Annu. Rev. Publ. Health 1991. 12:257–80

EPIDEMIOLOGIC SURVEILLANCE IN DEVELOPING COUNTRIES

Ralph R. Frerichs

Department of Epidemiology, University of California at Los Angeles, Los Angeles, California 90024-1772

KEY WORDS: microcomputers, sample surveys

INTRODUCTION

Developing countries are different from technologically-developed countries in many ways. Most people are poorer, less educated, more likely to die at a young age, and less knowledgeable about factors that cause, prevent, or cure disease. Biological and physical hazards are more common, which results in greater incidence, disability, and death. Although disease is common, both the people and government have much fewer resources for prevention or medical care. Many efficacious drugs are too expensive and not readily available for those in greatest need. Salaries are so low that government physicians or nurses must work after-hours in private clinics to feed, clothe, and educate their families. The establishment and maintenance of an epidemiological surveillance system in such an environment requires a different orientation from that found in wealthier nations. The scarcity of resources is a dominant concern. Salaried time spent gathering data is lost to service activities, such as treating gastrointestinal problems or preventing childhood diseases. As a result, components in a surveillance system must be justified, as are purchases of examination tables or radiographic equipment. A costly, extensive surveillance system may cause more harm than good.

In this article I will define epidemiologic surveillance. I also will describe the various components of a surveillance program, show how microcomputers and existing software can be used to increase effectiveness, and illustrate how

0163-7525/91/0501-0257$02.00

rapid microcomputer-assisted surveys can be used to supplement existing efforts.

EPIDEMIOLOGIC SURVEILLANCE

The word "surveillance" evokes images of mystery, police, saboteurs, and intrigue. The word was adopted by the English from a French term during the Napoleonic Wars (10). The original meaning was to watch over people who supposedly had subversive or criminal tendencies so as to prevent future problems. In contrast epidemiologists use the word in a way that is now common to public health. Epidemiologists follow disease cases or deaths with the intention of preventing future harm. Thus, surveillance is "the continued scrutiny of all aspects of occurrence and spread of disease that are pertinent to effective control" (11).

An expanded definition, which has been adopted by many health professionals in the United States, was published in 1986 by the Centers for Diseases Control (CDC) (37):

> [Epidemiologic surveillance is] the on-going systematic collection, analysis, and interpretation of health data essential to the planning, implementation, and evaluation of public health practice, closely integrated with timely dissemination of these data to those who need to know. The final link in the surveillance chain is the application of these data to prevention and control. A surveillance system includes a functional capacity for data collection, analysis, and dissemination to linked public health programs.

Again, the definition emphasizes the use of systematically collected data for control of disease.

Two terms that are related to surveillance are "monitoring" and "register" (or "registry"). Monitoring and surveillance are sometimes, but not always, used synonymously (10, 30, 37). Epidemiologic surveillance is a monitoring process. Both surveillance and monitoring feature continuous scrutiny and detection of change from expected levels. A surveillance system, however, goes one step further. Because surveillance also includes feedback and control, it is actually a monitoring and control process, rather than just a monitoring process. A register is similar to a surveillance program; both feature continuous gathering of data on all disease cases in a geographically-defined area. The primary focus of a register, however, is to list cases in need of long-term treatment (e.g. tuberculosis and leprosy registers), to describe patterns and time-trends of population-based incidence or mortality rates (e.g. diabetes registers) or to provide cases or deaths for etiologic studies (e.g. cancer registers).

Monitoring and Control Process

Perhaps the best way to explain a monitoring and control process is with the analogy of a building thermostat. Assume the thermostat regulates the heating and air-conditioning system. A person sets a temperature gauge to a comfortable level. The building temperature is continuously sensed by a thermometer in the thermostat. The actual temperature is compared with the desired, previously set level. If the observed temperature differs from the expected temperature, the thermostat sends a signal to activate the heater or air-conditioner. Either one remains on until the temperature returns to the desired level. The thermostat then sends a signal to turn off the heater or air-conditioner. The process of scrutiny and control continues until either the thermostat is turned off or the system breaks down.

An epidemiologic surveillance program is similar to the temperature-regulating system. Rather than sensing temperature, the surveillance program scrutinizes the occurrence of disease in specified populations. As shown in Figure 1, such a program typically has four components: sensor, monitor, level of expectation, and controller. The sensor identifies the state of health in the population under observation and sends a reference signal, which describes the state of health. The monitor receives the reference signal and compares it with the expected level. Usually, the expectations are set based on national or local policies or standards. If the measured level of disease differs from expectation, the monitor sends an error signal to the controller. The controller, who is a person or agency with some ability to take corrective actions, acts to reduce the level of disease in the population. The effectiveness of this action is measured by the sensor, which again sends a reference signal to the monitor. If the disease incidence falls below the expected level, no

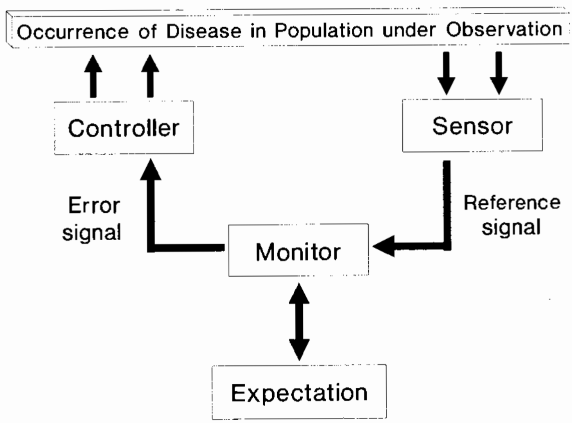

Figure 1 Components of an epidemiologic surveillance system.

further action is required. If the controller has not been successful, however, the error signal again points to the unacceptable level of disease and the process continues.

SENSOR The sensor may take several forms. In some surveillance programs, the sensor is local physicians who are required by law to report the occurrence of notifiable diseases. In other programs, it may be health workers at a hospital or clinic. If blood is routinely screened for evidence of genetic disorders or antibodies for various infectious diseases, then the laboratory is the sensor. Teachers or school nurses may represent a sensor if they routinely report illness-associated absenteeism. Similarly, industries that routinely report absenteeism to government health agencies may be a sensor. A sensor also may be an annual interview or examination for members of the military.

MONITOR AND EXPECTATION The monitor is usually an epidemiologist who reviews the information sent in from the field to decide if there is an acute outbreak or if the disease has reached epidemic levels. The epidemiologist compares the disease occurrence with a prior expectation and responds accordingly. For every disease, there is a threshold level of concern that might not be stated explicitly, but is there nevertheless. If a disease is uncommon, such as human rabies, or is subject to international health regulations, such as cholera, yellow fever, and plague, then the threshold to initiate action may be a single case. For more common diseases, such as malaria, influenza, and measles, the threshold may be many cases.

The level of expectation depends to a large extent on the feasibility of control and the resources of the control agency. For example, Walsh & Warren (40) note that the feasibility for malaria control in most countries is high because inexpensive drugs are available for prevention and treatment, and spraying programs work well to reduce the level of the vector. If a country has little money for control programs, the threshold may be set very high, so that only the occurrence of large epidemics will stimulate action.

CONTROLLER The controller is the person or organization that initiates prevention or control activities. At the regional level, the controller may be a local health officer. At the national level, it may be the chief medical officer in a department of disease control. Some countries provide direct funds to epidemiology units for disease control. Most developing countries, however, maintain a tight rein on funding; the authority to allocate resources is vested in administrators, rather than epidemiologists. In such instances, the epidemiologist must explain the surveillance findings and present recommendations for action to the administrator both in a timely manner and in terms that can be understood easily.

Total Count Versus Sampling

Most surveillance programs attempt to identify all cases, deaths, or injuries in a population for the disease of interest. This type of data gathering works well in developed countries that have sufficient resources and health professionals who comply with reporting requirements. Because surveillance diseases often are rare events in wealthy countries, no single practitioner is burdened excessively by the time-demands of reporting. When the disease is more common, however, even wealthy countries revert to sampling. For example, the United States maintains an active surveillance program of pneumonia-influenza deaths. The deaths, however, are tallied each week for a sample of 121 cities, rather than for a total count of the country (5).

Samples also are commonly used in developing countries. As has been noted by the Expanded Program on Immunization (EPI) of the World Health Organization (WHO), reporting systems that require total counts are notoriously incomplete for most diseases (9). Based on EPI's experience, routine reporting is reasonably complete only when a disease is recognized as a high priority problem and when cases are less common. Thus, sampling is often the only way to obtain timely, accurate information. For example, in Myanmar (formerly Burma), hospitals are the major sources of surveillance data for several diseases. Rather than a total count of all patients, the Department of Health only requires a 10% systematic sample of hospital patients discharged alive or dead (26). Surveillance data on nutritional status in Ethiopian refugee camps are provided solely by periodic cluster surveys (42).

When routine reporting systems are not functioning properly, EPI recommends that disease-specific data be obtained in surveys of sentinel sites, such as cooperative hospitals and clinics (9). Only a few sites are selected for sentinel reporting; selection is limited to those sites that have a capable and interested staff. For common diseases, such as measles and pertussis, the EPI program recommends the use of outpatient departments of pediatric hospitals or general health centers; for less frequent diseases, such as neonatal tetanus, diphtheria, and poliomyelitis, they recommend infectious disease hospitals or rehabilitation centers. Not all researchers agree, however, that sentinel sites are effective at surveillance. Walsh (39) has presented several caveats for those researchers who plan to use sentinel surveillance, the most important of which is that the selected sites should be representative of all sites in the country.

Later in this review, I will discuss the advantage for developing countries of using occasional sample surveys to supplement a total count surveillance program. Periodic surveys are especially useful for common diseases, injuries, or other health conditions for which the burden of reporting is excessive.

Diseases Included in Surveillance Programs

Many different diseases and health conditions are included in surveillance systems. The CDC has 98 on-going systems for surveillance (3). Although many of the diseases are communicable, the list is varied and includes such conditions as ectopic pregnancies, abortions, and occupational injuries. Similar diversity was apparent in a recent text on surveillance programs, primarily in economically more developed countries (10).

For developing countries, the list of disorders is much shorter and is usually limited to infectious diseases and, possibly, nutritional disorders. Most countries attempt to gather surveillance data on cholera, yellow fever, and plague, the three diseases subject to International Health Regulations. Other conditions typically included are those under active surveillance by WHO, such as AIDS, dracunculiasis (guinea worm), poliomyelitis, influenza, and malaria. Also included are those conditions that are required to be reported at the national level by most countries, such as various forms of encephalitis and hepatitis, measles, typhoid and paratyphoid, tuberculosis, leprosy, diphtheria, pertussis, human rabies, and tetanus. Other problems occasionally reported are accidents, low weight-for-age or weight-for-height, dog bites, and snake bites.

Reporting and Feedback

In the ideal, a surveillance program receives data on a regular basis from health professionals in the field and provides them with prompt feedback on the tabulated findings. For surveillance systems that rely on multiple reporting sites, the process of reporting and feedback is shown in Figure 2. Data are systematically collected from sensors, which include hospitals, clinics, private practitioners, school nurses, and company physicians. All data are sent to

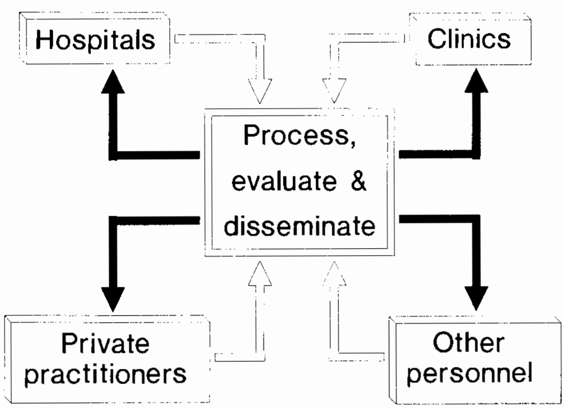

Figure 2 Reporting, processing, and feedback of surveillance data.

a central site for processing, and evaluation of disease changes over time. The information is then shared with program managers and other officials who are responsible for disease control and distributed back in summary form to the persons who initially provided the data.

The method of reporting can be either active or passive. In active surveillance, health department personnel contact sources in the field at regular intervals and requests specific data on diseases of interest. An active approach requires considerable time and effort by health department workers and is not often used in developing countries. In passive surveillance, the more common method, physicians and others in the public and private sectors complete a form for all reportable diseases. The form may be sent to the health department on a regular basis or when a disease case has been observed. Rather than relying on the salaried time of health department personnel, passive surveillance transfers part of the data acquisition cost to field professionals. Unfortunately, as noted by Evans (8), "time and lack of interest [by field personnel] greatly limit this [passive] system to a small percentage of reportable diseases." Thus, although the cost is low, so is the quality of the data.

Feedback is an extremely important component of a surveillance program. As EPI has stated, "without feedback, health workers soon realize that it makes no difference what they report or whether they report" (9). Information has value only if it is used for decision-making. If field personnel do not understand or see the results of the decision-making process, they will be less willing to contribute to future surveillance efforts.

Epidemiologic Surveillance and Health Information Systems

When a government assumes responsibility for health care, epidemiologic surveillance activities are usually included as part of the health information system. Field personnel are expected to complete reporting forms for the health information system on a regular basis and send the forms to the next highest official. At the community level, the person who delivers services also gathers data. These data are explicitly used for planning, surveillance, management, evaluation, and education. However, the implicit assumption is that the data will be used to improve health status and bring the population closer to "Health for All," the goal espoused by member nations of WHO. If health information systems did not exist, physicians and others would still deliver services and provide advice on prevention and control of various diseases. Additional knowledge should not hinder delivery; it should improve effectiveness without raising cost.

When too much time is spent gathering data, the efficiency with which a program is managed may worsen. Each health worker only has a finite amount of available time. The relationship between data-gathering time and management efficiency is shown in Figure 3. When a health worker spends little time gathering data, more time is available for the delivery of services, but not in the most efficient manner. As more data are gathered, managerial functions, such as scheduling patients, maintaining equipment, ordering supplies, arranging counseling sessions, and scheduling immunization and well-baby clinics, are improved. However, the gains in efficiency often are offset by the loss of service time. This point is labeled "threshold" in Figure 3. Once health workers have moved to the right of this threshold, too much time is spent gathering data and not enough time is spent delivering services or planning prevention or control programs. In the post-threshold situation on the right side of Figure 3, health workers become discouraged by excessive demands on their time and, as a result, no longer provide truthful information. Program administrators become overwhelmed with the volume of data and fail to review adequately the information. Finally, government health officials become cynical about quality and refuse to use the data. The cost in salary-time remains the same, but the effectiveness of the health care system is diminished.

An example of how data gathering demands can overwhelm a health care system is taken from Myanmar in 1985 (12). Peripheral-level health workers were expected to complete more than 30 sets of forms either on a daily or monthly basis. At the rural health center level, the local staff had to submit a set of forms with 1160 variables every month. Of these variables, 72% were requested by the Division of Disease Control, the unit responsible for epidemiologic surveillance. At the next administrative level, Township Medical Officers each month had to process and review these variables, plus 786 more variables on township-level activities, for a total of 1946 variables. The surveillance needs of the Division of Disease Control accounted for nearly 1000 of the variables sent forward by the Township Medical Officer. By the time the information was received at the national level, processed and analyzed, and included in a detailed report, three years would pass. The

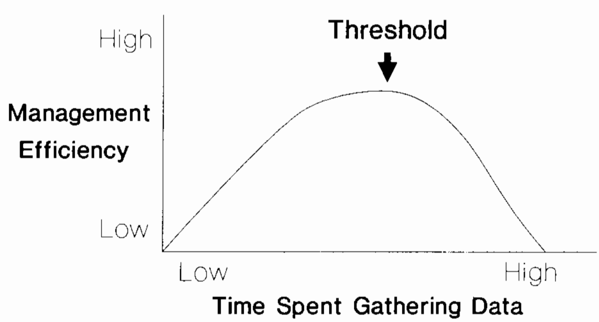

Figure 3 Effect of data gathering time on management efficiency.

information was then too dated to be used for decision-making. Although the surveillance report represented the time and effort of countless health workers, no one trusted the data.

When either too much information is requested or the local health worker does not understand the need for information, reporting quality and completeness decline. For example, in ten developing countries the estimated completeness of the existing surveillance program for poliomyelitis was 1–26%, with an average of 7–8% complete (41). Nearly 93% of poliomyelitis cases were typically unreported. In Myanmar, the best reporting system was a set of monthly forms with a 10% systematic sample of discharges from the nation's 614 hospitals (12). Unfortunately, even with this system that relied on government-funded physicians, only 38% of the hospitals submitted a complete set of 12 monthly reports.

Data, Information, Knowledge, and Action

Surveillance programs by definition are action oriented. They are intended to serve as an early warning system for both raging epidemics and small outbreaks. It is easy, however, for surveillance efforts to become too complicated and loose their focus on action. Realizing the danger of too much data, EPI has stated that disease-specific reporting forms should be simplified to include only the number of cases and be limited to those diseases for which control programs exist or for which control efforts would be initiated (9). A similar caveat is given by Walsh (39), who states that "data should be collected to provide information that is essential for making decisions."

The link between data and action is not always evident. As seen in Figure 4, the flow of information starts with data, a set of discrete observations or facts on cases, deaths, and people. When processed and analyzed by epidemiologists or statisticians, these data are converted to percentages (or proportions), incidence or mortality rates, prevalence or risk ratios, or indices, such as potential years of life lost (32) or days of healthy life lost (24). Nothing more happens, however, unless an administrator or program manager learns of the information and absorbs it as new knowledge. Action only follows if the administrator has the will and political power to act.

If data are to result in action, various requirements must be met. First, the epidemiologist can help determine a minimum data set essential for decision-making, but cannot control the quality or timeliness of data gathering. This requires highly motivated health workers and good supervisors. Once the data are collected, the epidemiologist must turn them into information that communicates clearly to the administrator. In much of the developing world, however, surveillance data are presented in complicated tables that are hard to read and do not point to any specific action. Thus, the administrator will often move on to other matters and either delay a decision or allocate the same

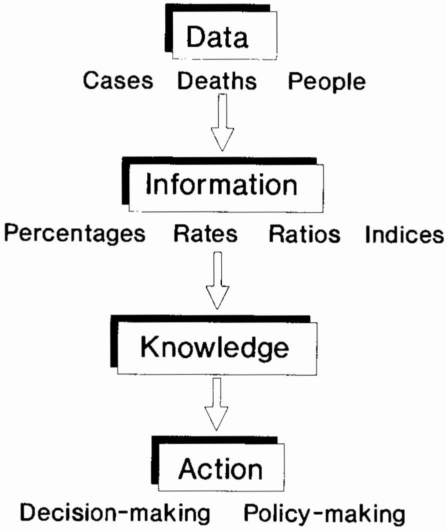

Figure 4 Link between data, information, knowledge, and action.

resources as in the past. Such decisions are especially harmful if diseases are flourishing at epidemic levels.

Graphs are the best way to present surveillance information. At minimum, every health region should have on their wall a graph showing the time trends for each disease in the surveillance system. If the disease is common, there also should be a line to distinguish epidemic levels from nonepidemic levels. Such information can easily be understood by health workers and administrators.

Action requires more than local health workers and an epidemiologist. Resources are usually scarce and there are many demands for program funds that must be satisfied by the administrator. The epidemiologist can do much, however, to guide program managers and administrators. Tugwell and colleagues have provided a framework to guide the decision-making process, which assumes that the goal is to reduce the burden that illness places on people at the community level (38).

USE OF MICROCOMPUTERS FOR SURVEILLANCE

Although much can be done in developing countries with minimal data sets, committed personnel, and active supervision, additional advances in surveillance programs can be brought about by the use of microcomputers. In the past decade, microcomputers have become increasingly important in developing countries. These small, inexpensive computers have the potential to revolutionize surveillance activities, as they have done for epidemiological research. In 1984, Frerichs & Miller (19) brought a microcomputer to Bangladesh and proved that health professionals who had never before used a computer could be taught in a short time to use this technology for epidemiologic research, data management, graphical presentations, and statistical analyses. In 1985, Gould & Frerichs (25) reported that the Bangladesh investigators were continuing to use the computer for both education applications and many research-oriented tasks. Also during the mid-1980s, Bertrand and colleagues were using microcomputers in Africa and South America for population-based surveys (1) and for management and planning purposes (2). They included an extensive surveillance program in the Sahel region of Africa to provide early warning of famine for relief projects (W. Bertrand, personal communication). Frerichs & Selwyn (20) recently have summarized many ways microcomputers serve the field of epidemiology, including data processing and analysis, program management, graphical presentations, and computer-assisted surveys.

Thacker & Berkelman (37) commented that the use of computers has increased the timeliness of data collection and analysis and has decreased the reliance of epidemiologists on programmers and statisticians for analysis and interpretation of findings. User-friendly software has allowed the epidemiologist to manipulate and analyze data for biologically- rather than statistically-significant relationships. Once, surveillance activities were hampered by lack of software; this problem no longer exists. Many types of database management programs are readily available from software vendors, along with teaching guides and reference manuals. Most of these programs, however, are costly and limited to data management. An exception is *Epi Info*.

Epi Info

Dean and colleagues at CDC have developed *Epi Info,* a multipurpose, public-domain program for word processing, database management, and statistical analysis (7). The price is minimal and includes the cost of printing the manual and preparing the computer disks for distribution. The program is used by epidemiologists around the world, including many who have been trained at CDC or by CDC personnel. The most recent version of the software was sponsored by both CDC and the Global Program on AIDS of WHO (see Ref. A). In the United States, the *Epi Info* program is of central importance to surveillance activities. Approximately 30 states use *Epi Info* to maintain, edit, and process data files containing disease reports. Every week, all 50 states send their local computer files to CDC to be merged in a standard format into a single national data file. *Epi Info* provides the standard format for the surveillance program.

For developing countries, the most recent version of the program assists with several important operations in disease surveillance. First, data from case records are entered into the computer and edited for minor errors, including duplicate entries. Second, the program processes the data and produces periodic tabulations of disease totals or disease cases. Third, frequency distributions and listings of any specified disease are produced by such variables as month, town or city, age, and sex. Fourth, the data are converted into a standard format for submission or transmission to the national reporting office. To customize the program for individual countries, CDC recommends contacting WHO representatives.

Spreadsheet Software

Among the most useful software for surveillance activities are spreadsheet programs, which resemble an accountant's ledger book with a large grid comprised of many rows and columns. The individual cells in the spreadsheet contain data, text, or formulas that relate the contents of one cell to others. The program allows health professionals to do a wide variety of functions without the assistance of professional programmers: data entry, editing, and analysis; modeling, problem solving, and decision analysis; preparing tables for annual reports; and generating many types of graphs. The following example will illustrate how a spreadsheet program can be used for surveillance purposes (see Ref. B).

Malaria has long been a problem in Southeast Asia. Cullen et al (6) have described an early warning system being used for malaria control in Northern Thailand for nearly a decade. The system is supposed to identify epidemic levels of the disease early enough so that remedial actions can prevent further spread. Following the scheme shown previously in Figure 1, the sensor is the local reporting of malaria cases (persons who are slide-positive for *Plasmodium falciparum* or *vivax*). The reference signal is the data flow from the field to the centralized processing unit at the district level. The monitor compares the observed number of malaria cases with the expected number based on nonepidemic monthly levels. Thus, the monitor attempts to determine if there are excess cases attributed to an epidemic. When the excess cases are detected, a reference signal is sent to the controller. The controller uses the information to initiate control measures, including the use of antimalarial drugs and vector control procedures.

A typical procedure for deriving excess cases is shown in Figure 5 with fourfold table common to epidemiology. The population in a given area is viewed during two time periods, when malaria is and is not at epidemic levels. During the epidemic period, persons living in the area who get malaria are listed in cell A while those who avoid the disease are in cell B. During the nonepidemic period the same two groups are cited in cells C and D. Thus, the total population in the area during the epidemic and nonepidemic periods is

A+B and C+D, respectively. All the persons in cell A had malaria. If the disease had not been epidemic, some of the persons in cell A might still have contracted malaria. These are expected or background cases. The expected cases, A_0, are derived by multiplying the malaria incidence rate during the nonepidemic period times the number of persons living in the region during the epidemic period. The excess cases, A_1, are the observed cases in cell A minus the expected cases, A_0. The excess cases represent the best estimate of cases attributed to the epidemic. By recognizing these excess cases at an early stage of the epidemic, the program manager can initiate control procedures to limit the size and duration of the epidemic.

Unfortunately, the calculation of excess cases used in actual surveillance programs is more complicated than the above-mentioned malaria example. Usually there is considerable variation from month to month in the occurrence of malaria. If the surveillance system is not very specific and identifies every increase as an epidemic, then control procedures will be initiated too frequently. That is, a nonepidemic rise will be falsely labeled as an epidemic (a false-positive epidemic). Conversely, if the surveillance program is too insensitive, then an actual epidemic will not be detected early enough (a false-negative epidemic). Cullen et al (6) also presented several, more realistic methods for determining excess malaria cases. One was favored, however, and is included here to illustrate the use of spreadsheet software (see Ref. B).

EXCESS CASES Record keeping has been excellent in the Li District of northern Thailand for many years, which has provided Cullen et al with accurate data on the number of malaria cases. The monthly cases for nine

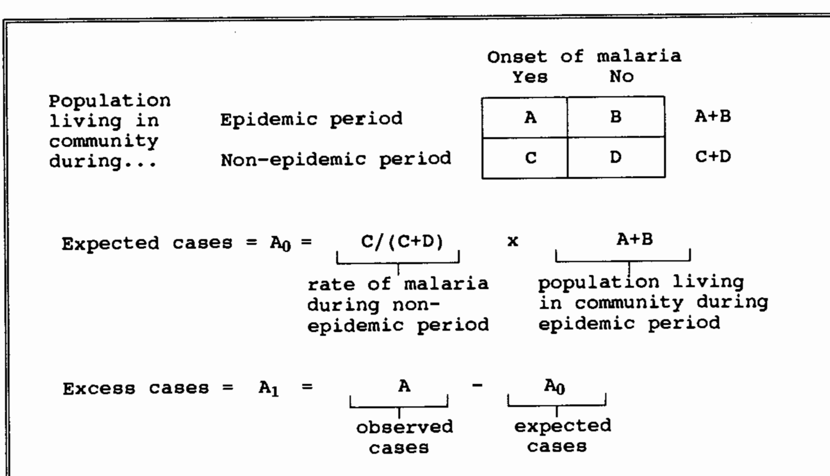

Figure 5 Determination of excess cases attributed to an epidemic.

years (1973–1981) are listed in Table 1 and shown in Figure 6. Although there was regular seasonal variation in malaria from 1973 to 1976, the peaks and valleys apparent in Figure 6 did not represent epidemic years. The data show that there were epidemics in 1977, 1978, and possibly 1981, with many more cases than expected. To derive the expected cases during 1977–1981, Cullen et al used monthly data from 1973–1976, the nonepidemic years. For each month, they determined the mean and standard deviation of cases that occurred during the four year period. Taking January as an example (see Table 1), there were 10 cases in 1973, 8 in 1974, 21 in 1975, and 24 in 1976. The mean for the four years was 15.8 cases (i.e. 63/4) with a sample standard deviation (SD) of 7.9 cases. Assuming case incidence is normally distributed, 95% of the cases per month during nonepidemic years would be expected to occur within plus or minus 1.96 SD of the mean (rounded to 2 by Cullen et al). This implies that 97.5% of the monthly malaria figures during nonepidemic months would fall below the mean plus 2 SD. Thus the mean plus 2 SD is the upper limit of monthly expected cases, which accounts for normal variation from one month to the next. Cases above this expected limit are termed excess cases.

EARLY WARNING Using the spreadsheet program, the expected and excess cases are derived for the Li District data for the years 1973–1981 (see Figures

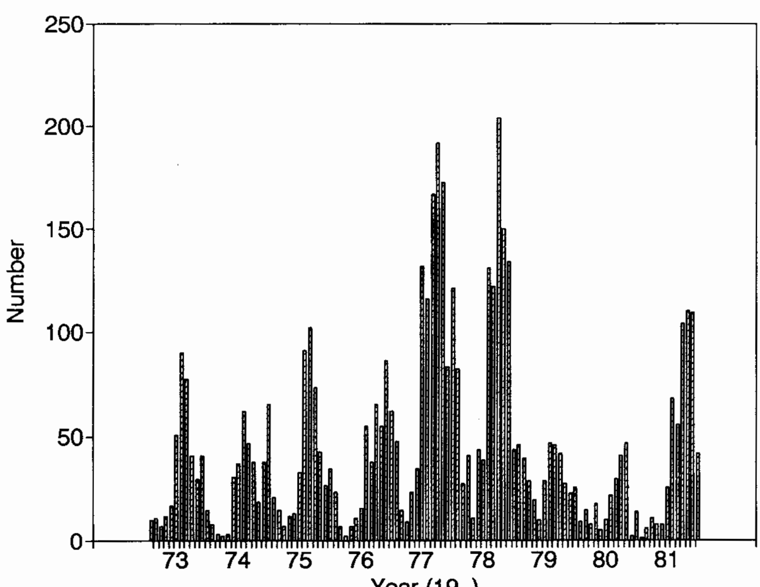

Figure 6 Slide positive malaria cases, Li District, northern Thailand, 1973–1981.

Table 1 Number of malaria cases per month for the years 1973–81 in Li district of northern Thailand (6)

Year	Jan.	Feb.	Mar.	Apr.	May	June	July	Aug.	Sep.	Oct.	Nov.	Dec.	Total
1973	10	11	7	12	17	51	90	77	41	30	41	15	402
1974	8	3	2	3	31	37	62	47	38	19	38	65	353
1975	21	15	7	12	13	33	91	102	73	43	27	35	472
1976	24	7	2	7	11	16	55	38	65	55	86	62	428
1977	48	15	9	24	35	132	116	167	192	173	83	121	1115
1978	82	28	41	11	44	39	131	122	204	150	134	44	1030
1979	46	40	29	20	10	29	47	46	42	28	23	26	386
1980	9	15	8	18	5	10	22	30	41	47	2	14	221
1981	1	6	11	8	8	26	68	56	104	110	109	42	549
Mean[a]	15.8	9.0	4.5	8.5	18.0	34.3	74.5	66.0	54.3	36.8	48.0	44.3	413.8
SD[b]	7.9	5.2	2.9	4.4	9.0	14.4	18.7	29.2	17.4	15.6	26.0	23.7	49.7
Mean + 2SD[c]	31.6	19.3	10.3	17.2	36.0	63.1	111.9	124.4	89.0	68.0	100.1	91.7	513.2

[a] Mean calculated for 1973–1978 (see text)
[b] Sample standard deviation for 1973–1978
[c] Mean plus two times the sample standard deviation for 1973–1978

A–D in the Appendix). Starting in January 1977, there were several months when observed malaria cases exceeded expection. As shown in Figure 7 there were 16 excess cases in January 1977, followed by both small and large increases in 1977–1978 and in 1981. Because the intention of a surveillance system is to provide early warning of an epidemic, a graph limited to excess cases provides a clearer reference signal for the controllers. Figure 7 shows such a graph, which includes an arrow pointing to the start of the epidemic. Figures 6 and 7 are the two graphs that would routinely be used by the malaria surveillance system.

In using this approach (although not with a spreadsheet program), Cullen et al pointed out that remedial actions could have been introduced in the first few months of 1977, rather than two years later in February 1979, as actually occurred (6). If the spreadsheet program had been available in Thailand and used to produce monthly tables and graphs, it might have been much easier to convince the local district administrator that action was needed. The ease and timeliness of producing these graphs points to the value of microcomputers for surveillance efforts.

Graphics Software

Graphics programs are another useful category of software for surveillance activities. Besides conventional line and bar graphs, which are featured in most spreadsheets programs, this software is used to display a wide variety of two- and three-dimensional graphs. The CDC uses many graphs in the

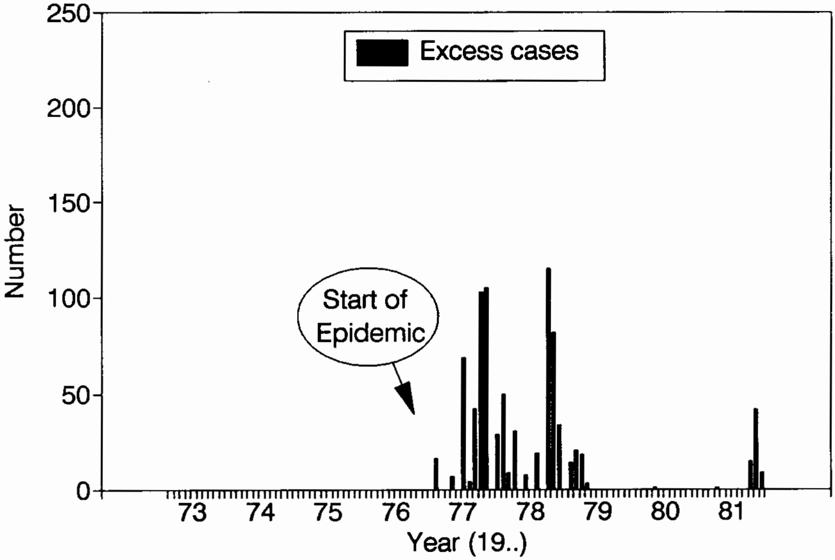

Figure 7 Excess malaria cases, Li District, northern Thailand, 1973–1981.

Morbidity and Mortality Weekly Report to notify local officials of pending health problems. Being a large agency in a wealthy country, they have computer programmers on their staff to develop specially-designed maps, graphs, and tables. Their output is generated by large computer and special graphics plotters and laser printers, none of which are available in most developing countries. Although the specific hardware and software CDC uses to display information are not readily available, commercial graphics software can often provide an acceptable substitute. Figure 8 shows a horizontal bar graph produced by one such program (see Ref. C) that is similar to the CDC surveillance graph for notifiable diseases (4). Each week, a ratio is derived of observed cases during the past four weeks to a historical average of comparable four-week periods during the prior five years. If the ratio is within normal limits, the horizontal bar for the ratio is black. If the disease has become epidemic with many excess cases, a hatched-line pattern is used to show the excess. This graph and the publication in which it appears would represent the error signal of a notifiable disease surveillance program.

A related software is the mapping program, which is used to construct and print maps of disease patterns for a given geographic area. Computer maps are a modern version of pin maps. Pins with colored heads are typically tacked in a map board to show where disease cases occurred. Different colored heads on the pins are used to designate different diseases or different groups, such as young children or males and females. Most surveillance programs in the developing world rely on both bar or line graphs and some form of pin map to show their findings. The former shows the trends over time for an entire

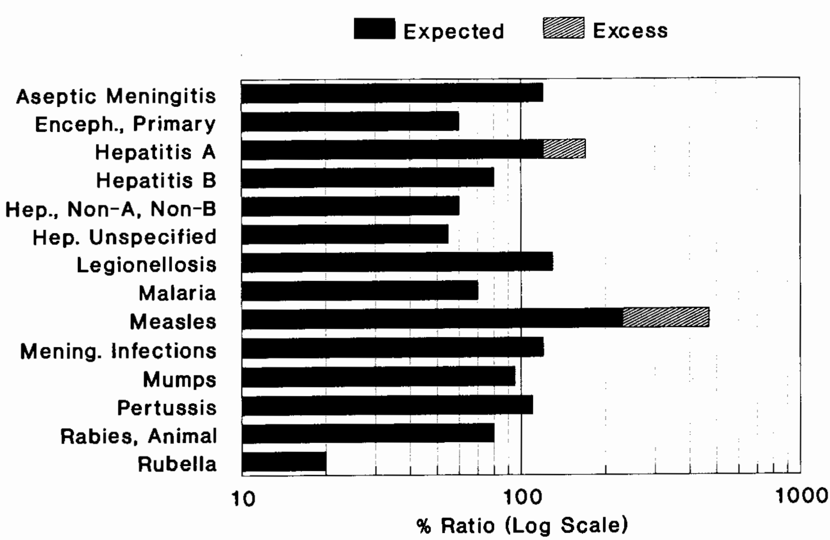

Figure 8 Expected and excess cases of notifiable diseases.

region; the latter shows the location of the individual cases. With graphics and mapping software, both figures could quickly be generated every week and shared with local decision-makers.

SAMPLE SURVEYS TO SUPPLEMENT SURVEILLANCE

Sample surveys can provide important data to supplement information derived from a surveillance system. Even when surveillance is effective, however, more information may be needed to plan and implement control efforts or to evaluate their impact on the population once the control program is in place. Surveys can provide this kind of information. Surveys may also be vital in countries that derive all their surveillance data from weekly hospital or clinic reports, rather than from a population sample. Such surveillance systems typically under-report disease cases by excluding persons who do not come to the facilities for medical assistance. Finally, surveys may provide the only source of factual information in the economically poorest countries; surveillance programs in these countries are often so inadequate that they are rarely used for decision-making. Here the governments may find that it is easier and less costly to hire and train a small survey team than to rely on a large surveillance system that cannot be effectively managed.

Large population-based surveys have been performed in many developing countries (28, 33). Examples include an extensive survey by Frerichs et al in a rural region of Bolivia (15–18), serological surveys for HIV infection in Rwanda (34), and more recently a series of demographic and health surveys being done at the national level in many developing countries by the Institute for Resource Development (27). Although these surveys do not represent surveillance efforts, they do provide useful supplementary information. Unfortunately, these large surveys of several thousand persons are costly and often require a year or more before a final report appears with the study findings.

Small, two-stage cluster surveys have been widely used by various groups in WHO to gather data on immunization coverage, diarrhea and diarrhea control procedures, neonatal tetanus, lameness (poliomyelitis), and related activities (29). These small surveys fill the information void left by limited or inadequate surveillance systems; they provide population-based data and are typically completed in a few months. The completion time was reduced even more by Frerichs and colleagues who used portable, battery-powered microcomputers to complete WHO-style surveys in a few weeks (13, 14, 21–23). Typically, they would take 3–4 days to conduct the interviews or examinations, and use a portable microcomputer to enter and edit data every afternoon or evening. On the fifth day, they present their findings in graphs and simple tables to the local medical officer and interested community members. During the following 4–5 days, they use the microcomputer and commercial software to complete the analysis, prepare tables and graphs, and produce a final written report. This form of sampling, termed "rapid survey methodology," provides yet another information source for administrators in developing countries who are faced with the time-delays and poor data quality of many surveillance programs.

UTILITY OF SURVEILLANCE PROGRAMS

For infectious disease, the early warning provided by surveillance systems allows program administrators time to start various control activities. These activities may include isolating infective cases, immunizing children, spraying pesticides, chlorinating wells, restricting food-handlers, and eliminating open sources of open water. The desired action, however, is less clear for many of the noncommunicable disease surveillance programs typically found in more developed countries.

Surveillance programs are rarely evaluated, even those programs that focus on infectious diseases. In their extensive review of surveillance in the United States, Thacker & Berkelman (37) point out that there have been no published studies on the impact of surveillance data on policies and interventions, and very few studies that relate the cost of acquiring surveillance information to any tangible benefits.

For developing countries, the need seems more apparent. However, information on the cost and usefulness of surveillance efforts is equally scarce. Smith (36) expressed this concern for surveillance programs in developing countries: "More work is . . . needed to critically evaluate the usefulness of already existing surveillance systems. If they are found not to provide useful data, then serious consideration should be given to using available resources in other ways, such as the development of sentinel sites or periodic surveys." A related concern was voiced at a recent national epidemiology conference in Chile, which concluded that the main problem with their current epidemiologic surveillance program is that it does not lead to decision-making, or else that decisions are taken too late to have an impact on the disease (31).

In the past decade, the United States Agency for International Development funded a series of methodologic studies on rapid epidemiologic assessment (REA) in developing countries. The specific intention was to develop REA methods for health planning and decision-making and to develop or validate more efficient or innovative epidemiologic methods for making timely decisions about health problems and program for health care or disease control (35).

Surveillance systems are expensive to set up and maintain. As a result, they should be evaluated along with other programs to determine if they are fulfilling their goals. This evaluation should relate the cost of the program to the effectiveness or usefulness. For most workers, timely feedback is proof enough that the system is working and that their contributions are valued. With proper graphs, they can anticipate epidemics and perceive the effects of control strategies. For administrators, program managers, and political leaders the best indication of usefulness is that epidemics are noted early and kept under control. This control is the promise of epidemiologic surveillance. However, we need to evaluate this promise.

CONCLUSION

Epidemics can devastate a population. Death and disability can drain the vitality of a community. Economic havoc follows despair as productivity decreases and even food and shelter become scarce. Surveillance systems are designed to prevent disease and empower the people by providing early warning of an epidemic. The resultant extra time can be spent battling the disease while it is still at a manageable level.

A surveillance system will be effective only if the quality of data is high, information is processed and analyzed rapidly, and the findings are clearly understood by those in power. In many developing countries, data collection and analysis are so difficult that only the most essential information should be collected: a listing of cases and deaths by cause, time of onset, and geographic area where the cases occurred. If resources are available, other related information should be considered, such as age, sex, characteristics of the agent, and circumstances leading to the disease onset. In many instances, however, the value of this additional information will not justify the added cost. When more extensive information is sought, it should be acquired in periodic community-based surveys.

Inexpensive microcomputers are being used by many national or district health departments throughout the third world. By using current computer hardware and software, surveillance data can be quickly processed and analyzed, and findings can be rapidly converted into simple, easy to understand tables and graphs. If surveillance data are believable and presented in a timely manner, administrators and program managers will be more inclined to use them to serve the community.

APPENDIX

Table 1 is entered into the spreadsheet program as numbers, text and formulas as shown in Figure A (see Ref. A). Rows 1 through 13 contain only text and numbers; rows 14 through 16 contain text and formulas. Figure B shows the same spreadsheet, but with formulae rather than calculations as shown in Figure A. The width of columns B through E was expanded to show the formulae. Elsewhere in the same spreadsheet, four long rows are created with the year in the top row, the observed malaria cases per month in the second row, the expected monthly cases in the third row, and the excess cases in the

File Edit Style Graph Print Database Tools Options Window

A1: U [W9] 'Table 3. Number of Malaria cases per month for the years 1973-81

Table 1. Number of Malaria cases per month for the years 1973-81 in Li d
of northern Thailand

Year	Jan.	Feb.	Mar.	Apr.	May	June	July	Aug.	Sep.	Oct.
1973	10	11	7	12	17	51	90	77	41	30
1974	8	3	2	3	31	37	62	47	38	19
1975	21	15	7	12	13	33	91	102	73	43
1976	24	7	2	7	11	16	55	38	65	55
1977	48	15	9	24	35	132	116	167	192	173
1978	82	28	41	11	44	39	131	122	204	150
1979	46	40	29	20	10	29	47	46	42	28
1980	9	15	8	18	5	10	22	30	41	47
1981	1	6	11	8	8	26	68	56	104	110
Mean	15.8	9.0	4.5	8.5	18.0	34.3	74.5	66.0	54.3	36.8
SD	7.9	5.2	2.9	4.4	9.0	14.4	18.7	29.2	17.4	15.6
Mean+2SD	31.6	19.3	10.3	17.2	36.0	63.1	111.9	124.4	89.0	68.0

LI.WQ1 [1] CALC READY

Figure A Image on computer monitor of spreadsheet entries.

File Edit Style Graph Print Database Tools Options Window

A1: U [W15] 'Table 3. Number of Malaria cases per month for the years 1973-8

Table 3. Number of Malaria cases per month for the years 1973-81 in Li d
of northern Thailand

Year	Jan.	Feb.	Mar.	Apr.
1973	10	11	7	12
1974	8	3	2	3
1975	21	15	7	12
1976	24	7	2	7
1977	48	15	9	24
1978	82	28	41	11
1979	46	40	29	20
1980	9	15	8	18
1981	1	6	11	8
Mean	@AVG(B4..B7)	@AVG(C4..C7)	@AVG(D4..D7)	@AVG(E4..E7)
SD	@STDS(B4..B7)	@STDS(C4..C7)	@STDS(D4..D7)	@STDS(E4..E7)
Mean+2SD	+B14+2*B15	+C14+2*C15	+D14+2*D15	+E14+2*E15

LI.WQ1 [1] CALC READY

Figure B Image on computer monitor of spreadsheet formulas.

third row (see Figure C). Excess cases are defined as 0 if the observed cases are less than the expected cases and observed cases minus expected cases if the observed cases are greater than the expected cases. The formulae for deriving the excess cases are shown in Figure D. In words the formulae read, if observed minus expected is less than 0, enter 0, otherwise enter observed minus expected.

	1973												
Observed	10	11	7	12	17	51	90	77	41	30	41	15	continued but
Expected	31.6	19.3	10.3	17.2	36.0	63.1	112.	124.	89.0	68.0	100.	91.7	shown on next
Excess	0	0	0	0	0	0	0	0	0	0	0	0	line
	1974												
	8	3	2	3	31	37	62	47	38	19	38	65	continued but
	31.6	19.3	10.3	17.2	36.0	63.1	112.	124.	89.0	68.0	100.	91.7	shown on next
	0	0	0	0	0	0	0	0	0	0	0	0	line
	1975												
	21	15	7	12	13	33	91	102	73	43	27	35	continued but
	31.6	19.3	10.3	17.2	36.0	63.1	112.	124.	89.0	68.0	100.	91.7	shown on next
	0	0	0	0	0	0	0	0	0	0	0	0	line
	1976												
	24	7	2	7	11	16	55	38	65	55	86	62	continued but
	31.6	19.3	10.3	17.2	36.0	63.1	112.	124.	89.0	68.0	100.	91.7	shown on next
	0	0	0	0	0	0	0	0	0	0	0	0	line
	1977												
	48	15	9	24	35	132	116	167	192	173	83	121	continued but
	31.6	19.3	10.3	17.2	36.0	63.1	112.	124.	89.0	68.0	100.	91.7	shown on next
	16.4	0	0	6.78	0	68.9	4.10	42.6	103.	105.	0	29.3	line
	1978												
	82	28	41	11	44	39	131	122	204	150	134	44	continued but
	31.6	19.3	10.3	17.2	36.0	63.1	112.	124.	89.0	68.0	100.	91.7	shown on next
	50.4	8.67	30.7	0	7.96	0	19.1	0	115.	82.0	33.9	0	line
	1979												
	46	40	29	20	10	29	47	46	42	28	23	26	continued but
	31.6	19.3	10.3	17.2	36.0	63.1	112.	124.	89.0	68.0	100.	91.7	shown on next
	14.4	20.7	18.7	2.78	0	0	0	0	0	0	0	0	line
	1980												
	9	15	8	18	5	10	22	30	41	47	2	14	continued but
	31.6	19.3	10.3	17.2	36.0	63.1	112.	124.	89.0	68.0	100.	91.7	shown on next
	0	0	0	.782	0	0	0	0	0	0	0	0	line
	1981												
	1	6	11	8	8	26	68	56	104	110	109	42	
	31.6	19.3	10.3	17.2	36.0	63.1	112.	124.	89.0	68.0	100.	91.7	
	0	0	.726	0	0	0	0	0	15.0	42.0	8.92	0	

Figure C Spreadsheet entries for observed, expected, and excess cases (entered but not shown as four long rows).

```
File  Edit  Style  Graph  Print  Database  Tools  Options  Window        ↑↓
P1: [W11]                                                               | ?
 J        P                  Q                      R
1                                                                       End
2   Observed                 10                     11                   ▲
3   Expected               31.6                   19.3                  ◄ ►
4   Excess      @IF((Q2-Q3)<0,0,Q2-Q3)  @IF((R2-R3)<0,0,R2-R3)           ▼
5
6                                                                       Esc
```

Figure D Spreadsheet formulas for observed, expected, and excess cases (entered as four long rows starting in cells P1 through P4; top row is year).

Literature Cited

1. Bertrand, W. E. 1985. Microcomputer applications in health population surveys: experience and potential in developing countries. *World Health Stat. Q*. 38(1):91–100
2. Bertrand, W. E. 1987. Use of microcomputers in health and social service applications in developing countries. *Crit. Rev. Med. Informatics* 1(3):229–40
3. Cent. Dis. Control. 1986. Comprehensive plan for epidemiologic surveillance. *General Surveillance Activities*. Atlanta: Cent. Dis. Control
4. Cent. Dis. Control. 1989. Proposed changes in format for presentation of notifiable disease report data. *Morb. Mortal. Wkly. Rep*. 38(47):805–09
5. Cent. Dis. Control. 1990. Update: Influenza—United States, 1989–90. *Morb. Mortal. Wkly. Rep*. 39:157–59
6. Cullen, J. R., Chitprarop, U., Doberstyn, E. B., Sombatwattanangkul, K. 1984. An epidemiological early warning system for malaria control in northern Thailand. *Bull. WHO*. 62:107–14
7. Dean, A. G., Dean, J. A., Burton, A. H., Dicker, R. C. 1990. Epi Info, Version 5.0: a word-processing, database, and statistics program for epidemiology on microcomputers. Atlanta: Cent. Dis. Control. 384 pp.
8. Evans, A. S. 1989. Surveillance and seroepidemiology. In *Viral Infections of Humans—Epidemiology and Control*, ed. A. S. Evans, 2:51–73. New York/London: Plenum. 829 pp.
9. Expand. Program Immun. 1987. Disease surveillance: Information for action. *WHO Update* Dec: 1–6. Geneva, Switzerland
10. Eylenbosch, W. J., Noah, N. D. 1988. The surveillance of disease. In *Surveillance in Health and Disease*, ed. W. J. Eylenbosch, N. D. Noah, 2:9–24. Oxford/New York/Tokyo: Oxford Univ. Press. 286 pp.
11. Eylenbosch, W. J., Noah, N. D. 1988. Introduction. See Ref. 10, pp. xv–xvii
12. Frerichs, R. R. 1985. Existing health information system. In *Burma Primary Health Care II Project Report*. USAID Project 482–004. West. Consort. Public Health, Berkeley, Calif.
13. Frerichs, R. R. 1988. Rapid microcomputer surveys. *J. Trop. Pediatr*. 34:147–49
14. Frerichs, R. R. 1989. Simple analytic procedures for rapid microcomputer-assisted surveys in developing countries. *Public Health Rep*. 104(1):24–35
15. Frerichs, R. R., Becht, J. N., Foxman, B. 1980. A household survey of health and illness in rural Bolivia. *Bull. Pan Am. Health Organ*. 14:343–55
16. Frerichs, R. R., Becht, J. N., Foxman, B. 1980. Prevalence and cost of illness episodes in rural Bolivia. *Int. J. Epidemiol*. 9(3):233–38
17. Frerichs, R. R., Becht, J. N., Foxman, B. 1981. Childbearing and breastfeeding in rural Bolivia—A household survey. *J. Trop. Pediatr*. 27:245–49
18. Frerichs, R. R., Becht, J. N., Foxman, B. 1981. Screening for childhood malnutrition in rural Bolivia. *J. Trop. Pediat*. 27:285–91
19. Frerichs, R. R., Miller, R. A. 1985. Introduction of a microcomputer for health research in a developing country—the Bangladesh experience. *Public Health Rep*. 100:638–64
20. Frerichs, R. R., Selwyn, B. J. 1991. Microcomputer applications in epidemiology. In *Oxford Textbook Public Health*, ed. W. W. Holland, R. Detels, E. G. Knox. Oxford/New York: Oxford Med. Publ. 2nd ed. In press
21. Frerichs, R. R., Tar, K. T. 1988. Use of rapid survey methodology to determine immunization coverage in rural Burma. *J. Trop. Pediatr*. 34:125–30
22. Frerichs, R. R., Tar, K. T. 1988. Breastfeeding, dietary intake, and weight-for-age of children in rural Burma. *Asian-Pac. J. Public Health* 2:125–30
23. Frerichs, R. R., Tar, K. T. 1989. Computer-assisted rapid surveys in developing countries. *Public Health Rep*. 104(1):14–23
24. Ghana Health Assess. Proj. Team. 1981. A quantitative method of assessing the health impact of different diseases in less developed countries. *Int. J. Epidemiol*. 10:73–80
25. Gould, J. B., Frerichs, R. R. 1986. Training faculty in Bangladesh to use a microcomputer for public health work: a follow-up report. *Public Health Rep*. 101:616–23
26. Health Inf. Serv. 1985. Hospital inpatient morbidity and mortality data, 1982. *Burma Health Stat. Rep*. Dep. Health, Minist. Health, Rangoon, Burma. 73 pp.
27. Inst. Resour. Dev. 1990. Status of demographic and health surveys as of March 1990. *Demogr. Health Surv. Newsl*. 3(1):6–7
28. Kroeger, A. 1983. Health interview surveys in developing countries: a review of the methods and results. *Int. J. Epidemiol*. 12(4):465–81

29. Lemeshow, S., Robinson, D. 1985. Surveys to measure programme coverage and impact: a review of the methodology used by the Expanded Programme on Immunization. *World Health Stat. Q.* 38:65–75
30. Lwanga, S., (Chairman). 1978. Statistical principles of monitoring and surveillance in public health. *Bull. WHO* 56:713–22
31. Pan Am. Health Organ. 1990. First Chilean meeting on epidemiology. *Epidemiol. Bull.* 10(4):1–16
32. Romeder, J-M., McWhinnie, J. R. 1977. Potential years of life lost between ages 1 and 70: an indicator of premature mortality for health planning. *Int. J. Epidemiol.* 6:143–51
33. Ross, D. A., Vaughan, J. P. 1986. Health interview surveys in developing countries: a methodological review. *Stud. Fam. Plann.* 17:78–94
34. Rwandan HIV Seroprevalence Study Group. 1989. Nationwide community-based serological survey of HIV-1 and other human retrovirus infections in a Central African Country. *Lancet* 1:941–43
35. Selwyn, B. J., Frerichs, R. R., Smith, G. S., Olson, J. 1989. Rapid epidemiologic assessment: the evolution of a new discipline—introduction. *Int. J. Epidemiol.* 18(Suppl. 2):S1
36. Smith, G. S. 1989. Development of rapid epidemiologic assessment methods to evaluate health status and delivery of health services. *Int. J. Epidemiol.* 18 (4):S2–S15
37. Thacker, S. B., Berkelman, R. L. 1988. Public health surveillance in the United States. *Epidemiol. Rev.* 10:164–90
38. Tugwell, P., Bennett, K. J., Sackett, D., Haynes, B. 1985. Relative risks, benefits, and cost of intervention. In *Tropical and Geographical Medicine,* ed. K. S. Warren, A. A. F. Mahmoud, 124:1097–13. New York: McGraw-Hill. 1175 pp.
39. Walsh, J. A. 1986. Prioritizing for primary health care: methods for data collection and analysis. In *Strategies for Primary Health Care,* ed. J. A. Walsh, K. S. Warren. Chicago/London: Univ. Chicago Press. 330 pp.
40. Walsh, J. A., Warren, K. S. 1980. Selective primary health care: an interim strategy for disease control in developing countries. *Soc. Sci. Med.* 14C:145–63
41. WHO. 1982. *Wkly. Epidemiol. Rec.* 57:361–68
42. WHO. 1990. Nutritional status of Somali refugees in Eastern Ethiopia, September 1988–May 1989. *Wkly. Epidemiol. Rec.* 65(13):93–100

SOFTWARE

A. Epi Info, Version 5. 1990. Div. Surveill. Epidemiol. Stud., Epidemiol. Program Off., Cent. Dis. Control, Atlanta, Ga. 30333
B. Quattro Pro, Version 2. 1990. Borland Int. Inc., 1800 Green Hills Road, P.O. Box 660001, Scotts Valley, Calif., 95067-0001
C. Harvard Graphics, Version 2.3. 1990. Software Publ. Corp., P.O. Box 7210, 1901 Landings Dr., Mountain View, Calif., 94039-7210