiatric Case Register and the data reported are shown on page 507.

These statistics aid the mental health program planner by describing for each facility the diagnostic and demographic characteristics of patients seen and the services received during a time period (*9–12*). The reports, however, represent only single episodes in the psychiatric career of individuals. Without linkage of records of all episodes and services for the same person, administrative and epidemiologic research is limited. In the evaluation of mental health programs, the increased multiplicity of facilities serving the mentally ill in recent years, the shorter hospital stay, and higher rates of readmission make linking of records imperative for obtaining unduplicated counts of patients and for studying the movement of patients between facilities. This linkage is possible in Maryland since name, birth date, social security number, and other identification are routinely reported. A register protocol is being established therefore to provide for continuous or longitudinal statistical records for each resident of the State reported as a patient of a psychiatric facility (*13*).

Volume. Some idea of the magnitude of the task may be obtained from the number of actions that must be processed annually and from the estimated size of the register. Currently, 19,000 psychiatric admission and 18,000 discharge reports are received each year (see box); in addition, about 4,500 punchcards refer to significant changes in the status of the hospital patient.

With regard to the size of the register, the count of reported episodes for persons under psychiatric care during a year is 31,000, but it is estimated that because of multiple episodes for some individuals, this figure represents only about 20,000 persons. Similarly, it is estimated that only about 5,000 to 10,000 new psychiatric cases will be reported each year. Depletion of the register will result principally from migration out of the State and death. Except for the senile and the severely brain damaged, the rate of decrement due to associated causes of death, either in the hospital or outside, is small. Since one-half of clinic admissions are children, the register is likely to grow at a substantial rate, probably reaching 50,000 to 75,000 within 10 years. Since one purpose of the register is to study the lifetime aspects of mental illness and psychiatric care, cases will be permanently removed from the active cumulative register only upon death.

Matching of records. One of the most difficult tasks in establishing a register of this size is the elimination of duplicate registration of the same person and the association of data based on name and descriptive information only. Without an electronic computer this can be an exceedingly difficult job. The extreme speed of a computer in sorting and processing makes possible the use of numerous and varied methods to eliminate duplicates.

The primary common factor would of course be the name. It is planned to Soundex code

automatically the surname (and also the maiden name of women) with a slight variation from the standard Soundex system. This variation is to code the first letter as well as the rest of the surname. Every case in the register will be checked to every other case with similar Soundex code and compared within specified tolerances according to sex, color, date of birth, marital status, maiden name, and residence. Matches which are possible duplicates will be

Reports to Maryland Psychiatric Case Register

TYPE AND NUMBER OF FACILITIES

Outpatient Clinics		*Hospitals and Institutions*	
County health department clinics	31	State mental hospitals	6
State mental hospital clinics	5	Institutions for mental defectives	1
University of Maryland and Johns Hopkins University clinics	10	Homes for mental defectives	5
Miscellaneous auspices (Maryland and District of Columbia)	17	Private mental hospitals	14
Veterans Administration clinics	2	General hospitals accepting psychiatric patients	9
		Veterans Administration hospital	1
Total	65	Total	36

ITEMS REPORTED ON ADMISSION

Name (and maiden name), case number	Name, case number
Address	Address
Birth date, sex, color, marital status	Birth date, sex, color, marital status
State mental hospital status [1]	Education, religion
Date of admission	Date of admission
Previous admission to this facility	Previous admission to this facility
Referral source	Previous inpatient admissions anywhere
	Type of commitment
Social security number	Social security number
Diagnosis or impression	Diagnosis

ITEMS REPORTED ON DISCHARGE

Symptoms of problem drinking	Symptoms of problem drinking
Type of service received and condition after treatment	Dates and type of movement in and out of hospital other than discharge [2]
Number of interviews with or about patient, by profession, by month of service	
Disposition	Cause of death, if died in hospital
Dates of discharge and final interview	Date of discharge

NUMBER OF ACTIONS [3]

Admissions annually	10,000	Admissions annually	9,000
Discharges annually	9,000	Discharges annually (including 1,000 deaths in hospitals)	9,000
		Placed on leave	3,000
		Returns from leave	1,500
Average on rolls at any time	5,000	Average on books at any time	14,000
Under care during year	13,000	Under care during year	18,000

[1] Reported by State mental hospital outpatient clinics only.
[2] Reported by State mental hospitals only.
[3] Not unduplicated.

printed for clerical check to the source documents and resolution of conflicting information. In addition, records of all persons with the same date of birth or the same birth month and day within a few years of each other will be checked for other similarities, such as sex and race, and reviewed as possible duplicates. These procedures for person matching of records are still in the experimental stage.

We are also collecting social security account numbers as an additional means of searching for duplicates. Since a large proportion of the adult population has a social security number, use of this method of checking for duplication is warranted, even though we are not able to obtain the social security number of all patients who have one.

After name matching, each admission report will be automatically assigned a new register number or, if the case has been previously reported, an old register number. The name-matching process can be bypassed for the processing of discharge reports, because tape records maintained by facility and case number order will provide ready linkage to the patient's register number. Register numbers can also be readily assigned in this manner to readmission reports for those facilities using unit case numbers.

Magnetic Tape Files

The following records will be maintained on tape:

Master identity record. This record contains register number and patient identification data, plus code to indicate current psychiatric care. During matching, this file will be sorted by the computer in several ways: by Soundex code and name order, by birth date, by current address, and by social security number. It can also be sorted by register number for linkage with other files. A supplementary identity record will be ordered by maiden name or alias. Because of the confidential nature of the information in the register, the magnetic tapes are stored in locked cabinets. These tapes can be referred to only by use of a Honeywell 800 computer with a program especially written to analyze the data. No other records are kept in name order and all punchcards containing names are destroyed as soon as the names are transferred to tape.

Master statistical record of diagnoses and services. This record, identified only by register number, contains the patient's psychiatric data. It consists of three parts: (I) the patient's current status and cumulative record or count of selected events and services; (II) a résumé, for each year, of the patient's admissions, hospital days, and other services; and (III) a chronological detailed record of each event or episode of psychiatric care, with associated diagnosis and services.

Two auxiliary records are maintained:

Facility and patient case number record. This record aids in assigning register numbers to discharge reports and readmissions.

Chronological record of patient's residences. This record is used for reference purposes and special studies on mobility.

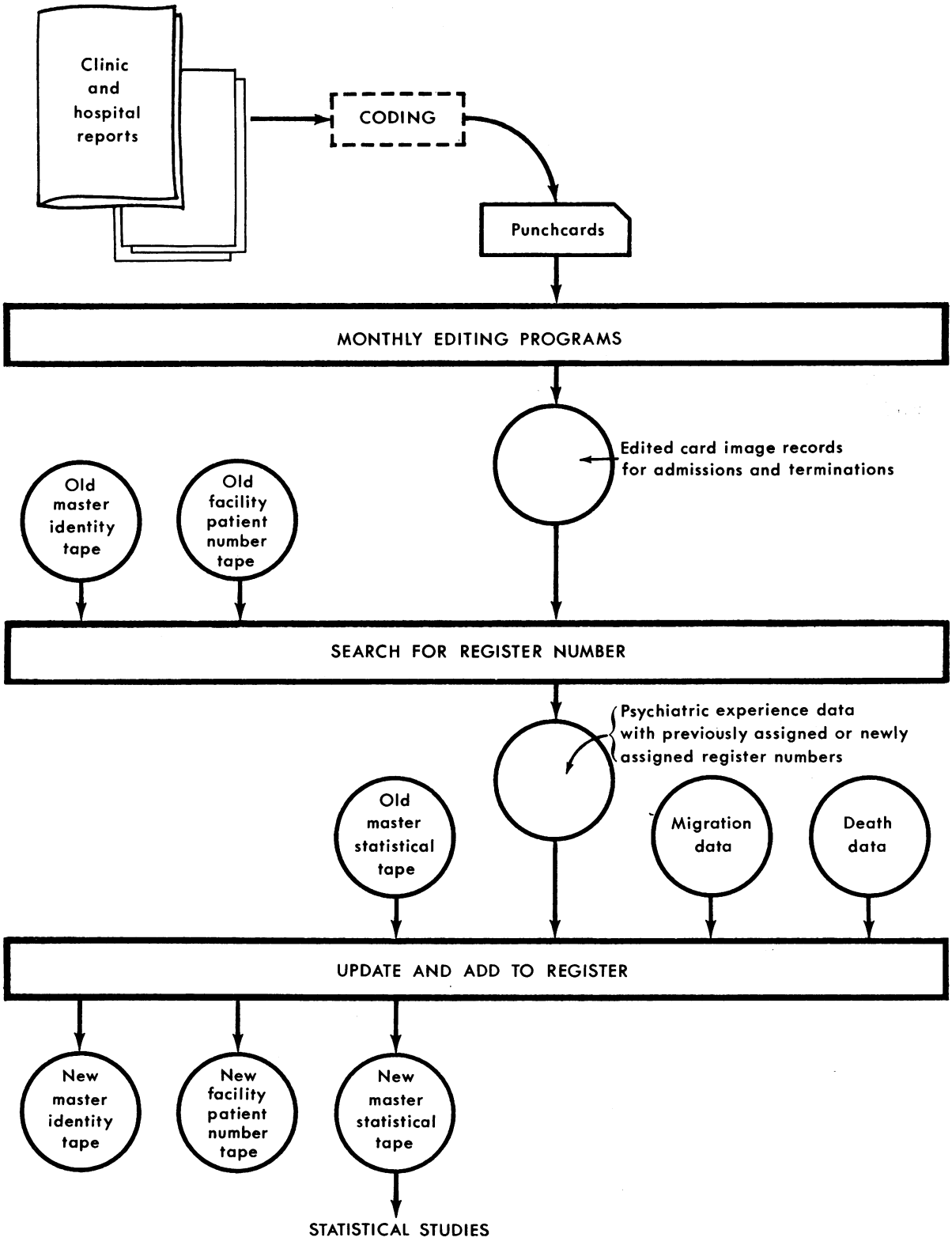
Input, Updating, and Clearance

The register will begin with all those enrolled in psychiatric facilities in Maryland as of July 1, 1961. Steps in future processing are illustrated in the chart. Ten different types of punchcards prepared from the various clinic and hospital reports will represent the basic input for maintenance and updating. It is planned to edit all records of admissions and separations monthly. A list of all errors and exceptions for clerical investigation and resolution will be produced. The monthly edited tape, as well as any corrections, will be input to a semiannual series of programs planned to match records by name and to update the master files. At the same time that the chronological portion (part III) of the patient's statistical record is brought up to date, the current status and yearly résumé (parts I and II) will be changed.

Updating of files on electronic computers is accomplished by merging the old tape with an input tape, making changes, and rearranging and adding information, thus producing a new or third tape. This may seem to be a waste, unless one realizes that data on magnetic tape can be erased and the tape used again. Also, the entire updating operation for the register will take about one-half to 1 hour of computer time.

When correction of data is necessary, it is not required to find, correct, and refile the

Processing of data, Maryland Psychiatric Case Register



original punchcard as in ordinary operations. Instead, a duplicate card containing the revised data and a special card type code will be used for automatic tape correction.

The register will be checked annually, using Soundex code, name, and other comparisons, against the 25,000 to 30,000 punchcards for deaths of residents. Upon determination of death of a registrant, all files will be annotated with the date of death. In addition, cause and place of death will be entered on the statistical record.

When persons move out of the State, they are no longer subject to readmission to a facility within the State and therefore must be removed from the patient population "exposed to risk" of readmission. We plan therefore to make periodic checks of addresses of former patients to determine migration out of the State. This will be accomplished by electronically listing the last reported addresses on perforated 3- by 5-inch card-weight paper and forwarding the resulting cards to the Post Office Department for verification of present residence. Changes of address will be added to the register by preparation of new punchcards. Since "outmigrants" may return to the State, their records will be retained in the master file for matching purposes but a separate file of "outmigrants" will be maintained for special studies. Persons of unknown address will be earmarked as "lost to followup" and may be subject to special followup studies.

Statistical Uses

The production of all types of reports based upon the psychiatric case register are facilitated by use of the computer. The entire statistical file could be run and pertinent statistical information selected in 15–30 minutes. Routine summaries based on the current status and cumulative counts and the yearly summary of experience from the master statistical record will describe for the current year the unduplicated numbers of patients admitted, discharged, and under care, by demographic and psychiatric characteristics, and the patient's annual and cumulative number of hospital days and number of admissions. Summaries will be prepared for each facility providing service and by the patient's geographic residence. Since data are obtained from all facilities serving residents of Maryland, the number of clinic and hospital patients can be related to the general population "at risk" of admission. Current population estimates for areas within the State will be on tape to permit automatic computation of rates per 100,000 population.

A principal administrative application of the register is the abstraction of data relating to patient "turnover," patient flow between facilities, and the chronicity of the disease. Typical questions are: How many patients served by psychiatric facilities this year had not been under psychiatric care at any time during the past 5 years and how many had been under continuous or intermittent care? How many and what kind of convalescent patients receive "aftercare" and what is the readmission rate of this group of former hospital patients compared with other convalescent patients? What is the subsequent psychiatric career of patients discharged as improved after treatment?

Probably the most interesting use of the case register will be the systematic study of changes in the psychiatric state of a cohort of new registrants using either life table methods or stochastic models. The date of each admission, discharge, and other event will be retained in permanent storage. It will be possible at any time therefore to abstract empirical data for fitting various mathematical models for describing the course of the disease. Thus, we might consider the conditions or distinct patient states as hospitalization, outpatient care, release from treatment, and death, or emigration, and assume that the stay of a given patient in a given state is a random variable. We might then attempt to answer such questions as: What is the distribution of time spent in any one occurrence of a given state? Given that a patient is in some state initially, what is the probability that he will be in a specified state at time "t"? What is the mean recurrence time for different types of patients? We might consider the various mental disorders as distinct states and ask: Is there a demonstrable progression of mental disorders from one type to another?

In addition, we are interested in using the register for a variety of special studies. For this purpose, it is planned to select a continuous sample of patients about whom additional in-

formation might be elicited. For example, some proportion of the register population will be untraced through ordinary means. The bias introduced by such persons "lost to followup" could be considerable. In addition, migration out of the State may be associated with greater or lesser mental health. It would seem important, therefore, to make intensive efforts to trace a sample of the patients lost or migrated so that selective factors associated with these events could be studied and sounder interpretation of data provided. The sample can also be used for the systematic followup of patients into the community in cooperative studies with the reporting facilities in order to determine the patient's current mental health status and level of adjustment. These findings would be used to obtain more complete estimates of prevalence of mental disorders and to determine factors which affect readmission. Also, the findings would aid in testing the validity of instruments designed for casefinding of mental disorders in community surveys.

In States without registers we attempt to determine the previous psychiatric care of the patient through information he or his family provides upon his admission to a psychiatric facility. In Maryland the reliability of such data can be ascertained by comparing the patient's report of previous care with his psychiatric register record. It may be possible also to use the register for studying family or household patterns of mental illness or to initiate other genetic studies based on the selection of patients with the same surname and address.

We would like to broaden our base for some Maryland communities by adding to our psychiatric facility reports data on mentally disturbed persons served by a wider spectrum of community agencies, such as social and welfare agencies, courts, and schools. With such data we could more adequately study the initial symptoms of disturbance in the problem individual and the natural history or life pattern of his mental disorders and determine the community services he receives in connection with his illness. Perhaps in this way we can learn at what point psychiatric intervention or some other type of community action could be most helpful in altering undesirable psychosocial patterns and can more closely approximate the true time and spatial distribution of mental diseases and their date of onset.

Discussion

We hope that the Maryland Psychiatric Case Register, in addition to producing a wealth of knowledge about mental disorders, can serve as a prototype for the maintenance and analysis of other chronic disease registers. Many if not all of the ambitious studies outlined could not be carried out without the newer methods of data processing.

Nevertheless, two questions about this research could justifiably be raised and should be answered—the problems of collecting data and the costs of programing:

1. Sustaining a high level of completeness and quality of data is one of the most formidable tasks in the maintenance of chronic disease registers. This includes the identification and reporting of all defined cases in a geographic area which meet the criteria for inclusion in the register, as well as the systematic collection of all pertinent followup information. Although data collection can be a major problem with some registers, to some extent, electronic methods can aid in this area. In the collection of followup information, forms which query regarding a patient's current status can be prepared automatically. Examples are: for this psychiatric register, annual queries to the post office regarding change in address; for cancer registries, annual queries to the physician inquiring about the patient's current status. By reducing through electronic methods the clerical work required to maintain the register, more clerical effort can be devoted to the control of reporting, such as carrying out procedures for followup on overdue reports. Furthermore, cooperation in reporting is more likely to be maintained at a high level where the reported data are used for extensive research.

2. The considerable time and cost necessary to develop the computer programing routines may be a second objection to their use. However, because the task to be accomplished is repetitive and the data are voluminous and complex, as in this instance, the labor and cost of writing the initial programs is in our opinion justified. In addition, it is planned that the

name-matching and other sorting routines that are developed for this register will be added to the library of subroutines. A meeting of persons participating in the development of electronic methods of maintaining psychiatric and other case registers is planned in the near future in order to share technical information.

REFERENCES

(*1*) Sartwell, P. E., and Merrell, M.: Influence of the dynamic character of chronic disease on the interpretation of morbidity rates. Am. J. Pub. Health 42: 579–584, May 1952.

(*2*) U.S. Public Health Service: State central case record systems and local case registers for tuberculosis. Washington, D.C., 1947.

(*3*) Bahn, A. K.: Cancer reporting and register program. Pub. Health Rep. 67: 395–400, April 1952.

(*4*) Cutler, S. J.: The role of morbidity reporting and case registers in cancer control. Pub. Health Rep. 65: 1084–1089, Aug. 25, 1950.

(*5*) Haenszel, W., and Hon, N. B.: Statistical approaches to the study of cancer with particular reference to case registers. J. Chronic Dis. 4: 589–599, December 1956.

(*6*) Bragg, K., Linden, G., and Breslow, L.: The California tumor registry; a summary report on the cancer experience in 36 hospitals. California Med. 87: 180–188, September 1957.

(*7*) Greenberg, R. A.: The Connecticut tumor registry. Connecticut Health Bull. 73: 2, February 1959.

(*8*) U.S. Public Health Service: A manual for a rheumatic fever registry. Washington, D.C.

(*9*) Kramer, M., Pollack, E. S., Locke, B. Z., and Bahn, A. K.: National approach to the evaluation of community mental health programs. Am. J. Pub. Health 51: 969–979, July 1961.

(*10*) Bahn, A. K.: Mental health clinic statistics: Needs, sources, methods. Pub. Health Rep. 69: 619–625, July 1954.

(*11*) Kramer, M.: Concepts in establishing mental health clinic reporting; workshop report. Am. J. Orthopsychiat. 27: 643–645, July 1957.

(*12*) Kramer, M., Pollack, E. S., and Redick, R. W.: Studies of the incidence and prevalence of hospitalized mental disorders in the United States: Current status and future goals. *In* Comparative epidemiology of the mental disorders. Grune & Stratton, Inc., New York, 1961, pp. 56–100.

(*13*) Bahn, A. K.: The development of an effective statistical system in mental illness. Am. J. Psychiat. 116: 798–800, March 1960.

World Forum on Syphilis

The World Forum on Syphilis and Other Treponematoses will be held September 4–8, 1962, in Washington, D.C. Sponsors of the forum are the American Venereal Disease Association, the American Social Health Association, and the Public Health Service. The World Health Organization and the International Union Against the Venereal Diseases and the Treponematoses will participate.

The purposes of the meeting are to set up a fund of current knowledge about syphilis control, to outline future courses of action, and to stimulate research and investigation in particular areas.

Special invitations to present papers will be issued to outstanding scientists throughout the world. Experts in both the social and medical sciences will reexamine present concepts of syphilis control and suggest new areas for development.

All persons or organizations engaged in or interested in venereal disease control anywhere in the world are invited to attend the forum. Organizations expecting to be represented and individuals planning to attend should write to Dr. William J. Brown, Venereal Disease Branch, Communicable Disease Center, Public Health Service, Atlanta 22, Ga.

The forum immediately precedes the meeting of the International Congress of Dermatology (September 9–16), also in Washington, enabling many specialists to attend both meetings.