# Development of a Computer Program for the Public Health Laboratory

ALFRED S. EVANS, M.D., M.P.H.

THE LARGE VOLUME of data coming into a public health laboratory can serve many purposes provided there is a convenient system for retrieval and analysis of the information. A well-organized computer program may be the answer. Four general objectives of its use would be to (*a*) provide epidemiologic information on disease distribution, (*b*) serve as a record of the tests done by the laboratory, (*c*) indicate geographic areas where control programs are most needed, and (*d*) establish a sound statistical basis for budget preparation.

The purpose of this paper is to report the experience of the Wisconsin State Laboratory of Hygiene in the evolution of a computer program for its laboratory data. After 5 years of experimentation a promising method has resulted that appears to answer most of our needs. We hope that this report on that method will serve both as a stimulus to, and as a basis for, initiation of computer programs in other public health laboratories. We should like to see such programs tied into a nationwide information network that would collect, analyze, and exchange laboratory data through a national center.

## Background

Some mention of the Wisconsin State Laboratory of Hygiene's background and functions may be pertinent because of its rather unique position under the university and because of the extensive services it offers to physicians of the State. Since its establishment in 1903, the laboratory has operated administratively under the University of Wisconsin and has been located on the university campus, near the medical school. Since 1959 it has been closely associated with the department of preventive medicine, and along with the medical school, nursing school, hospital, and psychiatric institute it forms one of the five major divisions of the university's medical center.

The laboratory has always served as the official public health laboratory of the State board of health, whose health officer and president are on the five-man policy committee. In addition to the usual environmental health and microbiological services to the State health department, the laboratory offers to physicians, without charge except for a handling fee in clinical chemistry, a varied and extensive menu of services. Included annually in these services are 50,000 throat cultures, 80,000 Papanicolaou tests for cervical cancer, and more than 100,000 tests in clinical chemistry. About a million tests are

*Dr. Evans is professor of epidemiology and director, World Health Organization Reference Serum Bank, Yale University School of Medicine, New Haven, Conn. At the time this paper was written he was director of the Wisconsin State Laboratory of Hygiene and chairman and professor, department of preventive medicine, University of Wisconsin, Madison.*

done annually. Their number and variety make the development of a data handling method, such as a computer program, a highly desirable aim.

Adoption in the laboratory of a diazo reproducing system (*A*) for making copies of combination request-report sheets released five employees, who had formerly typed the reports, for assignment to other work and permitted us to train them in computer activities. William H. Kenyon, who was with the State's bureau of administration when the computer program was being set up for the Wisconsin State Laboratory of Hygiene, assisted in its establishment. Edward O'Brien of the statistics section of the Wisconsin State Board of Health has provided continued aid in the program.

In our first computer system, initiated in January 1961, we used standard Remington Rand forms, which were machine-punched at the offices of the State board of health. The board also made its statistical services available to the laboratory.

Next the Univac optical scanner (*B*) and optical scanner (mark-sense) cards (*C*) became available to us through the board of health. In this system, a 45-column card was hand-marked with a black pencil, and from this card punchcards were automatically prepared. The data entered included all of the positive bacteriological reports on throat, urine, and stool cultures, both positive and negative cytology-pathology reports on the 80,000 or so Papanicolaou smears done, and periodic analyses for selected months of all tests done in clinical chemistry. Later on, rabies reports were entered on a separate form. This system was in operation from March 1962 until the end of 1964. The information obtained from it was useful in delineating epidemiologic patterns of streptococcal disease, *Salmonella* infection, cervical cancer, and rabies. In clinical chemistry, the system indicated which physicians were using our services and how many tests were done. The information was limited, however, to the amount which could be coded on one card.

In 1964, the Wisconsin State Government centralized its computer services under its bureau of administration, and the board of health discontinued its direct coding and computer operations. The bureau had an IBM 1410 data processing system (*D*) comprised of a processing unit, a card-read punch, a printer, a console, and a disk storage drive. In addition, an IBM optical scanner was purchased. This equipment permitted development of a new system for handling laboratory data which increased the space on the input sheet to 90 columns instead of the former 45. Thus, detailed laboratory results could be entered, rather than simply the words—positive, suspicious, or negative—to which we were limited with the Remington-Rand system. For data analysis, the bureau of administration used a program known as WISTAB (Wisconsin tabulator). This program had been developed by R. W. McCoy, director of the School of Commerce, Data Processing Center, University of Wisconsin, and by Kenyon, then systems analyst, department of administration, Systems and Data Processing Center, State of Wisconsin.

The first results from the new system for laboratory data were obtained in January 1965, and the rest of this paper is concerned exclusively with a description of this system.

## Numerical Coding

The first step in the development of the input data for the IBM 1410 processing system was the formulation of numerical codes that would be suitable for all types of laboratory information. Inquiries to the World Health Organization, the Communicable Disease Center of the Public Health Service, and a number of other State and city laboratories failed to reveal any precedent to follow. The usual codes, such as those employed in the "Standard Nomenclature of Disease," the "International Classification of Diseases," and the "Hospital Diagnostic Index," were not suited to our purpose because they were based on final clinical diagnoses. The information we collect does not consist of clinical diagnoses but of detailed pieces of laboratory data. We sought a basis for coding which (*a*) would be the one most applicable to all of our public health laboratory data, (*b*) would permit the recording of the details in identification that are necessary for epidemiologic purposes, and (*c*) would be flexible enough to permit inclusion of new data and new microbial species in the future.

A six-digit identification system was evolved

which could be used to code most of our laboratory results in line with our basic source references.

| *Section and major group* | *Code numbers assigned* | *Source* |
|---|---|---|
| Bacteriology | 000.000–299.000 | Reference 1. |
| Viruses | 5 digits—based on syndrome, virus group, subgroup, and type. | Derived from reference 2. |
| Mycology | 400.000–499.000 | Reference 3. |
| Parasitology | 500.000–599.000 | Based on reference 4, pp. 30–35. |
| Pathology | Morphology, 2-digit code; topography, 4-digit code. | Based on reference 5. |

In bacteriology, the latest edition of "Bergey's Manual of Determinative Bacteriology" available to us was the 1957 edition (*1*). Although outdated, it provided the most comprehensive listing of bacteria available and was therefore chosen for the coding system. The names of all the genuses of bacteria listed in the table of contents were numbered consecutively, beginning with 001 in a three-digit system. This method permitted inclusion of all genus categories by the number 208, including the *Rickettsia*. The family order, class, and species were then designated by the next three numbers. These were simply sequentially assigned as they appeared in the discussion of each individual genus.

For *Salmonella* serotypes, the numbers included in the Kauffmann-White schema reproduced in Bergey's manual (beginning with page 376) permitted the assignment of the numbers in that schema to 343 serotypes. *Salmonella* serotypes identified later were added arbitrarily to this list. Thus, *Salmonella blockley* became 102-345, *Salmonella ohio* 102-346, and *Salmonella molade* 102-347. In other instances it was necessary to make arbitrary decisions and additions in recognition of changes since Bergey's manual was printed, for example, in the identification of *Staphylococcus*, the mycobacterium genus, and atypical forms of *Mycobacterium tuberculosis*.

We had originally based virology coding on Andrewe's book (*2*), using six digits. After several trial runs, it became clear that the space needed to record all the tests under a six-digit code would be prohibitive. A different approach was therefore adopted in which a classification number for the basic clinical syndrome was used and subunits were designated for specific viruses. Under each of the subgroups, specific types can be identified up to 99. The way that the following coding system is used will become more apparent in the description of input data.

A. *Respiratory viruses*
   1. Influenza type A
   2. Influenza type B
   3. Adenovirus
   4. Parainfluenza 1
   5. Parainfluenza 3
   6. Parainfluenza 2
   7. Psittacosis
   8. Respiratory syncytial
   9. Rhinovirus

B. *Encephalitis viruses*
   1. Eastern equine
   2. Western equine
   3. St. Louis encephalitis
   4. Lymphocytic choriomeningitis
   5. Mumps
   6. California encephalitis
   7.–9. Unassigned

C. *Enteroviruses*
   1. Poliomyelitis I
   2. Poliomyelitis II
   3. Poliomyelitis III
   4. ECHO
   5. Coxsackie A
   6. Coxsackie B
   7. Rhinovirus
   8–9. Unassigned

D. *Exanthem*
   1. Measles
   2. Rubella
   3. Herpes simplex
   4. Smallpox
   5.–9. Unassigned

E. *Rickettsiae*
   1. Q fever (*Rickettsia burneti*)
   2. Rocky Mountain spotted fever
   3. Colorado tick fever
   4. Typhus, endemic
   5. Typhus, murine
   6–9. Unassigned

F., G., H., and I. Unassigned

J. Unidentified virus

The mycology code is based on the Public Health Service's Communicable Disease Center "Laboratory Manual for Medical Microbiology," published in 1963 (*3*). The series of numbers in this code is entered on the same code sheet as the one used in bacteriology. Similarly, the parasitology listing is also included on the

code sheet for bacteriology. The parasitology code is based on a reproduction from Faust and Russell's book on tropical medicine (*4*) of a list entitled "Zoological Names, a List of Filer Classes and Orders," pages 30–39.

The pathology-cytology sheet is separate. On it, the code numbers are those listed in the booklet "Systemized Nomenclature of Pathology," prepared by the Committee on Nomenclature and Classification of Disease of the College of American Pathologists (*5*). This coding system incorporates four major divisions—topography, morphology, etiology, and function. It was designed for coding hospital data and is excellent for this purpose. Its etiology section is not adaptable to public health laboratory information because the code numbers are not sufficiently detailed to record bacteria, viruses, and other pathogens encountered in a large reference laboratory. Nevertheless, for the area of cytology and pathology the system appeared adequate for our purposes and was adopted. A four-digit morphology code and a separate four-digit code for topography are used.

Based on the sources mentioned, a mimeographed guide and code sheet were prepared to assist the coders of the laboratory data. These aids listed all the common organisms likely to be encountered in 90 to 95 percent of our public health laboratory results. The basic reference books mentioned (*1–5*) are also at hand for the coders to consult if the classification desired is not included in the mimeographed guide.

For entering data into the computer scanner system, 8-by-11-inch, 90-column sheets were marked with a "magic felt marker." Separate entry forms were developed and printed for each of the following areas: bacteriology (including mycology and parasitology), immunology, pathology-cytology, viral and rickettsial infections, and rabies. No input form has yet been developed for the other operating sections of the laboratory—clinical chemistry, environmental health, and syphilis serology. Identifying information which would be largely common to all of the laboratory sections was placed in the computer system on the left-hand side of the coding sheet. This information, with the number of columns used for each item, was as follows:

| *Item* | *Number of columns* |
|---|---|
| Physician's or veterinarian's code number (includes identification of academic degree, type of practice, and location) | 9 |
| Patient's code number | 5 |
| Coder's number | 2 |
| Laboratory section where test was done (information was usually preprinted on sheet) | 3 |
| Age of patient | 2 |
| Sex | 1 |
| Specimen (code number) | 2 |
| Purpose of test (diagnosis, followup, or other) | 1 |
| Previous test and result | 1 |
| Laboratory technique used | 1 |
| Date specimen received | 5 |
| Laboratory identification number of specimen | 6 |

Data peculiar to each laboratory division were entered on the right-hand side of the entry sheet. Following are the descriptions of the data for each division with the number of columns used for each type of information:

| *Data by division* | *Number of columns* |
|---|---|
| *Bacteriology (including mycology, parasitology)* | |
| Clinical data | 7 |
| Test results | 1 |
| Identification of first organism by code | 6 |
| Quantitative titer | 2 |
| Number of colonies | 1 |
| Tuberculosis organism sensitivities | 3 |
| Antibiotic sensitivites (12 antibiotics) | 12 |
| Identification of any second organism | 6 |
| *Immunology* | |
| Clinical data | 7 |
| Test results | 1 |
| Quantitative titers for 15 tests | 15 |
| *Virology and rickettsiology* | |
| Clinical data | 5 |
| Clinical syndrome tested | 1 |
| Tested against more than 1 syndrome | 1 |
| Test used | 1 |
| Quantitative titer on acute serum against 10 possible antigens | 10 |
| Quantitative titers on convalescent serum samples | 10 |
| Isolation results: gross result | 1 |
| Virus group isolated | 1 |
| Virus subgroup isolated | 1 |
| Virus type isolated | 2 |
| Isolation systems used | 1 |
| Isolation systems yielding virus | 1 |
| *Rabies* | |
| Clinical data on patient | 1 |
| Type of exposure of human being | 1 |
| Part of body bitten | 1 |
| Severity of bite | 1 |
| Date of exposure or bite | 5 |
| Species of animal involved | 2 |
| County where bite occurred | 2 |
| Symptoms in animal | 12 |
| Test results by 4 laboratory techniques | 4 |
| Treatment of patient | 2 |
| *Pathology-Cytology* | |
| Symptoms and signs | 2 |
| Previous treatment or diagnosis | 2 |
| Previous test and when done | 2 |

| | |
|---|---|
| Results of previous test | 2 |
| Treatment or condition | 2 |
| Hormone evaluation | 2 |
| Pathological diagnosis, primary | 2 |
| Test results: positive or negative | 1 |
| Topographical diagnosis | 2 |
| Morphological diagnosis | 4 |
| Secondary diagnosis | 4 |

Because only five coders were available to record the large amount of laboratory information, compromises were necessary as to which data should be entered. First, it was decided to omit the physician's and the patient's code number as a routine entry. Duplicate examinations on the same patient were identified by the patient's age, sex, the section in which the laboratory work was done, which test was performed, and the county of origin of the specimen. Second, it was not possible to enter both the positive and negative data for all tests. Therefore negative data were recorded only in areas of special epidemiologic interest in which a denominator was needed to calculate the rate of positivity (number of tests positive over number done) and which might indicate age, sex, seasonal, or geographic trends of significance. The following entries were selected for inclusion of both positive and negative results: streptococcal throat cultures, antibiotic sensitivities of bacterial pathogens (including *Mycobacterium tuberculosis*), heterophile antibody tests, cytological examinations, and tissue diagnoses. In rabies tests and virology results, all tests done were recorded. As previously mentioned, no analyses are made in the clinical chemistry, environmental health, or syphilis serology sections, as they are beyond the capacity of our current personnel to code.

### Output System and Data Analysis

The information recorded on the IBM marksense data sheets was fed into the optical scanner and was simultaneously entered on punchcards and magnetic tape. The material to be programed and printed out was set up using the WISTAB program. This system was written so that the user could program his own tabulations with minimal assistance of a computer specialist. With it, tables showing the percentage distribution and two- or three-dimensional cross-tabulations are possible. Percentage distributions can be calculated based on row, column, or grand totals, as well as on row frequencies. The system was designed to provide an accurate and efficient means of data tabulation under the personal control of the user. Since some information needs are common to most of our laboratory divisions, these needs guided us in designing output requirements. Such common information requirements included the total number of tests done, their distribution according to the county of source of the specimen, the specific positives encountered (listed individually by age, sex, and county of the person tested), and analyses of the age, sex, and geographic distribution of persons whose tests were of special epidemiologic significance (in which instances, both positive and negative data were recorded and, for tissue diagnosis, the specific diagnosis made).

To date, computer programs have been completed in bacteriology (including mycology and parasitology), pathology-cytology, rabies, and immunology. Analysis of the data for virology is in process. The types of tabular material printed out routinely in the different divisions thus far programed are as follows:

| *Primary listing* | *Details included* |
|---|---|
| Bacteriology (monthly reports) | |
| 1. Individual listing of each pathogen by species. | Age and sex of person tested, date of test, specimen, and county of origin. |
| 2. Summary table of the number of all bacterial positives and streptococcal negatives, listed by organism. | |
| 3. Positive and negative cultures for hemolytic streptococci, listed by county of origin. | Number and percent of tests positive and negative by county. |
| 4. Age-group distribution of positive and negative hemolytic streptococci. | Sex for each age group; number of tests positive and negative. |
| 5. Same as item 3 but showing age-specific percentage distribution of positives. | Percentage distribution of all positives by age group. |
| Immunology (semiannual reports) | |
| 1. Distribution of the quantitative titers for each test done. | |
| 2. Distribution of individual heterophile tests in each age group. | Similar table with percentage data. |
| 3. Same as item 2 but summarizing tests as positive or negative by age group, by number and percent. | |
| 4. Analysis of symptoms in heterophile-positive cases. | |

| *Primary listing* | *Details included* |
|---|---|
| Immunology (Continued) | |
| 5. Analysis of symptoms in anti-streptolysin O positive cases. | |
| 6. Distribution of positive and negative results by county (separate tables for each test). | |
| Pathology | |
| 1. Distribution of all cervical cytology results by county. | Number and percentage of abnormal smears in each county. |
| 2. Distribution of results of all other cytology, listed by source of the specimen—sputum, gastric. washing, and so forth. | Number positive and negative. Recommendations made. |
| 3. Analysis of all results in cervical-vaginal cytology by age group. | Results (malignant, suspicious, atypical, negative) and recommendation (repeat smear now, repeat in 2 months, biopsy). |
| 4. Results of tissue examination listed by topography. | Number of lesions at each site which were benign and malignant and percentage distribution. |
| Virology | |
| 1. List of all positives either by isolation or by serology. | |
| 2. Age distribution of persons with positive results. | |
| 3. Distribution of positives by county. | |
| 4. Distribution of positives by specimen source (throat washing, stool, and so forth). | |
| 5. Distribution of positives by clinical syndrome. | |
| 6. Sympton analysis for each positive, by clinical syndrome. | |

## Workload and Costs

In 1965, five coders entered data on 218,796 test reports on IBM mark-sense cards, or about 44,000 per year per coder. This number may increase as greater familiarity with the forms develops. Following are the types and number of data sheets coded in 1965:

| | |
|---|---|
| Bacteriology | 59,407 |
| Cytology-pathology | 126,681 |
| Immunology | 27,917 |
| Rabies | 1,289 |
| Virus | 3,500 |
| Total sheets coded | 218,796 |

The cost of the computer program can be divided into initial costs and maintenance costs. The cost for the first year was approximately $29,000, including the cost of setting up the computer program and the cost of printing forms. The major item was $21,240 for coders, whose starting salary was $307 per month. These initial costs for the program during 1965 were as follows:

| | |
|---|---|
| Programing (personnel and machine time) | $500 |
| Cost of machine time for data processing | 4,490 |
| Cost of five coders | 21,240 |
| Forms—first print at $12.50 per 1,000; additional prints at $9.00 per 1,000 | 2,675 |
| Cost of marking pens | 60 |
| Total | 28,965 |

None of the coders had prior experience with the forms used in the program, but personnel can easily be trained in their use. In future years the cost of the program is expected to decrease by about $1,000, to about $27,715 annually. The cost per test coded was 12 cents; with improved experience in coding, it might approach 10 cents.

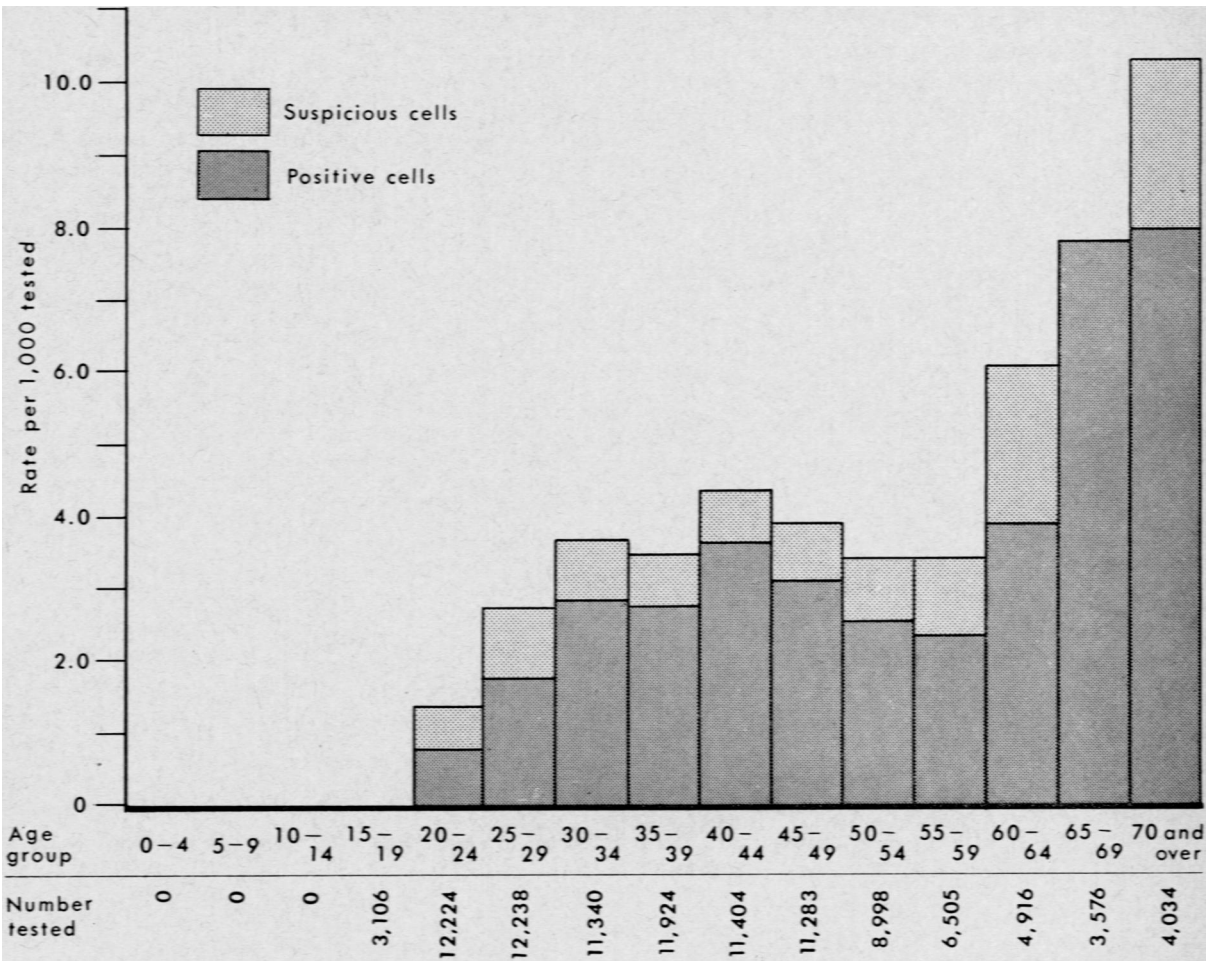
## Applications of Computer Data

The application and use of a computer system must justify the energy and funds expended in operation. Our current program has been operating for only 1 year so that it is premature to answer the question of justification decisively. Clearly, some of the results would have been extremely difficult or impossible to obtain by a manual tally system. Following are some of the ways in which the system has thus far been used.

*Workload analyses.* The computer analysis provides a method for determining the workload of each laboratory section if all tests are entered into the system. The cost of entering both positive and negative results may make such complete entry prohibitive, as was true in our laboratory. The data available, however, have proved useful in year-to-year comparisons of the workload for individual sections and for certain programs. The analysis by county has proved valuable in indicating the extent to which our State laboratory services overlap those of other public health or community laboratories serving the same area. Monthly analysis of certain tests like cultures for hemolytic streptococcus has enabled us to plan ahead in personnel and supplies for the peak months. Similar data in enteric bacteriology have afforded a basis for scheduling staff vacations in accordance with the workload.

*Planning.* Public health programs may be more intelligently planned, modified, or eval-

**Figure 1. Age-specific rates for positive and suspicious cancer cells in cervical cytology smears from 100,975 women, 1965**




uated if analytical data are available on which to base decisions. Of special interest is the identification of population groups at highest risk to specific diseases; such data can serve as a guide for the preferential provision of laboratory tests. Age analysis has proved most helpful in this regard. In figure 1 the age distribution of cervical cancer detected by the Papanicolaou tests is given. The low yield of positives in the under-25 age group gives a basis for eliminating this group and for expansion of public health screening programs in the older age group. Analysis of the frequency of hemolytic streptococcal infection at various ages has also provided useful data. The 5- to 14-year age group is the one most commonly yielding positive test results (fig. 2). Supplementary data obtained from questionnaires sent physicians using our services have also shown this age group to be the one in which cases are most frequently complicated by rheumatic fever (*6*). This information provides a basis for selective promotion of streptococcal testing programs in the 5–25-year age group. We have also used this information as a guideline in the accepting and grouping of specimens in peak-load months.

Geographic analysis of the source of specimens has in some instances been correlated with questionnaire data from pathologists as an aid for public health planning. For example, analysis of services for cervical cytology by this means indicated an under-utilization of those in the community cytological facilities and over-utilization of those of the laboratory of hygiene (*7*). The analysis also indicated that only 14.4 percent of the women over 20 years of

age were receiving a cervical cytological examination yearly. At that time about 175,000 of the women in Wisconsin were examined yearly, but it was estimated that about 350,000 could be examined if existing cytology services were fully utilized. An intensive drive therefore was begun to promote greater utilization of local facilities where they were available and to discourage use of State laboratory facilities in areas where local services were available. In 1965, a computer and questionnaire review of the status of Papanicolaou test usage in Wisconsin indicated that this objective had been largely fulfilled: 29 percent of the women over 20 were being screened yearly and the average cytologist was examining 5,000 smears yearly, close to the maximum (*8*). Thus, the problem of expansion of this test was no longer underutilization but too few cytologists. To meet this need we have enlarged our cytology school from 8 to 12 students.

*Reporting.* The computer program has supplemented the State health department's information on notifiable communicable disease and other public health problems. The monthly computer print-out of all bacteriological positives of public health importance is incorporated into the monthly laboratory newsletter sent to health officers, laboratory directors, and physicians using our services. The data are also available to the Wisconsin State Board of Health, to supplement its information on reportable diseases received from physicians and local health officers. In some instances, such as streptococcal sore throat, more cases are identified through the laboratory than through the physicians' reports. The results of our tests for cervical cancer are of epidemiologic importance because we identify 40 percent of all cancer of the cervix found by Papanicolaou smears in the State. The system also reinforces other reporting systems for salmonellosis and other enteric infections, tuberculosis, leptospirosis, histoplasmosis, and so forth.

Two routes of identification of persons with notifiable diseases are available to State health departments. One is through physician notification. The other is by identification of a case through a positive test on a specimen sent to the public health laboratory. If the names of

**Figure 2. Age-specific rates for hemolytic streptococcal sore throat in 29,217 persons, 1965**

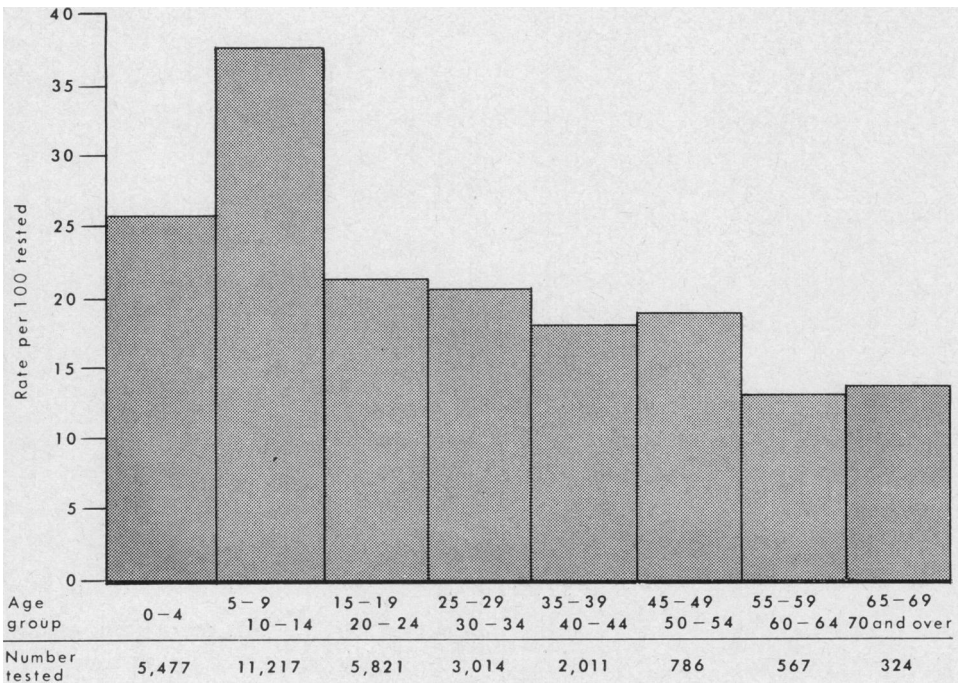
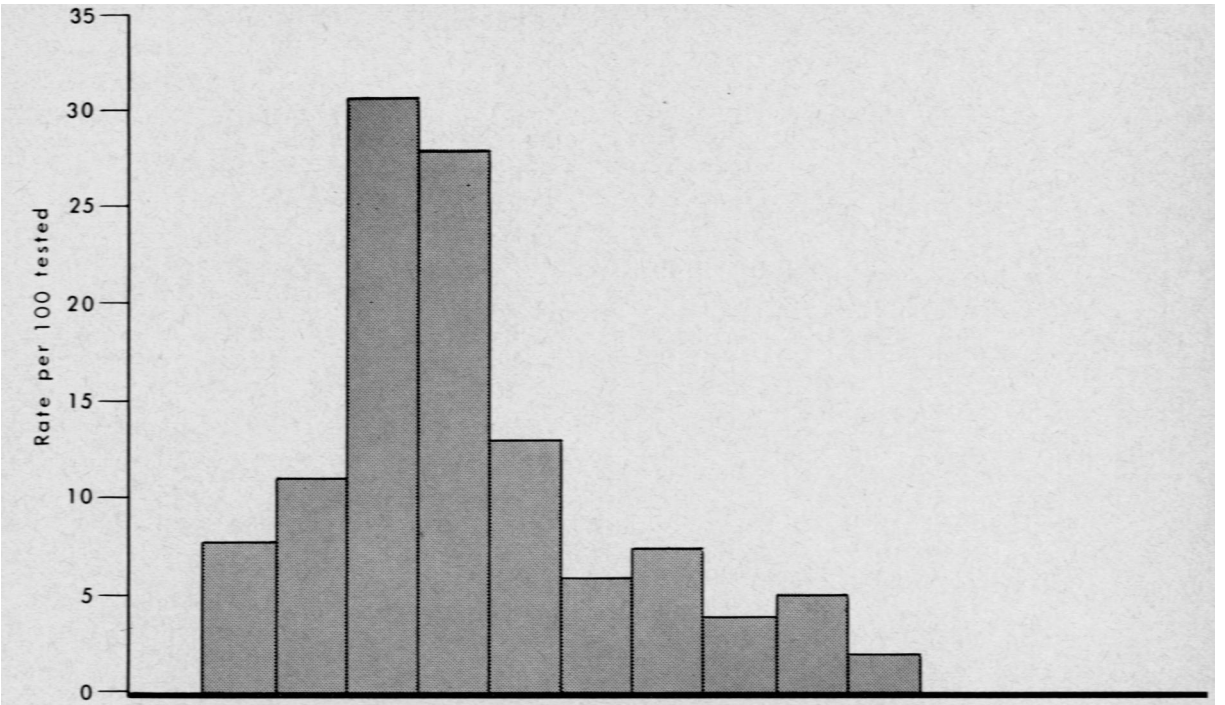


**Figure 3. Age-specific distribution of 942 persons with positive heterophile antibody tests for infectious mononucleosis among 4,598 persons tested, 1965**




| Age group | 0–4 | 5–9 | 10–14 | 15–19 | 20–24 | 25–29 | 30–34 | 35–39 | 40–44 | 45–49 | 50–54 | 55–59 | 60–64 | 65–69 | 70 and over |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Number tested | – | 415 | 667 | 1,737 | 847 | 290 | 183 | 147 | 99 | 76 | 46 | 38 | 21 | 14 | 18 |
| Number positive | – | 32 | 73 | 533 | 235 | 38 | 11 | 11 | 4 | 4 | 1 | 0 | 0 | 0 | 0 |

NOTE: Tests in which titers were at 1 : 16 or higher were considered positive.

the persons from both of these sources are entered into a computer, the names can be compared to identify duplicates. The addition of persons with positive laboratory tests but not reported by physicians provides a more complete picture of disease incidence and also measures the adequacy of physician reporting.

In some States, tests for phenylketonuria are done entirely in the public health laboratories. In these States, a comparison by computer of the names of babies tested in the laboratory with the names recorded on birth certificates in the same period of time would indicate the percentage of newborn babies being tested and the names of those who have been missed.

If it were possible to include both private and public laboratories in data going into a computer system, much better knowledge of certain diseases would result.

*Epidemiologic data.* A large number of epidemiologic applications of a computer program in the laboratory are possible. The place of epidemiologic studies in a public health laboratory is being reviewed separately (*9*). Briefly, demographic data, prevalence and incidence estimates, susceptibility of age groups to certain diseases, changing temporal patterns, and sometimes clues to causes can be derived from computer analysis of public health laboratory data.

*Budget preparation.* An obvious use of good laboratory data is in budget justification and in estimating the cost of specific programs. The cost of case detection for cervical cancer or for prevention of rheumatic fever can be derived by computer and the information used in the support and justification of laboratory programs.

*Laboratory operation.* Finally, the laboratory staff can compare new techniques with old ones. For example, the rate of cancer detection

when we used only a single cervical scraping was very close to that achieved when both a scraping and a vaginal aspiration were done. The far greater number of women who could be examined by the same personnel if only one test was done resulted in a large net gain in the total number of cases of cancer detected. In computer analysis of our streptococcal program, we learned that about 90 percent of all hemolytic streptococcal cultures were in group A, irrespective of the month or year of the test or the age group of the patient. On this basis, no practical loss would be encountered if grouping as a routine procedure were abandoned or, conversely, if only group A cultures were identified by the fluorescent antibody technique. Computer data on the age-specific distribution of positive heterophile antibody tests for infectious mononucleosis will indicate the age groups in which testing will reveal the greatest number of cases (fig. 3).

## Discussion

We are well satisfied with the concept of our present computer operation as it permits us to do our own coding and entering without any special equipment or highly trained personnel. Many improvements, however, are still needed. For special projects, a LINC computer (*E*) would be an asset, with its capability of immediate analysis and print-out. This computer could also be used as an input source for the IBM system.

In developing our program we were aware of the need for a system and coding method that would be flexible and which might be adapted to use in other public health laboratories. We believe that now is the time to draw plans for a national laboratory computer network. It might operate at the Communicable Disease Center or at some other central governmental computer center. (As of November 1966, an ad hoc committee appointed by the Center was working in this direction. Several other States are also in the process of developing a laboratory computer system, some using as a model the program outlined in this paper.)

A national laboratory computer network could act as a central processing unit for its own operation and for certain laboratory data for each State. Larger States would probably want their own programs, which could be made compatible with the one at the national center. Smaller States might wish to use the national center directly. It is still early enough in the development of a national program to plan carefully toward a coordinated, compatible system with common classification and coding methods. We hope that our pilot program will be of use in the realization of such a national network.

## Summary

A computer program for analysis of public health laboratory data using an IBM 1410 data processing system has been developed at the Wisconsin State Laboratory of Hygiene. Input data are entered on mark-sense 8- by 10-inch coding sheets. A six-digit code was evolved for handling microbiological data. Each of five laboratory sections has separate entry forms. On these forms, the 45 columns to the left are used for data common to all sections, while the 45 columns on the right are used for data peculiar to the individual section.

Output data are printed out monthly for bacteriological and cytological data; less frequently in the areas of immunology, rabies, and virology. All positive test results are recorded, as well as negative data for tests of special interest, such as streptococcal sore throats, heterophile antibody, bacterial sensitivities, and cervical cytology; both positive and negative tests in the immunology, rabies, and virology sections are also recorded. Print-out data include the number of tests and the distribution of the persons tested by county, age, and sex. Tabular data showing the percentage of specific tests positive, by month, age and sex of the tested, and serologic titer are an intrinsic part of the regular report.

Five coders made 218,796 entries during the first year of operation of the computer system, or about 44,000 entries each. Each test in the program cost about 12 cents.

### REFERENCES

(*1*) Breed, R. S., Murray, E. E. D., and Smith, N. R.: Bergey's manual of determinative bacteriology. Ed. 7. The Williams & Wilkins Company, Baltimore, 1957.

(2) Andrewes, C. H.: Viruses of vertebrates. Ed. 1. The Williams & Wilkins Company, Baltimore, 1964.
(3) U.S. Public Health Service: CDC manual for medical microbiology. Communicable Disease Center, Atlanta, Ga., 1963.
(4) Faust, E. C., and Russell, P. F.: Clinical parasitology. Ed. 7. Lea & Febiger, Philadelphia, 1964, pp. 30–39.
(5) College of American Pathologists: Systematized nomenclature of pathology. Chicago, Ill., 1965.
(6) Grover, A. A., Spoerk, K., and Evans, A. S.: Streptococcal infections. Observations from a public health laboratory. Amer J Public Health 55: 1609–1621 (1965).
(7) Evans, A. S.: Current status of cervical cytology in Wisconsin. Wisconsin Med J 62: 211–216 (1963).
(8) Laboratory Newsletter (Wisconsin State Laboratory of Hygiene) 4: 2–16 (1965).
(9) Evans, A. S.: Epidemiology in the public health laboratory. Amer J Public Health. In press.

EQUIPMENT REFERENCES

(A) Bruning diazo wet copier, model 250. Chas. Bruning Co., Division of Addressograph Multigraph Corp., Mt. Prospect, Ill.
(B) Univac optical scanner, model 5400. Sperry Rand Corp., Philadelphia, Pa.
(C) Univac optical scanner cards. Sperry Rand Corp., Philadelphia, Pa.
(D) IBM 1410 data system, including a 1410 computer for 80 K. storage, model 1230 optical scanner, model 737 hand punch, model 083 sorter, model 1301 disk storage unit, model 1403 printer, and model 1402 reader. International Business Machine Co., Armonk, N.Y.
(E) LINC computer. Digital Equipment Corp., Maynard, Mass.

## Construction Grants to Medical Schools

Six construction grants totaling nearly $14 million are being made to schools of medicine, osteopathy, and dentistry to expand teaching facilities.

The Public Health Service grants, made under provisions of the Health Professions Educational Assistance Act, will aid in construction of facilities to accommodate 62 additional first-year students annually.

Western Reserve University School of Medicine in Cleveland, Ohio, received $4,560,908 for an addition to the medical school building (21 first-year students).

Jefferson Medical College of Philadelphia, Pa., received $392,527 for equipment (14 first-year students).

The Chicago College of Osteopathy, Chicago, Ill., received $1,924,910 for additions to the teaching hospital and the basic science building (5 first-year students).

Emory University School of Dentistry, Atlanta, Ga., was awarded $3,498,582 for an addition to teaching facilities (22 first-year students).

Indiana University School of Medicine in Indianapolis was awarded $2,190,584 for the replacement of clinical teaching facilities at James Whitcomb Riley Hospital for Children. The grant will enable the university to replace obsolete facilities and, as a result, maintain an acceptable level of medical education.

The Woman's Medical College of Pennsylvania in Philadelphia received $1,354,233 for a new clinical laboratory and laundry building, an addition to the existing clinical science building, and the reconversion and renovation of certain areas within the existing building.

With these awards, 80 grants totaling about $180 million have been made under this program. These grants will result in a first-year enrollment increase of 2,688 students in various health professions.