The World Health Organization is a specialized agency of the United Nations with primary responsibility for international health matters and public health. Through this organization, which was created in 1948, the health professionals of some 160 countries exchanged their knowledge and experience with the aim of making possible the attainment by all citizens of the world by the year 2000 of a level of health that will permit them to lead a socially and economically productive life.

By means of direct technical cooperation with its Member States, and by stimulating such cooperation among them, WHO promotes the development of comprehensive health services, the prevention and control of diseases, the improvement of environmental conditions, the development of health manpower, the coordination and development of biomedical and health services research, and the planning and implementation of health programs.

These broad fields of endeavor encompass a wide variety of activities, such as developing systems of primary health care that reach the whole population of Member countries; promoting the health of mothers and children; combating malnutrition; controlling malaria and other communicable diseases including tuberculosis and leprosy; having achieved the eradication of smallpox, promoting mass immunization against a number of other preventable diseases; improving mental health; providing safe water supplies; and training health personnel of all categories.

Progress towards better health throughout the world also demands international cooperation in such matters as establishing international standards for biological substances, pesticides and pharmaceuticals; formulating environmental health criteria; recommending international nonproprietary names for drugs; administering the International Health Regulations; revising the International Classification of Diseases, Injuries, and Causes of Death; and collecting and disseminating health statistical information.

Further information on many aspects of WHO's work is presented in the Organization's publications.

The World health statistics quarterly replaces (since 1978) the monthly World health statistics report (published since 1947). It deals with the detailed analysis of selected health topics of current interest. Starting with Vol. 41 (1988), the Quarterly contains articles in either French or English with a summary in both languages.

Symbols used in tables

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>*</td>
<td>Preliminary, approximate or estimated data.</td>
</tr>
<tr>
<td>...</td>
<td>Data not available.</td>
</tr>
<tr>
<td>–</td>
<td>Nil or magnitude negligible.</td>
</tr>
<tr>
<td>–</td>
<td>Category not applicable.</td>
</tr>
<tr>
<td>T</td>
<td>Total.</td>
</tr>
<tr>
<td>M</td>
<td>Male.</td>
</tr>
<tr>
<td>F</td>
<td>Female.</td>
</tr>
<tr>
<td>N</td>
<td>Absolute numbers.</td>
</tr>
</tbody>
</table>

Explication des signes

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>*</td>
<td>Donnée préliminaire, approximative ou estimative.</td>
</tr>
<tr>
<td>...</td>
<td>Donnée non disponible.</td>
</tr>
<tr>
<td>–</td>
<td>Zéro ou quantité négligeable.</td>
</tr>
<tr>
<td>–</td>
<td>Catégorie non applicable.</td>
</tr>
<tr>
<td>T</td>
<td>Total.</td>
</tr>
<tr>
<td>M</td>
<td>Masculin.</td>
</tr>
<tr>
<td>F</td>
<td>Féminin.</td>
</tr>
<tr>
<td>N</td>
<td>Nombres absolus.</td>
</tr>
</tbody>
</table>

© World Health Organization 1991

© Organisation mondiale de la Santé 1991
## EPIDEMIOLOGICAL AND STATISTICAL METHODS FOR RAPID HEALTH ASSESSMENT

<table>
<thead>
<tr>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Epidemiological and statistical methods for rapid health assessment: introduction. Martha Anker</td>
<td>94</td>
</tr>
<tr>
<td>A simplified general method for cluster-sample surveys of health in developing countries. Steve Bennett, Tony Woods, Winitha M. Liyanage &amp; Duane L. Smith</td>
<td>98</td>
</tr>
<tr>
<td>Evaluation of the EPI survey methodology for estimating relative risk. Donald R. Harris &amp; Stanley Lemeshow</td>
<td>107</td>
</tr>
<tr>
<td>Lot quality assurance sampling: single- and double-sampling plans. Stanley Lemeshow &amp; Scott Taber</td>
<td>115</td>
</tr>
<tr>
<td>Lot quality assurance sampling techniques in health surveys in developing countries: advantages and current constraints. Claudio F. Lanata &amp; Robert E. Black</td>
<td>133</td>
</tr>
<tr>
<td>The potential of the case-control method for rapid epidemiological assessment. Jane C. Baltazar</td>
<td>140</td>
</tr>
<tr>
<td>The use of focus groups in social and behavioural research: some methodological issues. Mohammed Ejazzudin Khan, Martha Anker, Bella C. Patel, Sandhya Barge, Hemlatha Sadhwani &amp; Ranjana Kohle</td>
<td>145</td>
</tr>
<tr>
<td>The value of questionnaires aimed at key informants, and distributed through an existing administrative system, for rapid and cost-effective health assessment. Christian Lengeler, Hassan Mshinda, Donald de Savigny, Peter Kilima, Dominique Morona &amp; Marcel Tanner</td>
<td>150</td>
</tr>
<tr>
<td>The benefits of the application of geographical information systems in public and environmental health. Henk J. Scholten &amp; Marion J.C. de Lepper</td>
<td>160</td>
</tr>
<tr>
<td>Rapid assessment of health needs in mass emergencies: review of current concepts and methods. Debarati Guha-Sapir</td>
<td>171</td>
</tr>
</tbody>
</table>

## MÉTHODES ÉPIDÉMIOLIQUES ET STATISTIQUES POUR UNE ÉVALUATION RAPIDE DE LA SITUATION SANITAIRE

<table>
<thead>
<tr>
<th>Title</th>
<th>Pages</th>
</tr>
</thead>
<tbody>
<tr>
<td>[Méthodes épidémiologiques et statistiques pour une évaluation rapide de la situation sanitaire: introduction] [anglais seulement]. Martha Anker</td>
<td>94</td>
</tr>
<tr>
<td>Méthodes générales simplifiées pour les enquêtes sanitaires utilisant le sondage par groupes dans les pays en développement [résumé]. Steve Bennett, Tony Woods, Winitha M. Liyanage &amp; Duane L. Smith</td>
<td>105</td>
</tr>
<tr>
<td>Evaluation de la méthodologie d’enquête du PEV concernant l’estimation du risque relatif [résumé]. Donald R. Harris &amp; Stanley Lemeshow</td>
<td>114</td>
</tr>
<tr>
<td>Echantillonnage par lots pour l’assurance de la qualité: plans de sondage à un et à deux degrés [résumé]. Stanley Lemeshow &amp; Scott Taber</td>
<td>132</td>
</tr>
<tr>
<td>Le potentiel de la méthode cas-témoins pour une évaluation épidémiologique rapide [résumé]. Jane C. Baltazar</td>
<td>143</td>
</tr>
<tr>
<td>L’utilisation des groupes focaux en recherche sociale et comportementale: problèmes méthodologiques [résumé]. Mohammed Ejazzudin Khan, Martha Anker, Bella C. Patel, Sandhya Barge, Hemlatha Sadhwani &amp; Ranjana Kohle</td>
<td>149</td>
</tr>
<tr>
<td>L’utilité des questionnaires destinés aux répondants clés et distribués dans le cadre d’un système administratif existant, pour une appréciation rapide et rentable de la situation sanitaire [résumé]. Christian Lengeler, Hassan Mshinda, Donald de Savigny, Peter Kilima, Dominique Morona &amp; Marcel Tanner</td>
<td>158</td>
</tr>
<tr>
<td>Les avantages de l’utilisation des systèmes d’information géographique en santé publique et en hygiène de l’environnement [résumé]. Henk J. Scholten &amp; Marion J.C. de Lepper</td>
<td>169</td>
</tr>
<tr>
<td>Evaluation rapide des besoins sanitaires en situation d’urgence de grande ampleur: méthodes et concepts actuels [résumé]. Debarati Guha-Sapir</td>
<td>180</td>
</tr>
</tbody>
</table>
EPIDEMIOLOGICAL AND STATISTICAL METHODS FOR RAPID HEALTH ASSESSMENT

INTRODUCTION

Martha Anker

Health managers, especially in developing countries, need timely and accurate information on which to base their decisions. Yet many of the traditional epidemiological and statistical methods are not very well suited to an environment with extremely limited financial resources and with few people skilled in data collection and analysis.

For this reason, there is great interest in adapting traditional methods to conditions existing in countries, and in finding suitable alternatives to traditional methods. Since many of these new and adapted methods emphasize speed of execution, they have become known as methods of "rapid assessment". Rapid assessment refers to a broad collection of epidemiological, statistical and anthropological techniques which aim to provide accurate information quickly, at a low cost, in a simple format. There is special emphasis on providing information that can be useful at the local level.

It is easy to see why rapid assessment methods would be appealing to programme managers, who need information to monitor and improve the performance of the health-care system and the health of the population. However, rapid assessment methods should not be used indiscriminately; nor should they be considered as "quick and dirty" techniques to be carried out in a sloppy manner. Rapid assessments, appropriately used and carefully carried out, can result in accurate appraisals of the situation in a relatively short time. It goes without saying that no matter what methods are chosen for rapid assessment, it is important that the assessments be carefully executed. This is true in any study, but it is even more important for rapid assessments, since they are rarely subject to rigorous scientific review before being acted upon.

Because rapid assessment methods are relatively new in their application to health assessment, and because many of them sacrifice some statistical precision for the sake of speed and simplicity, it is important to understand the strengths and weaknesses of each method, in order to provide accurate information for policy decisions. Many of the methods proposed for rapid assessment are not yet on a firm scientific footing, and need to be further developed and evaluated. Considerable methodological work is still necessary if these techniques are to provide a sound basis for programme planning, monitoring and evaluation.

In view of the increasing interest in rapid assessment techniques, both within WHO (many WHO programmes have been actively involved in developing and using these techniques) and in Member States, there was a strongly-felt need to examine the existing methodologies of rapid assessment and to underline the uses and limitations of each method, with a view to encouraging the development of those methods which appear to be the most promising. Therefore in November 1990, WHO's Division of Epidemiological Surveillance and Health Situation and Trend Assessment organized an informal consultation on rapid epidemiological and statistical methods of health assessment, involving WHO staff members interacting with experts in the field of rapid assessment methodologies. Background papers for this consultation were prepared by WHO staff members, as well as outside experts. A selection of these papers is presented in this issue of the World health statistics quarterly.

The background papers for this consultation emphasized a number of themes common to many of the methods: (i) the need to clearly spell out the strengths and limitations of each method so that it is used only when appropriate; (ii) the need for further work in developing and validating new methods; (iii) the importance of presenting results to decision makers in a clear and understandable manner; and (iv) the importance of high standards of execution, so that the results can be effectively used by decision makers.

Literature on methodological aspects of rapid assessment is scarce and widely scattered, a notable exception being a 1989 supplement to the International journal of epidemiology entitled "Rapid epidemiologic assessment: the evolution of a new discipline" (1). During the consultation, participants repeatedly recommended that WHO intensify its dissemination of information on appropriate rapid assessment methodologies. For this reason, the current issue of the Quarterly is dedicated to epidemiological and statistical methods of rapid health assessment. The emphasis is on methodology — both theoretical and practical aspects. In one instance, there are complementary articles on the same rapid assessment technique: one dealing with methodology, the other with practical considerations. Articles emphasize the strengths and limitations of each method, as well as further work that needs to be done, or is being done, on validation, simplification, and further development of the methods.

Topics

Articles in this issue address two main topics. First a number of articles are concerned with sampling methods that reduce the time and resources required to collect and analyse data from individuals.

They deal in particular with extending the survey sampling technique used by the Expanded Programme on Immunization (EPI) to other areas, with applying lot quality assurance sampling, and with the application of case-control methodology to rapid health assessment.

* Statistician, Division of Epidemiological Surveillance and Health Situation and Trend Assessment, World Health Organization, Geneva.

Rapp. trimest. statist. sanit. mond., 44 (1991)
The second set of articles deals with the collection, organization and analysis of data at the community level (or at a higher level of aggregation). This includes methods of gathering data at the community level such as focus-group interviews, or using key informants, as well as methods of organizing, analysing and presenting community-level data such as geographical information systems.

The special problems of carrying out rapid assessment in emergency situations are addressed in the final article.

**Sampling methods for rapid health assessment**

EPI developed a rapid method for cluster sampling that is well suited to settings in which sampling frames are not readily available. Instead of developing a sampling frame for each cluster, and then choosing the sample randomly (as is usually done in traditional cluster sampling), the EPI method requires that one house be chosen at random and that subsequent houses be chosen according to fixed easy-to-follow rules. This method is used widely for sampling children of a particular age group to determine the proportion who have been vaccinated (2). Because of its success, this methodology has been extended to other situations. Two of the articles address the appropriateness of extending the EPI methodology for broader use.

The article by Bennett et al. describes a cluster-sampling approach for health surveys that retains as far as possible the simplicity of the EPI strategy. The article indicates how the effects of clustering can be taken into account for the estimation of means, rates and proportions and their associated standard errors, an aspect which is often forgotten or ignored during rapid assessment. In short, this article presents the scientific considerations which should be taken into account when using this method.

The second article on the EPI cluster-sampling technique investigates its usefulness in estimating relative risk (Harris & Lemeshow). Computer simulation indicated that the EPI method of cluster sampling produces relatively accurate results for estimating the proportions of children vaccinated (3). This article describes a simulation technique that was used to create artificial populations with characteristics typical of areas of Africa in terms of factors believed to be important determinants of the risk of HIV infection. The computer program selected samples in two ways: according to simple random sampling techniques and according to EPI sampling techniques. Estimated relative risks of HIV infection from both sets of samples were compared to the values from the underlying population. Only small differences were found between methods, indicating that the EPI method could be used to estimate the relative risk of HIV infection. There are two aspects of this article worth highlighting. Firstly, the simulation-model technique which is described in some detail is important in itself, because it allows for experimentation with different methods of sampling under many different circumstances. Secondly, the result of the validation is important since it paves the way for a new utilization of the EPI technique — namely the evaluation of relative risk. A general conclusion that can be drawn from the two articles on EPI methodologies is that the method can be used in a wider context than just vaccination coverage. Further work in modifying and validating the EPI methodology for other applications would seem promising.

Two articles address the techniques of lot quality assurance sampling. This technique, although first developed for industry, has recently been applied to health assessment. It has the potential advantage of requiring a relatively small sample; however, the sampling technique required is somewhat more complicated than simple random sampling or cluster sampling. One article outlines the basics of lot quality assurance sampling (LQAS) (Lemeshow & Taber) and the second article (Lanata & Black) presents an example of the application of this sampling methodology in a field setting, indicating the type of operational difficulties to be expected when carrying out lot quality assurance sampling. The article by Lanata & Black describes its use in monitoring and evaluating health programmes, and suggests that it is a useful method for allowing programme managers to pinpoint weaknesses in the health system and to allocate scarce resources where they are most needed.

The general conclusion drawn from the articles and from the discussions at the consultation indicate that because it requires small sample sizes, LQAS would be a worthwhile cost-effective method for rapid assessment in a number of situations.

However, the method has a number of limitations. It is based on a hypothesis-testing strategy rather than an estimation strategy. Due to small sample sizes, LQAS cannot be used for estimating point prevalence for small areas. It can only pinpoint areas which fall above or below a certain target. In addition, the sample design for LQAS (in particular for double sampling) is relatively complicated and requires considerable expertise to obtain the optimal sample sizes. In addition, LQAS may not necessarily be rapid, as it requires a sampling frame for each lot. If the sampling frame is difficult or expensive to obtain, LQAS will probably not be a rapid approach unless the same design is used periodically for monitoring purposes (in which case the repeated sampling of the lot makes the initial investment in developing a sampling frame worthwhile). In short, it seems that the small sample sizes required by LQAS make it an attractive, cost-effective method for use in a number of specific situations. But it is not necessarily cost-effective in all circumstances. More work is needed in testing LQAS in order to further develop useful applications of this method.

The case-control methodology provides another potential method of sampling for rapid assessment (by choosing appropriate samples of cases and controls). It differs greatly from the other techniques discussed in this issue, as the case-control method involves a retrospective study. Case-control methodology began as a means of identifying the risk factors of chronic diseases as an alternative to carrying out large-scale prospective studies. The article by Baltazar discusses the use of case-control methodology for rapid assessments, with particular reference to water and sanitation interventions in the

---

* A simplified general method for cluster-sample surveys of health in developing countries, page 98.
* Lot quality assurance sampling: single- and double-sampling plans, page 118.
* Lot quality assurance sampling techniques in health surveys in developing countries: advantages and current constraints, page 133.
* The potential of the case-control method for rapid epidemiological assessment, page 140.
Philippines. It includes methodological considerations involved in designing and carrying out case-control studies, highlighting both the strengths and limitations of the methodology. The general conclusion to be drawn from this article is that the use of case-control methodology for rapid assessment is still in the early stages of development. The method has a good deal of potential as it is much less time-consuming to carry out a case-control study than a traditional prospective study. However, designing a case-control study is not simple, and requires inputs from researchers well trained in epidemiology or biostatistics. Efforts are now being made to develop standardized protocols for case-control studies evaluating water and sanitation interventions. This would allow interventions to be evaluated in a similar way in several areas, and would reduce the burden of designing a different case-control study for each setting.

Collection, organization and presentation of aggregate-level data

Several articles in this issue deal with innovative methods of collecting, organizing and presenting aggregate-level data. Collecting data on an aggregate level can be time-saving because data are collected from a few individuals for the group as a whole. There are numerous ways of collecting data on an aggregate level, using groups of experts, groups of community leaders, groups of individuals from the community at large, in-depth interviews with key informants, to name a few. Two articles deal with methods of collecting information at the community or village level, the article by Khan et al. on focus-group interviews, and the one by Lengeler et al. on using questionnaires aimed at key informants to investigate local perceptions.

Focus-group discussions are receiving increased attention as a method of collecting community-level information. A focus group is a structured in-depth discussion with a small group of people from the population under study. Focus groups yield considerable information, and can be very useful in investigating people’s knowledge, behavior and attitudes, especially when discussing their experiences and feelings freely and openly. These methods, originally developed for investigating consumer motivations, have recently been used more and more in the health field. The article details the practical considerations of conducting and analyzing focus-group discussions in field settings, particularly in rural and slum areas, and highlights the need for further validation of this methodology. It concludes that focus groups can yield a good deal of important information when used in conjunction with more traditional statistical methods. However, it cautions against the use of focus-group discussions as a stand-alone method.

An experiment in gathering community-level data on local perceptions and priorities through community-level questionnaires is described in the article by Lengeler et al.9 The method involves using existing administrative systems (such as the school system or the political party system) to distribute simple, self-administered questionnaires aimed at soliciting the perceptions of key informants about diseases which are common and important in their community. Results from the questionnaires were then validated by epidemiological information, and it was found that the responses from schoolteachers on the prevalence of schistosomiasis in the community were adequate to pinpoint villages with high levels of the disease. There are two aspects of this article worth highlighting. The first is the use of existing administrative systems to distribute the questionnaires. This greatly facilitates the logistics of carrying out the study, and could encourage wider participation than if ordinary mail were used. The second aspect is that the questionnaire focused on local perceptions and priorities. Basing decisions about disease control on local perceptions and priorities is important, since control measures will be more likely to succeed if they deal with a problem considered important by the local residents themselves.

This methodology is still in the validation stage. It seems to be working well for pinpointing areas with high levels of schistosomiasis. It would be interesting to test it for other diseases with easily recognizable symptoms where information on prevalence is lacking at the central level.

Geographical information systems (GIS) can be used to organize, analyze and present data at the community level. Geographical information systems run the gamut from very sophisticated, well-developed systems requiring substantial inputs in terms of data and expensive equipment, to simple systems run on microcomputers, using economical, user-friendly software. The article by Scholten gives an overview of the types of geographical information systems currently in use, and their applicability to health research. The conclusion to be drawn from this article is that “the application of GIS is making significant contributions to public health because of its efficiency, integration and presentation of information”. A large-scale GIS cannot be considered as a rapid assessment method in itself, and much of its potential lies in its ability to integrate large amounts of data, and to provide epidemiological insights which cannot be obtained easily by other means. However, some aspects of GIS are worth highlighting in the context of rapid assessment. Firstly, once the initial investment of setting up the GIS has been made, information can be retrieved rapidly; this can be very useful for allowing the health sector to respond quickly, especially in the case of environmental accidents. Secondly, the facility of a small-scale GIS to present data in a map format can be tapped for rapid assessment. The fact that the maps are easy to produce and understand makes them attractive tools, providing structure and organization to the data.

It is clear from the consultation and from the article by Guha-Sapir that rapid assessment is particularly important for resource management during emergency situations. Information is needed quickly; errors are costly in terms of both health and resources. At present, however, methods of rapid assessment in emergency situations are grossly inadequate. More work is needed to establish frameworks for which essential information becomes

---

9 The use of focus groups in social and behavioral research: some methodological issues, page 145.
10 The value of questionnaires aimed at key informants, and distributed through an existing administrative system, for rapid and cost-effective health assessment, page 150.
11 The benefits of the application of geographical information systems in public and environmental health, page 160.
The following general conclusions can be drawn from discussion in the consultation and the background papers.

1. Rapid assessment appraisals are here to stay—it is not a passing fancy. Many programmes in WHO, and many national and international organizations as well as government ministries involved in health planning and appraisal, are using rapid assessment approaches; many others would like to use them. In such a situation it becomes very important for health statisticians to ensure that the rapid assessment methods being used have a sound scientific basis, in order that in turn the key information required to estimate needs more accurately can be better formulated.

2. Several of the most useful rapid assessment methods involve extensions or modifications of traditional statistical techniques. For example, the EPI method of cluster sampling modified the traditional methods of cluster sampling to make sample surveys easier to conduct in developing country settings where accurate sampling frames are difficult to obtain. LOAS, a traditional method used in industry for quality control, is now being modified and applied to monitoring health programme performance. Case-control studies, traditionally used to study rare diseases, are now being adapted to evaluate interventions. More work is required on the possibility of extending traditional methods to meet the information needs of health management in developing countries.

3. Several of the rapid assessment methods being used are not yet on a firm scientific footing. For example, the use of questionnaires aimed at key informants to obtain information about local priorities has produced promising results but is still in the validation process. It needs to be tested in various settings before it can be used with confidence as a tool for health management.

4. The qualitative rapid assessment methods, such as focus-group discussions, can complement quantitative methods by adding depth and insight, but it may be dangerous to use them as a stand-alone method for policy makers.

5. The most appropriate rapid assessment technique to use in a particular situation is partly dependent on the amount of time available to analysts and programme managers when making their decisions. In emergencies, for example, the time constraint is extremely severe, and special methods are required to deal with those situations.

6. High standards and scientific objectivity are critical for any assessment process. The need for rapid results is not an acceptable excuse for quick and dirty work.

Finally, rapid assessments have the potential for providing programme managers with necessary information. Many methods are still in relatively early stages of development. There is great scope for developing and testing methodologies. The process of testing and validating these methods is crucial, in order that they can confidently take their place alongside the traditional methods of health assessment.

**REFERENCES**


---

A SIMPLIFIED GENERAL METHOD FOR CLUSTER-SAMPLE SURVEYS OF HEALTH IN DEVELOPING COUNTRIES

Steve Bennett,* Tony Woods,† Winitha M. Liyanage‡ & Duane L. Smith‡

1. Introduction

In order to monitor the health status of the population and to evaluate the use and effectiveness of disease protection and control measures, up-to-date information is required. In developing countries in particular, the information needed is often provided by means of cross-sectional surveys. An example of such a survey is that developed by the Expanded Programme on Immunization (EPI) of the World Health Organization (WHO) (1, 2, 3) to estimate vaccination status among young children. This scheme is a type of cluster sampling, in which a sample of 30 clusters (villages or the like) is selected and 7 children of the required age are selected in each cluster. The scheme was designed to allow the estimation of vaccination status to within ±10 percentage points and achieves this aim very well (1, 3). It has been used for its intended purpose of estimating vaccination coverage in many parts of the world (1).

Such a cluster-sample design is the only practical solution for most surveys, where the idea of taking a simple random sample of individuals across the country would be practically impossible. The EPI design is appealing in its simplicity, and has been extended to other health surveys, where the aims were different. Sometimes the cluster-sampling scheme or the sample size have been modified to take account of the objectives of the new survey (4) but at other times the “30×7” design has been adopted uncritically. A sample size which is adequate to estimate vaccination status to within 10 percentage points will not be adequate if a more precise estimate is needed, or if a comparatively rare event like mortality is being studied. Single-stage cluster sampling may be quite unsuitable for a survey in which estimates are required for separate regions of the country.

In Section 2, we outline some of the concepts used in this article. Section 3 describes the selection of the sample and Section 4 discusses criteria of sample size. The analysis of data is described in Section 5 and some extensions to the basic design are considered in Section 6.

We shall consider the sampling and statistical aspects of such surveys: the sample design and selection method, the size of the sample and the estimation of standard errors. There are many excellent textbooks which describe complex designs and appropriate formulae for their analysis (5, 6), but a certain level of expertise is needed to make the most of these, and this is often not available to workers in the field. Many of the ideas in this article have been discussed in the context of EPI surveys (2) and have been used in guidelines produced for particular surveys by WHO and other organizations (7, 8), but these may not be readily available. The monograph by Lerneshow et al. (9) covers some of these issues in detail, and many of the points made here have also been discussed recently, by Frerichs & Tar (10) and Frerichs (11), who present a practical scheme for a rapid health survey making use of microcomputers, with a more specific sampling design. Details of other practical aspects of survey methodology such as field organization, questionnaire design, etc., may be found in a number of books (12, 13).

A need for “further research into possible alternatives to the currently-used 30×7 EPI survey” has been expressed (2) and the aim of this article is to present a more general approach to the design of cross-sectional health surveys, while retaining as far as possible the simplicity of the EPI strategy.

2. Aims and concepts

It is important in any survey to set out clearly in advance the aims of the investigation. This is particularly important in deciding the sampling strategy and the size of sample to be taken. The principal aim of the study will implicitly define the basic sampling unit or BSU (also known as the ultimate sampling unit (7), or listing unit (5)). For example, in an EPI survey the principal aim may be to measure the vaccination status of children aged between 12 and 23 months. In this case the BSU is the child aged 12-23 months: the sample size is determined in terms of numbers of these “index” children. Interviewers are instructed to visit sufficient households to achieve this number, and only to carry out interviews in households in which an index child is found. This is fine as long as the study is restricted to matters directly concerning children aged 12-23 months, but if the purpose of the survey is expanded to also ascertain for example the use of oral rehydration therapy for children aged 0-5, then the sample of such children may be unrepresentative because it will only comprise those who live in households containing a child aged 12-23 months.

Most surveys have multiple aims, and for this reason should be expected to use the household as the BSU. The only exception to this would be surveys which clearly are focused only on one specific type of individual, and do not involve other members of the household, except as they affect the individual under study. Even when this is the case, there are good reasons why the BSU should still be
the household. Sample-size calculations may be carried out in terms of the number of individuals of a particular type needed, and then translated into an approximate number of households. The term "household" may be interpreted according to local conditions; a convenient definition may be "those whose food is prepared by the same person".

For households there may exist a sampling frame, or list from which the sample may be drawn. If one approximate number of households can usually be established for choosing households one by one. Such a sampling frame is likely not to exist for BSUs other than households. It would be rare to find health records which are so complete and up-to-date that they contain the current population of children aged 12-23 months for example.

A survey will collect data on many different items, and most frequently its results will be presented in terms of rates which are the ratio of two counts. An example of this would be the estimation of usage of a health centre by children aged 5-14, which might be estimated in an appropriate sample by:

\[
\text{Number of children aged 5-14 in sample} = \frac{\text{Number of children aged 5-14 in sample}}{\text{Number of children aged 5-14 in sample}}
\]

In a survey in which the household was the BSU, not only the numerator of this ratio (the number of children who have visited a health centre), but also the denominator (the number of children aged 5-14 in the sample), would be an unknown quantity until the survey had been carried out. Both would be different if a different sample of households had been selected.

Finally, it should be noted that we shall use the term cluster in its standard sampling sense to mean a natural grouping within the population, such as a village, district or other community, from which a subsample may be selected, and not in its EPI usage as that subsample itself. Although we talk in terms of "communities" the reader may interpret this as villages, urban blocks or enumeration districts or whatever grouping is appropriate.

### 3. Selecting the sample

Selection of the sample may be done in several stages: for example a country may be split into regions, a number of districts chosen from each region, a few communities from each district and a number of households from each community. However, the basic principles for deciding sample size and structure and the methods for estimating rates and their standard errors are the same. They will be demonstrated first for the simplest situation where a selection of communities is made directly within some country (or region), and estimates are obtained for that country.

The extension to several stages of sampling is straightforward and is described in Section 6. The number of communities and households to be chosen will be discussed in Section 4. Here we only discuss how the selection should be made.

**Selection of communities**

The strategy used for the selection of communities is the same as that used in the EPI method. It will be necessary to have a list of all the communities in the region where the survey is to take place. Some approximate measure of the number of households in each community is also necessary. If one can assume that the mean size of household will not vary greatly from one community to another, then any general measure of community population size will do. The relative size of the communities is more important than their absolute size, so even an out-of-date census will be adequate if some allowance is made for known variations in population growth rate since then. If some communities are too small to provide an adequate sample of households, they should be combined with other neighbouring communities before making the list.

Selection of a sample of communities is then performed by sampling with probability proportional to size (PPS). As in the EPI methodology, this is carried out by creating a cumulative list of community populations and selecting a systematic sample from a random start. For example, suppose it is required to take a sample of three communities from the list of 10 communities shown in Box 1. Divide the total population of the communities (6 700) by the number of communities to be selected (3) to obtain the sampling interval (6 700/3 = 2 233). Choose a random number between 1 and 2 233. Suppose this number is 1 814. This should be fitted into position in the list to identify the first community in the sample. Since 1 814 lies between 1 601 and 1 900, community 4 will be chosen. Now add the sampling interval to the initial random number: 1 814 + 2 233 = 4 047, and so community 6 is chosen. Add the sampling interval again: 4 047 + 2 233 = 6 280 and community 10 is chosen.

This procedure leads to communities being selected with probability proportional to size. It is desirable if, in addition, a constant number of households is selected within each chosen community. Then, overall, each household in the population will have an equal probability of being in the sample. Such a sampling procedure is said to be self-weighting and leads to the simplified formulae for analysis given in Section 5. If some other scheme is used it is unlikely that the sample will be self-weighting, and a weighted analysis will be necessary. Even the straightforward unweighted value of a proportion taken from such a sample would be a biased estimator of the true population value.

It should be noted that in selecting a PPS sample as described above it is possible for the same community to be selected twice, if that community has a population greater than the sampling interval. This is unlikely to happen if the proportion of communities

**Box 1. A cumulative list of community sizes**

<table>
<thead>
<tr>
<th>Community</th>
<th>Population size</th>
<th>Cumulative population size</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1 000</td>
<td>1 000</td>
</tr>
<tr>
<td>2</td>
<td>400</td>
<td>1 400</td>
</tr>
<tr>
<td>3</td>
<td>200</td>
<td>1 600</td>
</tr>
<tr>
<td>4</td>
<td>1 000</td>
<td>2 600</td>
</tr>
<tr>
<td>5</td>
<td>1 200</td>
<td>3 800</td>
</tr>
<tr>
<td>6</td>
<td>1 000</td>
<td>4 800</td>
</tr>
<tr>
<td>7</td>
<td>1 600</td>
<td>6 400</td>
</tr>
<tr>
<td>8</td>
<td>200</td>
<td>6 600</td>
</tr>
<tr>
<td>9</td>
<td>350</td>
<td>6 950</td>
</tr>
<tr>
<td>10</td>
<td>450</td>
<td>7 400</td>
</tr>
</tbody>
</table>
selected is small (the sampling fraction), unless one community is very much bigger than all the others. If it should happen, the correct procedure to follow would be to select two subsamples of households from within this community. It is equally valid (though less informative) to take only one subsample and count each observation twice over. It is not appropriate to select another community instead, or to repeat the whole sampling procedure until no communities are repeated, since either of these approaches invalidates the required probabilities.

If no measure of community population sizes is available at all, it will be impossible to carry out PPS sampling, and communities must be selected by simple random sampling. In this case a fixed number of households should still be taken from each selected community, but the responses obtained will have to be weighted in the analysis (see Section 5). This will necessitate a count of the total number of households in each selected community.

**Selection of households**

The ideal procedure for the selection of households would be to have a list of all households in the community and to choose a selection from the list at random. If such a list does not exist, and if the community is small, then a list can be created by carrying out a quick census, or perhaps by consulting community leaders.

If this is not practicable then some means has to be used which ensures that the sample is as representative as possible. This will usually involve two stages: a method of selecting one household to be the starting point and a procedure for selecting successive households after that.

The EPI recommendation for the first household is suitable (2); that is, to choose some central point in the community, such as the market; choose a random direction from that point, count the number of households between the central point and the edge of town in that direction, and select one of these houses at random to be the starting point of the survey.

The remaining households in the sample should be selected to give as widespread a coverage as possible of the community consistent with practicality. It is possible to follow the EPI strategy of simply going to the household whose door is nearest to the current household, but whereas this procedure is adequate for the purposes of EPI sampling (3) (where children of the right age are found only in a small proportion of households visited) it is unlikely to be adequate in general. It would be better to choose, say, the fifth nearest household, and better still to select all the households completely at random.

Some procedure needs to be adopted for dealing with dwellings which contain several households. If these are infrequent, it is best to select all the households within the selected dwelling, as this prevents households in multi-household dwellings from being underrepresented. If most dwellings contain more than one household, as for example in the compounds common in some parts of Africa, then the compound may be treated as a cluster and multistage sampling used (see Section 6). In large communities it would be a good idea to spread the sample around by having more than one starting point in different parts of the community. This would also reduce the underrepresentation of households in the outer parts of the community inherent in having just one central starting point.

The above ideas should be seen only as suggestions. Any method which achieves a random or near-random selection of households, preferably spread widely over the community, would be acceptable as long as it is clear and unambiguous, and does not give the field worker the opportunity to make personal choices which may introduce bias. In every situation a solution should be sought which is appropriate to local conditions.

**4. Sample size and precision**

**Precision, clustering and variability**

In deciding on an appropriate sample size for a survey, one is faced with the need to strike a balance between precision and cost. Ideally, one would decide on the precision needed and calculate the sample size accordingly. In practice, however, resources are always limited and often the best one can do is to calculate what sort of precision can be achieved with the resources available. This is valuable: in particular if the achievable precision is poor then perhaps the decision should be made not to carry out the survey at all.

The precision of the estimates made from the survey will depend on the size of the sample and the amount of clustering, and the item whose value is being measured. The size of the population from which the sample is selected has little effect in practice, and may be ignored. The larger the sample, other things being equal, the more precise any estimates will be. For the same overall total sample size, however, a survey in which a large number of clusters is selected, and a few households visited in each, will give more precise results than a survey in which a larger number of households is visited in each of a smaller number of clusters. For example, a survey in which 300 mothers are interviewed will usually give more precise results than one in which 200 mothers are interviewed, but if the 300 are distributed as 30 clusters of size 10. In opposition to this, a larger sample size and more clusters (even if somewhat smaller) will lead to an increased workload, which in turn means increases in costs and time.

The precision of an estimate also depends on the item itself and how even is its distribution across the population. For example, suppose the overall (unknown) proportion of households with a pit latrine in the region were 40%: if the proportions in each community in the region varied very little (say from 35% to 45%) then a small number of clusters selected would give a reasonably precise estimate; if, on the other hand, the proportions in each community varied more widely (say from 0% to 80%) then one would need a considerably larger sample to be sure of obtaining the same precision. This variability is measured by the rate of homogeneity (roh) which will be discussed in detail below (6).

The usual way to measure the precision of an estimate is by its standard error. We can then construct a 95% confidence interval for the true value.
from (estimate – 2 standard errors) to (estimate + 2 standard errors). If we denote the average number of responses achieved to an item per cluster by b and the total number of responses to the item in the survey by n, then the standard error of an estimated proportion p may be written in the form

\[ s = \sqrt{p(1-p)/n} \]  

(1)

Note that this is an extension of the simpler formula used when the data are assumed to come from a simple random sample, the binomial formula

\[ s = \sqrt{p(1-p)/n} \]  

(2)

The value of \( \sqrt{D} \) measures the increase in the standard error of the estimate due to the sampling procedure used.

D is known as the design effect and is given by

\[ D = 1 + (b-1)roh \]  

(3)

where \( roh \) is the rate of homogeneity mentioned above and b is the average number of responses to the question per cluster (see below). The value of \( roh \) will be estimated in the light of experience of previous surveys of similar design and subject matter. Such a value may be used for guidance on sample size decisions before the current survey is carried out, but once the analysis is underway, standard errors should be calculated using the formula for \( D \) (or equivalently of \( roh \)) which will be estimated from the data of that survey, from item to item. One possible contributing factor to the variation between clusters as compared to the variation within clusters. In a single-stage cluster design effect of 1.6, whereas a sample of 30 from each cluster would lead to a design effect of 3.9. Use of the formula (3), however approximate, is more likely to be appropriate than the value of 2 often used for the design effect regardless of cluster size or type of item.

The value of \( roh \) may be thought of as a measure of the variability between clusters as compared to the variation within clusters. In a single-stage cluster sample such as the one described here, \( roh \) is equivalent to the “intra-cluster correlation” \( \rho \); in a more complex design such as a stratified multistage survey, \( roh \) is composed of the components of variability from all stages of the design.

The value of \( roh \) will be higher for those items whose value varies more between clusters. For example, because families in the same area tend to have broadly similar socioeconomic status, variables such as “husband’s occupation: clerical” will be more likely to produce the same response for two individuals in the same cluster than for individuals in separate clusters. Such socioeconomic variables may have a relatively high value of \( roh \), around 0.20 (14).

Demographic items such as “currently married” and measures of mortality will be hardly more likely to produce the same answer from two respondents in the same cluster than from two respondents in different clusters, and will have \( roh \) very close to 0, around 0.02 (14). Questions of general morbidity such as “ill in past two weeks” may have similarly low values, but morbidity from specific infectious diseases may have much higher values, up to 0.3 (4). For questions of health-care practice and of use of health-care services such as “use of ORS for last episode of diarrhoea” or immunization coverage, responses will depend on the level of services locally and on local custom, and the value of \( roh \) may be from 0.1 to 0.2 (14). Although experience is limited, \( roh \) can take values up to 1, in practice values above 0.4 are uncommon, except for variables which are specific to the locality rather than the household, and hence clustered by definition, such as for example “health centre within 30 minutes walk”. That value declines slowly with the clustering of 

1 W.M. Liyanage, unpublished MSc thesis.
to 30, and \( n = 30 \times 20 = 600 \). If we have some idea of the proportion \( p \) in advance, we should use it in the formula, but if not it is best to use \( p = 0.5 \) as a guess since this maximizes \( s \) and hence error on the safe side. The value of roh is hardest to estimate, but is likely to be high, with more variation in such an item between communities than within each community, so we may take \( \text{roh} = 0.20 \). Using the formula (3) we obtain a design effect of

\[
D = 1 + (29 \times 0.20) = 6.8
\]

and from (1) the estimate of the standard error is

\[
s = \sqrt{[0.5 \times 0.5 \times 6.8 / 600]} = 0.05
\]

or 5%. This indicates that with such a sample size we can be 95% certain that the true proportion of households with latrines will lie within \( \pm 10\% \) (2 standard errors) of our estimate. Whether or not this precision is adequate depends on the purpose of our survey. If the design effect had been ignored, we would have predicted a standard error of

\[
s = \sqrt{[0.5 \times 0.5 / 600]} = 0.02,
\]

encouraging us to believe that our survey would give much more precise results than would actually be the case.

Suppose that in the same survey we also wished to estimate the proportion of children aged 12-23 months who had been adequately vaccinated by their first birthday. If we could assume that such children are found in about one-quarter of all households, then we would expect to get about 7 responses from each cluster, and we would take this as the value of \( b \). The values of \( n \) would be \( 7 \times 20 = 140 \). We might take the value of roh to be \( 0.10 \) and following the above calculations would obtain \( D = 1.6 \) and \( s = 5.3\% \), giving a 95% confidence interval of about 11%. Ignoring \( D \) would have led us to underestimate the width of the confidence interval as 8%.

**Estimating sample size**

If the investigator knows that a certain precision is required from the survey, then the necessary sample size may be calculated. Usually it will be a matter of deciding how many cluster samples of a given size \( b \) will be necessary. The design effect \( D \) should be calculated from (3) as before, and then the number of clusters necessary is given by \( c \) where

\[
c = \frac{p(1-p)D}{s^2b}. \tag{4}
\]

For example if \( p \) is expected to be around 20% for some measure of disease prevalence, for which we expect roh to be about 0.02, and suppose that we wish to estimate \( p \) to within \( \pm 5\% \). If we expect to have 20 responses from each cluster, then the value of \( D \) will be 1.36 (from (3)). For a confidence interval of \( \pm 5\% \) we shall need \( s = 0.025 \), then from (4) we need \( c = 18 \) clusters.

If we had failed to take account of the design effect we would have estimated the sample size from equation (4) as 13 clusters. Using equation (1), we see that our result would then have had a predicted standard deviation of 0.029 and a confidence interval of \( \pm 6\% \), a little less precise than desired. The small size of the loss of precision in this example is due only to the small value of \( D \). In many cases, \( D \) will be considerably larger, and the precision achieved considerably less than desired. In general, ignoring the design effect in estimating the sample size required will lead to confidence intervals which are wider than desired by a factor of \( \sqrt{D} \).

Such calculations should be made for the most important items in the survey schedule. Ideally \( c \) should be chosen to be the largest value given by these calculations in order to satisfy all the requirements. If the sample sizes necessary for different items are grossly different (as may happen in a study which covers both disease prevalence and usage of health-care facilities), it may be advisable to just use a subsample for those questions requiring fewer responses. However, the increase in complexity of the instructions given to interviewers means that this should be used with caution.

One should note that if the prevalence of an item under consideration is expected to be quite low, for example HIV seropositivity which may in some countries be around 2%, then it is not sensible to design a survey to achieve an absolute precision of 5%. In such a case the standard error desired needs to be considered relative to the expected prevalence rate, and would be much smaller, say 0.5% in absolute terms.

If the survey has been stratified (see Section 6) then each stratum should be considered as a separate survey, and sample-size calculations performed for each one to give the precision necessary for that stratum. The precision of the overall national estimate will then be somewhat better than that for any single stratum.

If the survey is one of a series, and the purpose is to estimate the change in some measure since the previous survey, then one needs to estimate the standard error of the change. This will be larger than the standard error of the new estimate of the measure, because of the imprecision of the estimate of the measure from the previous survey. To allow for this, the sample size may need to be doubled that calculated by the usual methods.

### 5. Analysis of data

This section describes the methods used to provide estimates of proportions or rates, together with standard errors of those estimates so that confidence intervals can be calculated. A mean value may also be estimated in the same way. We also describe how to calculate \( D \) and roh. The methods described below can be carried out on a simple calculator having a square-root key, and the use of a spreadsheet is illustrated in the Annex. The calculations in this and earlier sections may also be programmed easily on a computer using a spreadsheet package, as shown by Frerichs (11).

**Estimation of a proportion**

Suppose that a number of households have been selected in each of \( c \) communities with a view to estimating (by examining their record cards) what proportion of children aged 12-17 months were fully vaccinated on their first birthday. Suppose that in the \( j \)th community \( j = 1, \ldots, c \) these were \( x_i \); children whose record cards were examined, and that \( y_i \) of these were fully vaccinated as defined by the study.
Then the proportion of children in the $i^{th}$ community who were fully vaccinated will be given by

$$p_i = \frac{y_i}{x_i}.$$  

In the survey population as a whole the proportion who are fully vaccinated will be estimated by

$$p = \frac{\sum y}{\sum x};$$  

(5)

i.e. the total number of children vaccinated divided by the total number of children whose cards were examined. This is the straightforward ratio of the sample totals. Note that it is not the same as the average of the $p_i$'s, which would be incorrect since it does not take account of the variation in the $x_i$'s.

The standard error, $s$, of $p$ is obtained from the formula

$$s = \sqrt{\frac{\sum y_i^2 - 2\sum y_i x_i + p^2 \sum x_i^2}{[c(c-1)]}}.$$  

(6)

A spreadsheet for calculation of $s$ is given in the Annex, with an example of its use. This formula is more complex than the formula (2) usually used by standard computer packages in that it takes account of (i) the clustering of the sample and (ii) the variability between clusters of the denominator $x_i$. This value, (the number of record cards examined in the $i^{th}$ community) will have been unknown before the survey began and would probably be different if a different sample of households were taken from the same community. Failure to take account of these factors would lead to underestimation of $s$, and consequent overconfidence in the precision of the results (see Annex for an example). In many cases $x_i$ will not vary much between communities, for example when $x_i$ is the number of households selected, and then the simpler formula

$$s = \sqrt{\frac{\sum (p_i-p)^2}{[c(c-1)]}}.$$  

(7)

may be used instead of (6).

**Estimation of means**

At times one will collect data on values which are not simply "yes/no" attributes of the household or person, but counts or other measurable quantities, for example "number of children ever born" or "number of rooms". In this case one may wish to estimate the mean value over the population, for example the mean number of children ever born (although of course one may also estimate a proportion, for example the proportion of women who have given birth to more than 3 children). Estimation of the mean and its standard error are carried out in exactly the same way as for a proportion (Section 5) except that $y_i$ will now be equal to the sum of the numbers of children ever born to all of the $x_i$ mothers interviewed in the $i^{th}$ community.

**Weighted analysis**

In many situations there will be a need to weight the observations to allow for different probabilities of selection or different levels of non-response. For example suppose clusters were chosen with PPS as in Section 3, and it was intended to visit 25 households in each one, but because of staff illness it was only possible to visit 16 households in one of the clusters. If this fact is ignored, it will lead to that cluster being underrepresented in the calculation of the proportion $p$ and its standard error. The solution is to weight the responses from this community by multiplying them up by 25/16. In more general terms, this means replacing $x_i$ and $y_i$ each time they occur in formulae (5) and (6) with $w_i x_i$ and $w_i y_i$, giving the more general formula

$$p = \frac{\sum w_i y_i}{\sum w_i x_i}$$

and

$$s = \sqrt{\frac{\sum w_i^2 y_i^2 - 2\sum w_i x_i y_i + p^2 \sum w_i^2 x_i^2}{[c(c-1)]}}.$$  

where $w_i$ is the weight attached to the $i^{th}$ cluster. An unweighted cluster has $w_i = 1$.

If clusters are sampled with probability proportional to size and $x_i$ represents the number of BSUs (households) selected, then the proportion is estimated by $\hat{p}$, the average of the $p_i$'s, and we can use formula (7) for its standard error with $p$ replaced by $\hat{p}$. In other cases the approximate formula (7) ignores the size of the cluster and should not be used if weighting is necessary.

**Weighting may also be used to allow for clusters not being selected with probability proportional to size, for example when current size was not known at the time of their selection and they were selected with simple random sampling (or with probability proportional to a poor or very out-of-date measure of size). In this case the weight will be proportional to the actual population of the cluster (or the ratio of this to its old estimate).**

**Estimation of design effect**

The results of any survey may be used to estimate design effects, for use in the same or future surveys. The design effect is estimated by

$$D = \frac{s^2 \text{ from equation (6) or (7)}}{s^2 \text{ from equation (2)}}$$

The rate of homogeneity, $\rho$, may then be estimated as

$$(D - 1)/(b - 1)$$

where $b$ is as defined earlier. An example is given in the Annex.

**Imputation of standard errors**

In a large survey it may not be feasible to use the correct formulae (6) or (7) to estimate the standard error of every variable. In such a case one may calculate exact standard errors for a few variables of each type (socioeconomic, health status, etc.). Dividing each standard error by the corresponding binomial value (2) gives a new estimate of the design factor (the square root of the design effect $\sqrt{D}$). For the remaining variables of the survey one may calculate the binomial value (2) as given by calculator or standard software can be used, and just multiplied by the most appropriate value of $\sqrt{D}$ obtained for variables of similar type.

6. Extensions

The previous sections describe cluster-sampling procedures in a simple context: a sample of communities is selected from the whole region under consideration and a sample of households is visited in each selected community. Such a sampling scheme will be inadequate if the region is very large or if separate estimates are needed for different...
geographical areas. In this section we show how the techniques described above can be extended to allow for multistage sampling and stratification.

Multistage sampling

In a large region or country where an overall estimate is required, it will usually be sensible to select the sample of communities in at least two stages. For example, if the country is split into a number of administrative districts one would take a sample of districts by the systematic PPS method described in Section 3 (i.e. by making a list with cumulative population sizes). Within each selected district, communities would be selected, again by the systematic PPS method. The same number of communities must be selected in each district. If some districts are very small it may be sensible to combine them. Households would be selected in the usual way, with again the same number selected in each community.

With the systematic PPS method described here it is possible that the same district may be selected twice. This will happen if the population of the district is larger than the sampling interval. In this case two independent samples of communities should be selected from this district.

Decisions on the sample size will be made exactly as in Section 4, except that b will now be the expected number of responses per district and c will be the number of districts in the sample. The value of roh is now an indicator of the ratio of between-district variances to within-district variances. In theory, this requires an estimate of roh from a survey of similar multistage design. In practice, such estimates are not available, and the best one can do is probably to use the values given in Section 4 as guidelines, and bear in mind that they will be overestimates, as the value of roh is likely to decline slowly with the size of the primary cluster used.

The analysis will follow exactly the same pattern as in Section 5 except that \(x_i\) and \(y_i\), now refer to the number of responses and the number of positive responses respectively in the \(i^{th}\) district, summed over all communities selected in that district.

The method of sampling described here may be extended in exactly the same way to more stages if required.

Stratification

It may be necessary to obtain separate estimates for, say, the urban and rural sectors of the population, or for different provinces or ecological zones. Each province (etc.) will be a stratum, and a sample should be selected independently from each stratum. The sample size and structure for each stratum should be chosen with the conditions and needs of that stratum in mind, as if a separate survey were being carried out in that stratum alone. The samples may be of a different type and/or size for each stratum.

An estimate for each stratum may be calculated together with its standard error by treating each stratum as a separate survey. A stratified estimate for the whole country may then be calculated by weighting the sub-total estimates by the stratum populations. For example, suppose there are three strata and the estimates from them are \(p_1\), \(p_2\) and \(p_3\), with standard errors \(s_1\), \(s_2\) and \(s_3\) respectively. Then the estimate for the whole country would be

\[
p = V_1p_1 + V_2p_2 + V_3p_3
\]

with standard error

\[
s = \sqrt{\left(V_1^2s_1^2 + V_2^2s_2^2 + V_3^2s_3^2\right)}
\]

where \(V_i\) is the proportion of the country’s population which belongs to stratum 1, and so on (\(V_1 + V_2 + V_3 = 1\)). The standard error \(s\) for the national estimate will be somewhat less than the standard errors for the individual strata.

Implicit stratification

Stratification usually leads to a small reduction in the standard error of the overall estimate \(p\), compared to the error that would have been obtained if the survey had not been stratified. Another way of obtaining such a reduction is by implicit stratification. This is simply carried out at the time of selection of communities (or districts) by ensuring that the list of communities from which the systematic sample is to be taken is ordered by some measure which is correlated with the main purpose of the survey. For example, in a survey of the utilization of mother-and-child health facilities, there may have been a previous study carried out some years ago on the same subject, or there may be other knowledge available which indicates which communities may be expected to have high levels of utilization and which communities low levels. If not, one may be able to guess that those communities which are further from the regional capital, or which cover a more widely-scattered population, will have lower levels of utilization than others. Whatever the measure chosen, if the communities can be listed in approximate order from a high to a low level of expected utilization, then the sample selected will contain communities with a spread of utilization levels, and the estimated proportion \(p\) will be more precise. The standard error will be reduced, and its estimate \(s\) given by (6) will be somewhat of an overestimate (15). The improvement in precision cannot be quantified adequately to allow its use in sample-size calculations.

7. Conclusion

A simplified approach to survey design has been presented, with no attempt to cover all possible types of estimation. We have rather aimed to provide a set of guidelines which will enable the practitioner to plan a survey in a way which will give a reasonably representative sample, without any great bias, and of a suitable size to give adequate precision without wasting resources. The values given for the rate of homogeneity have of necessity been approximate, but variability between surveys and between variables is such that precise advice is impossible. The methods of analysis presented here offer an improvement on the common practice of assuming that the data came from a simple random sample and using the standard errors given by a calculator or standard computer package.

Acknowledgements

This article had its origins in work carried out for WHO/SHS on sampling in primary health care reviews. We are grateful to S. Lwanga, S.A. Sapirie, J. L. Tulloch and J. E. Dowd of the World Health Organization for their contributions, and to many colleagues for helpful comments.

Rapp. trimestr. statist. sanit. mond., 44 (1991)
General guidelines are presented for the use of cluster-sample surveys for health surveys in developing countries. The emphasis is on methods which can be used by practitioners with little statistical expertise and no background in sampling. A simple self-weighting design is used, based on that used by the World Health Organization’s Expanded Programme on Immunization (EPI). Topics covered include sample design, methods of random selection of areas and households, sample-size calculation and the estimation of proportions, ratios and means with standard errors appropriate to the design. Extensions are discussed, including stratification and multiple stages of selection. Particular attention is paid to allowing for the structure of the survey in estimating sample size, using the design effect and the rate of homogeneity. Guidance is given on possible values for these parameters. A spreadsheet is included for the calculation of standard errors.

**ANNEX**

**Estimating the standard error of a ratio and its design effect**

The use of a simple spreadsheet for the calculation of an estimate and its standard error using the precise formula (6) is demonstrated using the following example. The use of the approximate formula (7) for the standard error is also shown, and the design effect is calculated. The sample size is much smaller than those encountered in practice but all the important steps in the calculation are demonstrated.

Six communities are selected using the systematic PPS procedure. Twenty households are chosen in each community in order to estimate, for the population, the proportion of recently-pregnant mothers who have received postnatal care.

The data are:

<table>
<thead>
<tr>
<th>Community</th>
<th>Number of recently-pregnant women</th>
<th>Number receiving postnatal care</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>7</td>
<td>5</td>
</tr>
<tr>
<td>3</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td>6</td>
<td>3</td>
</tr>
<tr>
<td>5</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>6</td>
<td>3</td>
<td>0</td>
</tr>
</tbody>
</table>

The spreadsheet is constructed as follows:

<table>
<thead>
<tr>
<th>(y_i)</th>
<th>(x_i)</th>
<th>(y_i^2)</th>
<th>(x_i^2)</th>
<th>(x_iy_i)</th>
<th>(p_i)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>2</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>1.00</td>
</tr>
<tr>
<td>5</td>
<td>7</td>
<td>25</td>
<td>49</td>
<td>35</td>
<td>0.70</td>
</tr>
<tr>
<td>3</td>
<td>4</td>
<td>9</td>
<td>16</td>
<td>12</td>
<td>0.75</td>
</tr>
<tr>
<td>3</td>
<td>6</td>
<td>9</td>
<td>36</td>
<td>18</td>
<td>0.50</td>
</tr>
<tr>
<td>1</td>
<td>4</td>
<td>1</td>
<td>16</td>
<td>4</td>
<td>0.25</td>
</tr>
<tr>
<td>0</td>
<td>3</td>
<td>0</td>
<td>9</td>
<td>0</td>
<td>0.00</td>
</tr>
</tbody>
</table>

Total \(A=14\) \(B=26\) \(C=48\) \(E=130\) \(F=73\)

Here \(c=6\) is the number of communities; \(y_i\) is the number of recently-pregnant mothers in the \(i^{th}\) community who have received postnatal care; \(x_i\) is the number of recently-pregnant mothers in the sample from the \(i^{th}\) community.

The estimated proportion is

\[p = \frac{A}{B} = 0.5385.\]
The standard error $s$, as given by (6), is calculated as follows:

<table>
<thead>
<tr>
<th>New quantity</th>
<th>Calculated as</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p^2$</td>
<td>$p \times p$</td>
<td>0.2900</td>
</tr>
<tr>
<td>$G$</td>
<td>$2 \times p \times F$</td>
<td>78.621</td>
</tr>
<tr>
<td>$H$</td>
<td>$p^2 \times E$</td>
<td>37.70</td>
</tr>
<tr>
<td>$J$</td>
<td>$C - G + H$</td>
<td>7.079</td>
</tr>
<tr>
<td>$K$</td>
<td>$J/[c \times (c-1)]$</td>
<td>0.2360</td>
</tr>
<tr>
<td>$L$</td>
<td>$\sqrt{K}$</td>
<td>0.4858</td>
</tr>
<tr>
<td>$s$</td>
<td>$c \times L / B$</td>
<td>0.1121</td>
</tr>
</tbody>
</table>

The 95% confidence interval for the true proportion is $0.5385 \pm (2 \times 0.1121)$, i.e. 0.314 to 0.763.

The approximate formula (7) gives $s = 0.1482$. The difference between this figure and that given above arises because the $x_i$'s are very variable.

The standard error assuming a simple random sample is given by (2) as

$$s_{ssr} = \sqrt{(0.5385 \times (1-0.5385)/26)} = 0.0978,$$

thus ignoring the design of the study would have led us to assign our estimate a confidence interval from 0.343 to 0.734, which is 13% narrower than the correct value.

The design effect is estimated as

$$D = s^2/s_{ssr}^2 = (0.1121)^2/(0.0978)^2 = 1.31.$$

Since $b = \Sigma x_i/6 = 4.333$, $roh$ may be estimated in this case by $(D - 1)/(b - 1) = 0.093$.

**REFERENCES — RÉFÉRENCES**


Repp. trimest. statist. sanit. mond., 44 (1991)
Since 1978, the Expanded Programme on Immunization (EPI) of the World Health Organization (WHO) has advocated a modification of classical probability-proportional-to-size (PPS) cluster sampling for surveys of immunization coverage. In this methodology, random selection is not used for the selection of individuals at the second stage of sampling (1-3). Instead, only the first household visited at the second stage of sampling is chosen at random. Then as many nearby households as necessary are visited until the total desired number of eligible children has been selected from a particular cluster.

The decision to deviate from classical sampling theory is based primarily upon logistical and managerial constraints. Investigators concerned about the impact which this lack of random selection might have on estimates resulting from EPI-type surveys have evaluated the procedure using Monte Carlo computer simulation models (3). These studies suggested that the specified goal of the EPI methodology, to produce estimates of population rates accurate to within 10% of the true level, was satisfied in most of the artificial populations created.

This article reports the results of a Monte Carlo simulation study developed to evaluate the ability of a modified EPI sampling design to estimate not a population rate, but rather a population relative risk. In the context in which the EPI-like sampling scheme has been used, the actual measure of association estimated would be the prevalence ratio since these surveys have been cross-sectional in design. However, the computer simulation avoids consideration of specific study designs and uses the term relative risk to denote any of a variety of risk-ratio measures.

Methodology

The computer simulation program was designed to enable the user to create artificial populations which mimic typical African populations with respect to certain characteristics. Samples from these populations are then chosen by the program to serve as the basis for evaluating the proposed sampling scheme relative to simple random sampling (SRS) for estimating relative risks associated with hypothetical risk factors. A number of different populations were examined, some constructed to represent actual populations while others were fashioned to represent extreme population types.

While the basic design of the simulation program has been described elsewhere (3), several noteworthy modifications have been made for studying the estimation of relative risks.

All adults were assigned a value indicating whether or not they were exposed to each of two hypothetical risk factors. The exact assignment of these factors for infected and noninfected was set according to the specified level of association, based on a random process similar to that used in establishing other cluster characteristics. For evaluative purposes, one risk factor strongly related to HIV infection (RR=3.0) and a second unrelated to HIV infection (RR=1.0) were included.

The simulation program also included a means of evaluating the influence of several forms of potential bias commonly encountered in public health surveys. For instance, actual household surveys are almost certain to include individuals who are either unavailable or unwilling to participate. This can be particularly important since high levels of nonresponse by particular subgroups can potentially bias the results of a survey to a considerable degree. For each adult a variable indicating whether or not they participated in the survey was included in the simulation, in order to evaluate the influence of nonresponse upon estimates obtained using the two sampling strategies; a range of different levels of differential and nondifferential nonresponse were tested for this purpose.

Other forms of potential bias were evaluated in the simulation by associating additional parameters with each adult. To examine the effect on precision in the estimation of the relative risk that can potentially result from information bias, the ability to misclassify certain adults with respect to exposure status was incorporated within the program.

Having created a population of clusters, each with a particular set of characteristics, 500 independent samples of adults from each of the randomly-selected clusters were chosen using the EPI-like sampling strategy, and 500 additional independent samples were selected using SRS. Each individual sample provided an estimate of the relative risk associated with each hypothetical risk factor for a particular cluster. Collectively, information gathered on all clusters sampled in a single replication of the sampling process was used to derive an estimate of this parameter for the entire population.

The relative risk was calculated by the program as the ratio of the incidence, or as may be the case prevalence, in exposed individuals divided by the prevalence in those that are not exposed. Actually, two different approaches were included in the program to estimate the population relative risk. The first approach estimates the population relative risk as the simple average of the cluster-specific relative risks; the formulae for deriving the estimated population relative risk from the $j^{th}$ replication of the
Two means of estimating the variance for this particular method of estimating the relative risk were included in the program. The first represented the usual manner of calculating the variance from a PPS cluster sample as the sum of the squared differences of the cluster-specific estimates of the relative risk about their means, divided by \( m \) (\( m-1 \)). The second method employed a first-order (or linear) Taylor series approximation to estimate the variance of the log relative risk. The variance of the log relative risk \( \text{var} \{\ln(\text{RR})\} \) was calculated for each cluster sampled in a single repetition and the average of these cluster-specific variances was then determined. By constructing a confidence interval estimate (CIE) for the \( \ln(\text{RR}) \) and later exponentiating the endpoint of the interval, a confidence interval for the relative risk was obtained for each repetition.

The second approach for estimating the relative risk combined all data collected in a given repetition (i.e. all 50 clusters) and calculated the log of the relative risk from the resulting 2x2 table. The formula used to calculate the log relative risk for the \( j \)th replication of the sampling is displayed below:

\[
\ln(\text{RR}_j) = \ln \left( \sum_{i=1}^{m} \hat{a}_{ij} \right) - \ln \left( \sum_{i=1}^{m} \hat{b}_{ij} \right) + \ln \left( \sum_{i=1}^{m} \hat{c}_{ij} \right) - \ln \left( \sum_{i=1}^{m} \hat{d}_{ij} \right)
\]

where \( m \) is the number of clusters selected and \( \hat{a}, \hat{b}, \hat{c}, \) and \( \hat{d} \) are the frequencies observed in each of the cells of the 2x2 table classifying exposure by disease.

Exponentiating the log relative risk provided an estimate of the population parameter. The variance for this particular approach to estimating the relative risk was computed in two different fashions. The first method employed a first-order Taylor series approximation to estimate the variance of the log relative risk from the pooled 2x2 table. The second variance calculation employed a \( \delta \) method to approximate the variance of the log relative risk from the pooled 2x2 table.

The \( \delta \) method involves calculation of the product of the inverse \( d \) matrix times the \( V \) matrix times the \( d \) matrix. The \( d \) matrix is comprised of four elements which were derived by differentiating the log relative risk with respect to each cell of the 2x2 table resulting from cross-classifying disease by exposure. The \( V \) matrix is a four by four matrix having the variances of the individual elements of the cross-classification of adults by disease and exposure status (2x2 table) on the diagonal and the covariances off-diagonal.

For each approach to estimating the population relative risk, the program constructed tables displaying the distribution of estimates; these were used to assess whether the EPI-like sampling strategy met certain stated objectives with regard to accuracy in estimation of the population relative risk. Results presented in the tables in this article are based on the second method of estimating the RR along with the variance estimate based on the first-order Taylor series approximation.

Results

Precision in estimation using the EPI-like sampling design relative to SRS was evaluated with several different study populations. These populations, constructed both with and without pocketing of infection, included varying levels of nonresponse and misclassification of exposure. For each population created, performance was assessed based on selection of all household adults and on selection of only one adult per household for inclusion in the survey.

Tables 1-6 display the total number of sample estimates falling within the range defined by specific stated objectives. Each entry is based on a total of 500 replications of the sampling, each involving selection of a separate PPS sample of clusters. Consequently, these tables may be used to evaluate the likelihood that the results of a given survey will meet certain standards of performance.

Estimation of the relative risk

The first four columns of Table 1 display the total number of sample estimates accurate to within a range of \( \pm 0.10 \) of the actual population relative risk for risk factor 1. The observed differences between the two survey methods with respect to this objective in precision were typically small whether sampling from populations constructed with or without pocketing of infection.

The next series of entries in Table 1 present the total number of sample estimates falling within the less restrictive range defined by the true relative risk \( \pm 20\% \) of the actual value. As should be expected, the totals displayed in these columns tend to be larger than those described above. The results reported for the two survey methods were remarkably similar with respect to this objective, regardless of population type or presence of pocketing of infection.

The last four columns of Table 1 display the total number of confidence interval estimates (CIE) which included the true population parameter. Although SRS, in general, tended to perform somewhat better than the EPI design, the magnitude of the reported differences would likely be of little practical significance with regard to field application of this methodology.

Table 2 displays the total number of sample estimates meeting each of the three objectives in precision when the actual population relative risk was set at approximately 3.0. While the values reported in this table for the number of estimates within \( \pm 0.10 \) of the actual value are not as large as those observed when estimating the risk associated with the first factor, the observed differences between the two survey methods remained typically small.

The total numbers of sample estimates falling within the range defined by the actual relative risk \( \pm 20\% \) indicate that the EPI-like survey methodology was nearly always able to estimate a relative risk of this
### TABLE 1. NUMBER OF SAMPLE ESTIMATES OF THE POPULATION RELATIVE RISK OF EXPOSURE TO RISK FACTOR 1 (RR=1.0) THAT MET STATED OBJECTIVES IN PRECISION BASED ON SAMPLING OF ALL HOUSEHOLD ADULTS

<table>
<thead>
<tr>
<th>Population</th>
<th>No pocketing of infection</th>
<th>Pocketing of infection</th>
<th>Relative risk within 95% CIE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>EPI a SRS b</td>
<td>EPI a SRS b</td>
<td>EPI a SRS b</td>
</tr>
<tr>
<td>Standard</td>
<td>426 427</td>
<td>403 428</td>
<td>498 499 499 497</td>
</tr>
<tr>
<td>Rural – Rurale</td>
<td>423 443</td>
<td>447 444</td>
<td>494 499 499 499</td>
</tr>
<tr>
<td>Mixed – Mixte</td>
<td>458 462</td>
<td>445 443</td>
<td>499 498 499 499</td>
</tr>
<tr>
<td>City – Ville</td>
<td>448 470</td>
<td>447 447</td>
<td>499 499 498 500</td>
</tr>
<tr>
<td>Urban – Zones urbaines</td>
<td>437 441</td>
<td>424 435</td>
<td>472 475 469 483</td>
</tr>
</tbody>
</table>

* Expanded Programme on Immunization – Programme élargi de vaccination.
* Simple random sampling – Sondage aléatoire simple.

### TABLE 2. NUMBER OF SAMPLE ESTIMATES OF THE POPULATION RELATIVE RISK OF EXPOSURE TO RISK FACTOR 2 (RR=3.0) THAT MET STATED OBJECTIVES IN PRECISION BASED ON SAMPLING OF ALL HOUSEHOLD ADULTS

<table>
<thead>
<tr>
<th>Population</th>
<th>No pocketing of infection</th>
<th>Pocketing of infection</th>
<th>Relative risk within 95% CIE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>EPI a SRS b</td>
<td>EPI a SRS b</td>
<td>EPI a SRS b</td>
</tr>
<tr>
<td>Standard</td>
<td>164 188</td>
<td>218 211</td>
<td>498 499 499 499</td>
</tr>
<tr>
<td>Rural – Rurale</td>
<td>180 195</td>
<td>186 173</td>
<td>495 493 494 497</td>
</tr>
<tr>
<td>Mixed – Mixte</td>
<td>200 217</td>
<td>198 216</td>
<td>499 500 495 500</td>
</tr>
<tr>
<td>City – Ville</td>
<td>208 207</td>
<td>195 210</td>
<td>499 497 495 499</td>
</tr>
<tr>
<td>Urban – Zones urbaines</td>
<td>205 211</td>
<td>175 202</td>
<td>475 479 458 475</td>
</tr>
</tbody>
</table>

* Expanded Programme on Immunization – Programme élargi de vaccination.
* Simple random sampling – Sondage aléatoire simple.
magnitude to within this level of precision. With respect to the total number of confidence interval estimates which included the true population parameter, the observed data fail to demonstrate a significant difference between the two sampling methods regardless of population type or presence of pocketing of infection. In most of the situations evaluated, both methods of sampling were able to meet this study objective for more than 90% of the replications of the sampling.

**Estimation of the relative risk based on responders only**

Examination of the total number of sample estimates accurate to within a range of ±0.10 of the actual population relative risk for responders only indicated that although the number of successful estimates out of the 500 replications was not as large as that observed when estimating the relative risk based on all adults sampled, the observed differences between the two survey methods remained typically small; a probability of nonresponse of 20% was applied to infected adults while a nonresponse rate of only 5% was used to determine the response status of non-infected adults.

With respect to the total number of sample estimates falling within the range defined by the true relative risk ±20% of the actual value, for responders only, there was no substantial difference between the two methods of sampling with respect to either risk factor. Over 95% of all estimates produced with either method of sampling fell within the range defined by this objective. The presence of pocketing of infection did not impact significantly on precision in estimation with respect to either objective, nor did varying the study populations or selection of one vs. all household adults for inclusion in the survey.

When the variance based on information from responders only was used to derive a confidence interval for the estimated population relative risk, there was no indication that either method of sampling covered the true population RR with greater frequency. Both survey techniques appeared equally capable of meeting this objective with a high degree of confidence.

The effect of varying certain population characteristics on estimation of relative risks

Since variations between the study populations seemed to have little influence on precision in estimation using either sampling technique, evaluation of the effect of varying certain population characteristics was based on sampling from a standard population only.

The total number of estimates of the relative risk within certain ranges of precision of the actual value when sampling from the standard population constructed with varying levels of differential and non-differential nonresponse indicated that in most of the situations evaluated, SRS tended to perform somewhat better than the EPI-like method, though the differences between them were generally quite small.

Varying the levels of misclassification of exposure status generally indicated that SRS tended to produce greater numbers of estimates meeting each of the objectives of precision in estimation. However, the differences between the two sampling designs were typically small and not likely to be of any practical significance.

The total number of estimates of the relative risk of exposure to risk factor 1 that met certain stated objectives of precision when the population seroprevalence rate was varied is displayed in Table 3.

The total number of estimates of the population relative risk falling within the first two ranges of interest tended to increase with increases in the level of the seroprevalence rate. However, construction of confidence intervals about estimates of the relative risk was largely unaffected by increases in the rate of infection; using either method of sampling enabled confidence interval estimates to be generated which included the true population value in over 92% of samples. More importantly, throughout this table a high degree of comparability is observed between the results obtained with the two sampling designs. When differences were detected, they tended to be too small to be of any practical significance.

**TABLE 3. NUMBER OF ESTIMATES OF THE POPULATION RELATIVE RISK OF EXPOSURE TO RISK FACTOR 1 THAT MET STATED OBJECTIVES IN PRECISION WHEN THE TRUE SEROPREVALENCE RATE WAS VARIED**

<table>
<thead>
<tr>
<th>Seroprevalence rate (%)</th>
<th>RR ± 0.1</th>
<th>RR ± 20% (RR)</th>
<th>RR within CIE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>EPI a - PEV a</td>
<td>SRS b - SAS b</td>
<td>EPI a - PEV a</td>
</tr>
<tr>
<td>2</td>
<td>301</td>
<td>281</td>
<td>429</td>
</tr>
<tr>
<td>4</td>
<td>329</td>
<td>353</td>
<td>458</td>
</tr>
<tr>
<td>6</td>
<td>402</td>
<td>400</td>
<td>493</td>
</tr>
<tr>
<td>8</td>
<td>430</td>
<td>439</td>
<td>499</td>
</tr>
<tr>
<td>10</td>
<td>434</td>
<td>458</td>
<td>499</td>
</tr>
<tr>
<td>15</td>
<td>466</td>
<td>484</td>
<td>500</td>
</tr>
<tr>
<td>20</td>
<td>494</td>
<td>494</td>
<td>500</td>
</tr>
<tr>
<td>25</td>
<td>498</td>
<td>499</td>
<td>500</td>
</tr>
<tr>
<td>30</td>
<td>497</td>
<td>500</td>
<td>500</td>
</tr>
<tr>
<td>40</td>
<td>499</td>
<td>500</td>
<td>500</td>
</tr>
<tr>
<td>50</td>
<td>500</td>
<td>500</td>
<td>500</td>
</tr>
</tbody>
</table>

a Expansed Programme on Immunization – Programme élargi de vaccination.
b Simple random sampling – Sondage aléatoire simple.

Rapp. trimestr. statist. sanit. mond., 44 (1991)
The total number of estimates of the relative risk associated with exposure to risk factor 1 that met the stated objectives of precision in estimation when the population seroprevalence rate was varied is reported in Table 4, for responders only; a probability of nonresponse of 20% was applied to infected adults and a rate of only 5% used to determine the response status of noninfected adults. The number of EPI-like estimates falling within the range defined by the true rate ±0.10 tended to be quite similar to those obtained using SRS. With either method of sampling, the number of estimates meeting this objective rose as the level of infection increased.

The total number of estimates within the range defined by the true relative risk ±20% times the actual value, indicated that either sampling strategy was quite capable of meeting this objective with a high degree of confidence at virtually every seroprevalence level evaluated. Even in the lower range of infection, which is more likely for HIV infection, the EPI-like method performed nearly equally to or better than SRS.

Varying the level of infection appeared to have little impact on construction of confidence intervals. Confidence interval estimates generated using either method of sampling included the true population value in over 90% of samples. Overall, the data observed in this table fail to demonstrate a clear advantage of one sampling technique over the other in meeting the stated objectives.

The total number of estimates of the relative risk of exposure to risk factor 2 that met stated objectives of precision in estimation when the population seroprevalence rate was varied is presented in Table 5. It should be recalled that the magnitude of association between exposure and disease status for the second risk factor was set at approximately 3.0.

### Table 4.
**Number of Estimates of the Population Relative Risk of Exposure to Risk Factor 1, Based Solely on Responding Adults, That Met Stated Objectives in Precision When True Seroprevalence Rate Was Varied**

<table>
<thead>
<tr>
<th>Seroprevalence rate (Taux de seroprévalence (%))</th>
<th>Objective in precision – Precision recherchée</th>
<th>RR within CIE – RR dans un intervalle de confiance donné</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>RR ± 0.1ification – RR ± 0.1</td>
<td>RR ± 20% (RR) – RR ± 20% (RR)</td>
</tr>
<tr>
<td></td>
<td>EPI • PEV a – SAS b</td>
<td>EPI • PEV a – SAS b</td>
</tr>
<tr>
<td>2</td>
<td>240 ± 0.10</td>
<td>383 ± 0.10</td>
</tr>
<tr>
<td>4</td>
<td>313 ± 0.10</td>
<td>456 ± 0.10</td>
</tr>
<tr>
<td>6</td>
<td>394 ± 0.10</td>
<td>488 ± 0.10</td>
</tr>
<tr>
<td>8</td>
<td>417 ± 0.10</td>
<td>498 ± 0.10</td>
</tr>
<tr>
<td>10</td>
<td>416 ± 0.10</td>
<td>492 ± 0.10</td>
</tr>
<tr>
<td>15</td>
<td>450 ± 0.10</td>
<td>499 ± 0.10</td>
</tr>
<tr>
<td>20</td>
<td>486 ± 0.10</td>
<td>500 ± 0.10</td>
</tr>
<tr>
<td>25</td>
<td>487 ± 0.10</td>
<td>500 ± 0.10</td>
</tr>
<tr>
<td>30</td>
<td>495 ± 0.10</td>
<td>500 ± 0.10</td>
</tr>
<tr>
<td>40</td>
<td>499 ± 0.10</td>
<td>500 ± 0.10</td>
</tr>
<tr>
<td>50</td>
<td>500 ± 0.10</td>
<td>500 ± 0.10</td>
</tr>
</tbody>
</table>

a Expanded Programme on Immunization – Programme élargi de vaccination.
b Simple random sampling – Sondage aléatoire simple.

### Table 5.
**Number of Estimates of the Population Relative Risk of Exposure to Risk Factor 2 That Met Stated Objectives in Precision When True Seroprevalence Rate Was Varied**

<table>
<thead>
<tr>
<th>Seroprevalence rate (Taux de seroprévalence (%))</th>
<th>Objective in precision – Precision recherchée</th>
<th>RR within CIE – RR dans un intervalle de confiance donné</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>RR ± 0.1ification – RR ± 0.1</td>
<td>RR ± 20% (RR) – RR ± 20% (RR)</td>
</tr>
<tr>
<td>2</td>
<td>113 ± 0.10</td>
<td>462 ± 0.10</td>
</tr>
<tr>
<td>4</td>
<td>132 ± 0.10</td>
<td>471 ± 0.10</td>
</tr>
<tr>
<td>6</td>
<td>174 ± 0.10</td>
<td>497 ± 0.10</td>
</tr>
<tr>
<td>8</td>
<td>183 ± 0.10</td>
<td>493 ± 0.10</td>
</tr>
<tr>
<td>10</td>
<td>196 ± 0.10</td>
<td>492 ± 0.10</td>
</tr>
<tr>
<td>15</td>
<td>199 ± 0.10</td>
<td>499 ± 0.10</td>
</tr>
<tr>
<td>20</td>
<td>179 ± 0.10</td>
<td>499 ± 0.10</td>
</tr>
<tr>
<td>25</td>
<td>173 ± 0.10</td>
<td>497 ± 0.10</td>
</tr>
<tr>
<td>30</td>
<td>171 ± 0.10</td>
<td>496 ± 0.10</td>
</tr>
<tr>
<td>40</td>
<td>129 ± 0.10</td>
<td>486 ± 0.10</td>
</tr>
<tr>
<td>50</td>
<td>107 ± 0.10</td>
<td>461 ± 0.10</td>
</tr>
</tbody>
</table>

a Expanded Programme on Immunization – Programme élargi de vaccination.
b Simple random sampling – Sondage aléatoire simple.

Wld hlw statist. quart., 44 (1991)
The number of estimates of the relative risk that fell within the range defined by the true rate ±0.10 was generally rather small using either method of sampling. However, if the range of interest was instead defined on the basis of the true value ±20% times the actual value, the total number of estimates meeting this objective in precision was generally shown to increase along with increases in the seroprevalence rate. Indeed, either sampling technique appeared quite capable of meeting this objective in precision with a high degree of confidence at the levels evaluated.

The ability of either method of sampling to construct confidence intervals which included the true relative risk tended to diminish as the prevalence of infection rose; this was particularly true of estimates employing the EPI survey strategy. When the infection rate exceeded 15%, the EPI-like method generated confidence intervals which included the actual parameter value for less than 70% of the samples.

The total number of estimates of the relative risk of exposure to the second risk factor that met the stated objectives of precision in estimation when the population seroprevalence rate was varied is reported in Table 6, for responders only.

The total number of estimates within the range of the actual value ±0.10 indicated that neither method of sampling was capable of meeting this objective. The number of estimates within the range defined by the true relative risk ±20% times the actual value demonstrated marked improvement over the previous study objective; the data suggest that either sampling strategy was capable of meeting this objective with a high degree of confidence.

Varying the level of infection appeared to substantially influence the construction of confidence interval estimates. Generally, either method of sampling enabled confidence intervals to be generated that included the true population value with a high degree of confidence only when the infection rate did not exceed 20%. However, throughout the range of infection examined, the observed data generally failed to demonstrate a clear advantage of one sampling technique over the other with regard to confidence interval estimation.

Discussion

The use of the EPI-like survey technique for estimating relative risks was evaluated with respect to three different objectives in precision in estimation. The levels of precision employed for the first and second objectives included estimation of the relative risk (RR) within the range defined by the true level ±20% times the actual RR and estimation of the RR to within the narrower range defined by the RR ±0.10. A 95% confidence interval was used in assessing the third objective.

In the absence of pocketing of infection, results obtained from five simulated populations failed to provide sufficient evidence to suggest that the EPI-like survey strategy, when applied at the second stage of sampling, was less able to meet the stated objectives in precision than SRS with respect to estimation of the population relative risk.

When pocketing of infection was included within the population, although differences were noted between the two methods with respect to the stated objectives in precision in estimation of relative risks, their magnitude was insufficient to recommend one method over the other. Moreover, pocketing of infection did not appear to differentially affect the precision of the two survey methods.

Though not included in the tables, when sampling from the various study populations, the number of adults selected per household for inclusion in the survey (1 vs. all) did not appreciably affect the precision in estimation of relative risks using either method of sampling.

The observed estimates of relative risks based solely on responding adults indicated that while the total number of estimates meeting the stated objectives tended to decrease, the precision of the two sam-

<table>
<thead>
<tr>
<th>Seroprevalence rate (%)</th>
<th>RR ± 0.1</th>
<th>RR ± 20% (RR)</th>
<th>RR within C/E</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>EPI - PEV</td>
<td>SRS - SAS</td>
<td>EPI - PEV</td>
</tr>
<tr>
<td>2</td>
<td>103</td>
<td>90</td>
<td>449</td>
</tr>
<tr>
<td>4</td>
<td>98</td>
<td>121</td>
<td>451</td>
</tr>
<tr>
<td>6</td>
<td>165</td>
<td>152</td>
<td>454</td>
</tr>
<tr>
<td>8</td>
<td>179</td>
<td>181</td>
<td>451</td>
</tr>
<tr>
<td>10</td>
<td>182</td>
<td>188</td>
<td>486</td>
</tr>
<tr>
<td>15</td>
<td>174</td>
<td>169</td>
<td>493</td>
</tr>
<tr>
<td>20</td>
<td>191</td>
<td>211</td>
<td>495</td>
</tr>
<tr>
<td>25</td>
<td>123</td>
<td>126</td>
<td>480</td>
</tr>
<tr>
<td>30</td>
<td>144</td>
<td>76</td>
<td>485</td>
</tr>
<tr>
<td>40</td>
<td>103</td>
<td>9</td>
<td>433</td>
</tr>
<tr>
<td>50</td>
<td>93</td>
<td>0</td>
<td>399</td>
</tr>
</tbody>
</table>

* Expanded Programme on Immunization – Programme élargi de vaccination.

b Simple random sampling – Sondage aléatoire simple.

Rapp. trimest. statist. sanit. mond., 44 (1991)
Sampling methods generally did not appear to differ markedly. Differences in precision in estimation of relative risks using the two sampling methods were of little practical significance when nonresponse was differentially or non-differentially distributed, even in the presence of pocketing of infection.

In order to investigate the impact of other forms of potential bias on precision in estimation when using the EPl-like sampling design, sampling was also conducted using populations constructed with varying levels of misclassification of exposure status. For exploratory purposes, a range of values was chosen for designating misclassification of exposure. Differences in precision between the two methods when estimating the relative risk where there was misclassification of exposure were found to be of little practical significance.

To determine the effect on precision in estimation that results from varying the seroprevalence rate, the simulation was carried out using a standard population constructed with different levels of infection; infection rates ranging between 2% and 50% were investigated. Varying the level of infection did not appear to differentially affect the two means of survey sampling with regard to estimation of the population relative risk when it was set equal to 1.0. In this case, the difference in the number of estimates meeting each of the stated objectives using the two survey methods did not differ widely enough to be of any practical significance.

In attempting to estimate a larger magnitude of risk (RR=3.0), the SRS strategy was generally better able to meet the objectives in precision, especially with higher seroprevalence rates.

Obviously, the real world is much more complicated than can be represented in a computer simulation. While the resulting simplifications should be viewed as an arbitrary representation of reality, they can potentially influence the interpretation of the results. For instance, there was no attempt to construct clusters with varying numbers of households and the chosen cluster size of 600 households is, no doubt, larger than many areas where such surveys may be conducted.

Pocketing of infected adults was also represented in an exaggerated fashion. Pockets of infection were produced in the simulation models reported by infecting 75% of all adults residing in households within the pockets and none elsewhere within the cluster; although the program allowed other choices. Such extremes are likely to be rare in real populations.

The computer algorithm used to simulate movement of interviewers from house to house represents yet another example of the simplification of reality. This algorithm was based on an orderly arrangement of households within a matrix which clearly does not reflect conditions typically encountered in the field.

In actual field operations, non-sampling errors (such as interviewer, recall, and reporting bias) can be a significant problem. The simulation attempted to evaluate the impact of nonresponse and misclassification bias only. It was assumed that the potential for a selection bias was minimized through repeated sampling from the populations created. Aside from these potential biases, no attempt was made to address other potential sources of non-sampling errors that may arise under field conditions.

While attempts were made to reflect a number of population characteristics in the simulation, necessity it was not possible to include all features that characterize real populations. Moreover, since each execution of the simulation required considerable time to complete, only certain combinations of these characteristics were explored. Nevertheless, for the characteristics that were included, a range of possible values was examined.

Having stated the various limitations of the simulation, it is important to note that it is unlikely that any of these would differentially affect results obtained from the two methods of sampling.

In summary, if the principal outcome of interest in a particular study is to estimate the relative risk associated with exposure to some factor with a high degree of precision, the EPl-like survey method appears to be a reasonable alternative to the use of SRS at the second stage of sampling. In virtually all of the situations evaluated, estimates of the relative risk obtained using the two sampling methods failed to demonstrate a clear advantage in the use of one survey design over the other. However, the added cost and difficulty encountered in implementing SRS at the second stage of sampling certainly warrant further consideration of the EPI methodology for use in estimating relative risks.

Evidence from this study also suggests that sampling involving selection of only one adult per household may improve the precision of estimates derived using the EPl-like method. However, since selection of only one adult per household would require that a greater number of households be visited, any gain in precision must be weighed against the likely increase in cost of household visitation.

**SUMMARY**

Precision in estimation of relative risks using a standardized sampling method proposed by the WHO Global Programme on AIDS was evaluated using a Monte Carlo model simulating actual populations; the proposed survey design represents a modification of the methodology used by the WHO Expanded Programme on Immunization (EPI) to estimate immunization coverage among children. This study suggests that in actual populations the proposed survey strategy is a reasonable alternative to the use of simple random sampling (SRS) at the second stage of cluster sampling. Although varying such population characteristics as the seroprevalence rate, nonresponse rate, and rate of misclassification of exposure failed to demonstrate a clear advantage of one method over the other, the added cost and difficulty of implementing SRS under field conditions warrant further consideration of the EPI-like methodology for use in estimating relative risks.
On a évalué, au moyen d’un modèle de Monte Carlo simulant des populations réelles, la précision de l’estimation du risque relatif obtenue par la méthode de sondage normalisée que propose d’utiliser le Programme mondial OMS de lutte contre le SIDA. Le plan d’enquête proposé est une variante de la méthodologie utilisée par le Programme élargi de vaccination (PEV) de l’OMS pour estimer la couverture vaccinale des enfants. Il semble, au vu des résultats, que dans le cas de populations réelles, la méthodologie proposée puisse, au second degré d’un sondage en grappes, se substituer valablement au sondage aléatoire simple (SAS). Même si, en faisant varier des caractéristiques de la population telles que le taux de séroprévalence, le taux de non-réponse et le taux d’erreur sur le type d’exposition, on ne peut pas trancher en faveur de l’une des deux méthodes, le coût supérieur et la difficulté d’exécution du SAS sur le terrain justifient que l’on étudie plus avant la méthodologie utilisée par le PEV pour l’estimation du risque relatif.

REFERENCES — RÉFÉRENCES


LOT QUALITY ASSURANCE SAMPLING: SINGLE- AND DOUBLE-SAMPLING PLANS

Stanley Lemeshow & Scott Taber

One goal of health professionals concerned with curbing the spread of HIV infection is targeting communities with high seroprevalence rates. Choosing a sampling procedure is a critical step in achieving this goal. Lot quality assurance sampling (LQAS) has been suggested for use in targeting such communities because of the economy that may be derived from obtaining the required information with smaller sample sizes than would typically be needed when using more traditional sampling procedures (1,2).

The economy of LQAS may be improved while sustaining statistical precision by using double sampling. This approach selects the sample in two stages. At the first stage, a relatively small sample is selected. If the results from this first sample are "extreme", sampling stops and conclusions are drawn from the smaller sample. On the other hand, if the results of the preliminary sample are equivocal, then a second sample is chosen and conclusions are based upon the results of the combined sample selected.

The theoretical foundations, statistical qualities, relative economy and application of LQAS single sampling, and double sampling, will be discussed in this article.

Lot quality assurance sampling: an example

LQAS was originally used for manufacturing inspection where it was necessary to keep sampling costs to a minimum. LQAS is identical to stratified sampling, but the samples are too small to provide what are usually considered acceptably narrow confidence intervals for stratum, or "lot", specific parameter estimates. Rather, a decision is made about the quality of a particular lot based on the probability that the number of defective items in the sample selected from that lot is less than or equal to some critical value. This is sometimes referred to as acceptance sampling.

The strategy and goals of LQAS in the health field are similar to those in the manufacturing field. The manufacturer does not want to release a lot with more than a certain percentage defective. Similarly, a health planner does not want to conclude that a community has an acceptably low HIV seroprevalence when it actually exceeds some target seroprevalence above which the community should be targeted for intervention.

Suppose a researcher is interested in whether a community has a 10% prevalence of HIV infection. Traditional sampling procedures, such as simple random sampling, stratified sampling and cluster sampling would be useful if we wanted to estimate the prevalence of HIV infection. However, we are most interested in hypothesis testing relative to a threshold prevalence level beyond which health planners will intervene. LQAS addresses this issue. In order to provide a more familiar framework for health workers, we will develop the HIV example further.

Firstly, lots must be defined for our study. Usually, a lot is an operationally useful unit, and the lots form a mutually exclusive and exhaustive set. A public health planner might define a lot as a community that is socially and geographically distinct. Additionally, each lot may represent a distinct health care delivery area.

Secondly, the planner must decide whether to use one- or two-stage sampling. Under certain circumstances, two-stage sampling may require fewer subjects to be involved in the study, thereby reducing sampling costs. This choice will be developed in a later section.

Thirdly, sample sizes and the critical number of HIV seropositive subjects must be specified for each lot. The fundamental problem in LQAS sampling is not so much simply determining sample size as choosing an appropriate combination of sample size and critical value. The criteria for choosing a combination depends on the health planner's concern about controlling both the chance of concluding that a community has low HIV seroprevalence when it is actually high, as well as the chance of concluding that a community has high HIV seroprevalence when in fact it is low. These decisions determine the manner in which available health-care resources will be allocated.

The choice of sample size n and critical value d* depends on the following parameters specified by the health planner:

- **n** – lot size (i.e. population size of the community).
- **R₀** – upper threshold proportion of infected people beyond which intervention is deemed necessary.
- **Rₚ** – lower threshold proportion of infected people below which health planners have determined it to be more economical to continue surveillance while focusing intervention resources on more needy communities.
- **α** – probability of concluding that a community has low HIV seroprevalence when, in fact, it has a high level.
- **β** – probability of concluding a community has high HIV seroprevalence when, in fact, it has a low level.

The theoretical foundations, statistical qualities, relative economy and application of LQAS single sampling, and double sampling, will be discussed in this article.
has a low level. $1 - \beta$ is the desired statistical power to target those communities which are most needy.

A later section will demonstrate the statistical theory involved in calculating values for $n$ and $d^*$ under specified levels of the above quantities.

Fourthly, the observed number of HIV seropositive subjects in the sample is compared to the critical value for each lot. Subsequently, the health planner can make a decision as to whether each community requires intervention or continued surveillance.

In public health work a serious error would be made if a community were judged to have low HIV seroprevalence when, in fact, it did not. Since in statistical hypothesis testing the $\alpha$-error is specified by the experimenter, the probability of failing to detect a highly-infected community can be made as small as desired. The procedure is set up as a one-sided test as follows:

the null hypothesis is $H_0: P \geq P_0$ (i.e. proportion of infected subjects $\geq 0.05$) vs. the alternative hypothesis $H_a: P < P_0$ (i.e. proportion of infected subjects $< 0.05$ (e.g. $P_0 = 0.01$)).

The four-celled table presented in Fig. 1 describes the consequences of the testing procedure.

---

**FIG. 1**

**CONSEQUENCES D'UN TEST D'HYPOTHÈSE DANS LA MÉTHODE LQAS**

<table>
<thead>
<tr>
<th>Population réelle</th>
<th>Forte séroprévalence VIH</th>
<th>Faible séroprévalence VIH</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ne rejette pas Ho</td>
<td>Le test reconnaît/test sensible à un fort taux d'infection</td>
<td>“Risque pour le fournisseur”</td>
</tr>
<tr>
<td>“fort taux d'infection”</td>
<td>$1 - \alpha$</td>
<td>$\beta$</td>
</tr>
<tr>
<td>Sensibilité</td>
<td>Taux de faux positifs</td>
<td></td>
</tr>
<tr>
<td>Rejette Ho</td>
<td>“Risque pour la communauté”</td>
<td>Le test reconnaît un faible taux d'infection</td>
</tr>
<tr>
<td>“faible taux d'infection”</td>
<td>$\alpha$</td>
<td>$1 - \beta$</td>
</tr>
<tr>
<td>Taux de faux négatifs</td>
<td>Spécificité</td>
<td></td>
</tr>
</tbody>
</table>

---

**FIG. 1**

**CONSEQUENCES OF HYPOTHESIS TESTING IN LQAS PROCEDURE**

<table>
<thead>
<tr>
<th>Actual population</th>
<th>High HIV seroprevalence</th>
<th>Low HIV seroprevalence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fail to reject Ho</td>
<td>Test recognizes or is sensitive to the high level of infection</td>
<td>“Provider risk”</td>
</tr>
<tr>
<td>“high level of infection”</td>
<td>$1 - \alpha$</td>
<td>$\beta$</td>
</tr>
<tr>
<td>Sensitivity</td>
<td>False positive rate</td>
<td></td>
</tr>
<tr>
<td>Reject Ho</td>
<td>“Community risk”</td>
<td>Test recognizes low level of infection</td>
</tr>
<tr>
<td>“low level of infection”</td>
<td>$\alpha$</td>
<td>$1 - \beta$</td>
</tr>
<tr>
<td>False negative rate</td>
<td>Specificity</td>
<td></td>
</tr>
</tbody>
</table>

---

$d$ The level 5% is chosen here as an example. Actually, any level could have been selected.
Note that this figure reflects the fact that our hypothesis test is one-sided, and we assume the community has a high HIV seroprevalence \( P \geq P_0 \) unless our sample results yield significant evidence that we are sampling from a community with an appreciably lower seroprevalence level, in which case we reject \( H_0 \). We consider the hypothesis test this way because the health planner is able to set \( \alpha \), thus controlling the probability of concluding a community has an acceptably low seroprevalence level when it actually has a high level. This error, the type I error, is considered most serious because a health planner would decide not to target this community for intervention and the disease would continue spreading until the next survey. Hence, we consider the cost of declaring that a community has low HIV seroprevalence, when in fact it is not, to be high.

On the other hand, the type II error, rejection of an acceptable lot, is judged not to be as serious since the result of concluding a community has high HIV seroprevalence, when in fact it is low, would be to concentrate programme resources on an already low seroprevalence area. Though this would be an inefficient allocation of resources, the old adage “better safe than sorry” tells us that this is not as serious a scenario as the type I error.

The mechanics of sample size and critical region determination will be demonstrated in a latter section. We will be able to demonstrate the effect of various distributional assumptions, and the choice of a one- or two-stage sampling plan, on the statistical precision of a study.

In order to better understand the issues involved in sample size and critical value determination, it will be useful to introduce the operating characteristics curve.

The operating characteristics (OC) curve

Most LOAS literature refers to the operating characteristics (OC) curve as a tool for sample size and critical region determination because it summarizes the statistical relationship amongst the parameters mentioned previously. Although our discussion of LOAS will focus on creating and using tables for sample size and critical value determination, the OC curve will be an important tool for illustrating the relationship between the quantities introduced previously. More specifically, each OC curve will incorporate the following information for a unique combination of \( n \) and \( d^* \) (referred to as a candidate sampling plan):

- The horizontal axis of an OC curve corresponds to the proportion \( P \) in the population who are seropositive. These values may range from 0 to 1.
- The vertical axis of an OC curve corresponds to the probability of rejecting the null hypothesis, \( H_0: P \geq P_0 \), and concluding that the seroprevalence is significantly below \( P_0 \). This axis will be used to assess the \( \alpha \) level, probability of making a type I error, of a candidate sampling plan. These values will also range between 0 and 1.
- The slope of an OC curve indicates how much the probability of rejecting a null hypothesis, or detecting that a community has an alternate level of HIV seroprevalence, changes as we consider different values of \( P_0 \). If \( Pr(d \geq d^*) \) increases substantially over a relevant range of possible seroprevalence levels, \( P < P_0 \), we would conclude that we have a high level of statistical power to detect these seroprevalence levels relative to \( P_0 \).

The choice of a sampling plan consists of selecting that OC curve which contains a point corresponding to both the hypothesized seroprevalence and \( \alpha \).
levels of interest in a study, while also having a steep enough slope to indicate sufficient statistical power.

Some typical OC curves are presented in Fig. 2. Notice that the curves are labelled for the specific combination of \( n \) and \( d' \). We could also graph other candidate sample plans on the same axes and compare them based on the above criteria.

Several important dynamics are captured in an OC curve. As \( P_0 \) increases, the probability of observing less than or equal to some specified number of defective people \( d' \) decreases. As mentioned previously, the slope of the OC curve in the region to the left of \( P_0 \) indicates the statistical power of the sampling plan. As sample size increases, or critical value decreases, the OC curve will shift downward indicating that, for a given \( \alpha \) level, we will be able to test for a lower value of \( P_0 \). These interrelationships should be considered carefully in order to gain a better intuitive sense of the relationship amongst these parameters.

We are now ready to consider the theoretical underpinnings which capture these relationships and allow precise calculations of sample sizes, critical values and probabilities.

**LOAS single sampling – the theory**

Our goal is to calculate values of \( n \) and \( d' \) that limit the chance of making a type I error to some specified level of \( \alpha \). We need to distinguish between two possible scenarios of LOAS sampling. Under the first scenario, we will also be interested in controlling the chance of making a type II error. In order to control this probability, \( \beta \) and \( P_0 \) need to be specified. These values will depend on how concerned the health planner is about intervention resources being spent on communities which do not need them as badly as others.

In the second scenario, the cost of making a type II error, concluding a community has a high HIV sero-prevalence when it is actually low, is not serious. For instance, the cost of intervention may be low, or there may be enough resources available, to administer proper interventions to any community with a seroprevalence level in the neighbourhood of the cut-off value \( P_0 \). Therefore, the health planner may choose a sampling plan in which \( \beta \) and \( P_0 \) are not specified. However, under this second scenario, the health planner must designate a value for \( N \), the population size of the community being sampled. Both situations will be discussed.

An important aspect of these calculations, in either case, is the probability distribution used in making choices that are sensitive to the probability parameters stated by the researcher, such as the \( \alpha \) and \( \beta \) levels. Relevant distributions will be discussed, followed by a numerical example.

The hypergeometric distribution will be important for representing the probability of observing \( d \) infected people in a sample of size \( n \) from a population of size \( N \) in which \( NP_0 \) members are hypothesized to be infected. An important feature of the hypergeometric distribution is that it accounts for the fact that the probability of selecting an infected member of the population changes as members are sampled without replacement (2). The hypergeometric distribution is represented by the following formula:

\[
P(d \leq d^*) = \sum_{d=0}^{d^*} \binom{NP_0}{d} \frac{(N(1-P_0))}{n-d} \binom{N}{n}
\]

If this probability is small relative to \( \alpha \), then we may conclude that it is unlikely the proportion of infected people in the population is as high as \( P_0 \). In other words, the number of infected people in our sample is significantly less than that expected if the true proportion of infected people in the population was actually \( P_0 \). As a result, the community would be accepted as a low-prevalence area.

The hypergeometric distribution may be used for sample size determination in the sense that we will choose the value of \( n \) that will yield a hypergeometric probability less than or equal to \( \alpha \) given the values of \( P_0 \), \( d' \) and \( N \) stated by the researcher.

Approximations are often used in an attempt to simplify calculations and represent the information in a more intuitive manner. When the ratio of sample size to population size is smaller than approximately 10% (3, 4), or when \( N \) is large (5), these approximations are expected to yield very similar probability calculations to the exact hypergeometric probability. We will discuss the two most important approximations used for the hypergeometric distribution, namely the normal and binomial approximations.

The normal approximation to the hypergeometric distribution is valuable because many researchers have been exposed to the basic characteristics of the normal distribution and, therefore, greater intuition about the issue at hand may be gained. Also, use of the normal distribution allows us more easily to incorporate the statistical power to distinguish between a community with \( P = P_0 \) and one with \( P = P_1 \), where \( P_0 < P_1 \). The mean and variance of the hypergeometric distribution with sample size \( n \) and hypothesized proportion \( P_0 \) are as follows (6):

\[
\text{mean (d)} = nP_0
\]

\[
\text{var (d)} = nP_0(1-P_0)\frac{N-n}{N-1}
\]

[Note that if \( n \) were small relative to \( N \), the finite population correction term could be dropped from the variance expression because it becomes effectively 1. In this situation, the hypergeometric distribution is approximated well by the binomial distribution.]

We may standardize the observed number of infected people by using the standard normal transformation:

\[
z = \frac{d - nP_0}{\sqrt{nP_0(1-P_0)\frac{N-n}{N-1}}}
\]

This expression represents the number of standard deviations our observation \( d \) is from the expected value \( nP_0 \). The value of \( z \) can be used to estimate the probability of the observed scenario relative to our hypothesized scenario by using a standard normal probability table.

Rapp. trimest. statist. sanit. mond., 44 (1991)
The normal approximation can be used to calculate sample size, given values for $P_0$, $d^*$, $\alpha$, $\beta$, $P_a$ and $N$, by solving a system of equations. Alternatively, we can calculate a critical value, given values for $P_0$, $n$, $z_{1-\alpha}$ and $N$, while recognizing that we are testing a one-sided hypothesis, as follows:

$$d^* = nP_0 - z_{1-\alpha} \sqrt{nP_0(1-P_0)(N-n)/N}.$$ 

Since $d^*$ is unlikely to be an integer, the value chosen will be the largest integer $\leq d^*$.

The binomial approximation can also be used to represent the probability of observing $d$ people with a characteristic out of a sample of size $n$. The binomial distribution is the statistical distribution that describes the probability of a particular configuration of dichotomous outcomes when the total number of trials is finite (e.g. the number of times a "head" appears in seven tosses of a coin). An important assumption when using the binomial approximation is that the probability of an event occurring is constant for each consecutive person sampled. This assumption can, for all practical purposes, be considered true when $N$ is large because the probability will not change appreciably as people are sampled from a community without replacement. The binomial probability can be expressed as follows:

$$P(d \leq d^*) = \sum_{d=0}^{d^*} \binom{n}{d} P_0^d (1-P_0)^{n-d}$$

Tables 1A-4 use the normal approximation to the hypergeometric distribution to carry out the necessary calculations. These tables were created using standard spreadsheet software. Interested readers may contact the authors for templates.

### TABLE 1A. LQAS SINGLE-STAGE SAMPLE SIZES FOR ALPHA = 0.05 AND $d^* = 0$

**TABLEAU 1A. TAILLES D’ÉCHANTILLONS POUR LQAS À 1 DEGRÉ POUR ALPHA = 0,05 ET $d^* = 0$**

<table>
<thead>
<tr>
<th>$d^*$</th>
<th>$Z(1/\alpha)$</th>
<th>0.01</th>
<th>0.02</th>
<th>0.03</th>
<th>0.04</th>
<th>0.05</th>
<th>0.075</th>
<th>0.1</th>
<th>0.125</th>
<th>0.15</th>
<th>0.175</th>
<th>0.2</th>
<th>0.225</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1.645</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>100</td>
<td></td>
<td>74</td>
<td>58</td>
<td>47</td>
<td>40</td>
<td>35</td>
<td>26</td>
<td>20</td>
<td>17</td>
<td>14</td>
<td>12</td>
<td>10</td>
<td>9</td>
</tr>
<tr>
<td>200</td>
<td></td>
<td>115</td>
<td>80</td>
<td>62</td>
<td>50</td>
<td>42</td>
<td>33</td>
<td>29</td>
<td>22</td>
<td>18</td>
<td>15</td>
<td>13</td>
<td>11</td>
</tr>
<tr>
<td>1,000</td>
<td></td>
<td>212</td>
<td>118</td>
<td>81</td>
<td>62</td>
<td>49</td>
<td>33</td>
<td>24</td>
<td>19</td>
<td>16</td>
<td>13</td>
<td>11</td>
<td>10</td>
</tr>
<tr>
<td>2,000</td>
<td></td>
<td>237</td>
<td>125</td>
<td>84</td>
<td>63</td>
<td>51</td>
<td>33</td>
<td>25</td>
<td>19</td>
<td>16</td>
<td>13</td>
<td>11</td>
<td>10</td>
</tr>
<tr>
<td>2,500</td>
<td></td>
<td>243</td>
<td>126</td>
<td>85</td>
<td>64</td>
<td>51</td>
<td>33</td>
<td>25</td>
<td>19</td>
<td>16</td>
<td>13</td>
<td>11</td>
<td>10</td>
</tr>
<tr>
<td>5,000</td>
<td></td>
<td>255</td>
<td>130</td>
<td>87</td>
<td>65</td>
<td>51</td>
<td>34</td>
<td>25</td>
<td>19</td>
<td>16</td>
<td>13</td>
<td>11</td>
<td>10</td>
</tr>
<tr>
<td>10,000</td>
<td></td>
<td>261</td>
<td>131</td>
<td>87</td>
<td>65</td>
<td>52</td>
<td>34</td>
<td>25</td>
<td>19</td>
<td>16</td>
<td>13</td>
<td>11</td>
<td>10</td>
</tr>
<tr>
<td>15,000</td>
<td></td>
<td>264</td>
<td>132</td>
<td>87</td>
<td>65</td>
<td>52</td>
<td>34</td>
<td>25</td>
<td>19</td>
<td>16</td>
<td>13</td>
<td>11</td>
<td>10</td>
</tr>
<tr>
<td>20,000</td>
<td></td>
<td>265</td>
<td>132</td>
<td>88</td>
<td>65</td>
<td>52</td>
<td>34</td>
<td>25</td>
<td>19</td>
<td>16</td>
<td>13</td>
<td>11</td>
<td>10</td>
</tr>
<tr>
<td>25,000</td>
<td></td>
<td>266</td>
<td>132</td>
<td>88</td>
<td>65</td>
<td>52</td>
<td>34</td>
<td>25</td>
<td>19</td>
<td>16</td>
<td>13</td>
<td>11</td>
<td>10</td>
</tr>
<tr>
<td>50,000</td>
<td></td>
<td>267</td>
<td>133</td>
<td>88</td>
<td>65</td>
<td>52</td>
<td>34</td>
<td>25</td>
<td>19</td>
<td>16</td>
<td>13</td>
<td>11</td>
<td>10</td>
</tr>
<tr>
<td>1,000,000</td>
<td></td>
<td>268</td>
<td>133</td>
<td>88</td>
<td>65</td>
<td>52</td>
<td>34</td>
<td>25</td>
<td>19</td>
<td>16</td>
<td>13</td>
<td>11</td>
<td>10</td>
</tr>
</tbody>
</table>

### TABLE 1B. LQAS SINGLE-STAGE SAMPLE SIZES FOR ALPHA = 0.05 AND $d^* = 1$

**TABLEAU 1B. TAILLES D’ÉCHANTILLONS POUR LQAS À 1 DEGRÉ POUR ALPHA = 0,05 ET $d^* = 1$**

<table>
<thead>
<tr>
<th>$d^*$</th>
<th>$Z(1/\alpha)$</th>
<th>0.01</th>
<th>0.02</th>
<th>0.03</th>
<th>0.04</th>
<th>0.05</th>
<th>0.075</th>
<th>0.1</th>
<th>0.125</th>
<th>0.15</th>
<th>0.175</th>
<th>0.2</th>
<th>0.225</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.645</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>100</td>
<td></td>
<td>101</td>
<td>88</td>
<td>75</td>
<td>64</td>
<td>56</td>
<td>42</td>
<td>34</td>
<td>28</td>
<td>24</td>
<td>21</td>
<td>18</td>
<td>16</td>
</tr>
<tr>
<td>200</td>
<td></td>
<td>176</td>
<td>129</td>
<td>100</td>
<td>82</td>
<td>69</td>
<td>49</td>
<td>38</td>
<td>31</td>
<td>26</td>
<td>22</td>
<td>19</td>
<td>17</td>
</tr>
<tr>
<td>1,000</td>
<td></td>
<td>347</td>
<td>195</td>
<td>135</td>
<td>103</td>
<td>83</td>
<td>56</td>
<td>41</td>
<td>33</td>
<td>27</td>
<td>23</td>
<td>20</td>
<td>17</td>
</tr>
<tr>
<td>2,000</td>
<td></td>
<td>391</td>
<td>207</td>
<td>141</td>
<td>106</td>
<td>85</td>
<td>56</td>
<td>42</td>
<td>33</td>
<td>27</td>
<td>23</td>
<td>20</td>
<td>17</td>
</tr>
<tr>
<td>2,500</td>
<td></td>
<td>401</td>
<td>210</td>
<td>142</td>
<td>107</td>
<td>85</td>
<td>57</td>
<td>42</td>
<td>33</td>
<td>27</td>
<td>23</td>
<td>20</td>
<td>17</td>
</tr>
<tr>
<td>5,000</td>
<td></td>
<td>422</td>
<td>216</td>
<td>144</td>
<td>108</td>
<td>86</td>
<td>57</td>
<td>42</td>
<td>33</td>
<td>27</td>
<td>23</td>
<td>20</td>
<td>18</td>
</tr>
<tr>
<td>10,000</td>
<td></td>
<td>434</td>
<td>219</td>
<td>146</td>
<td>109</td>
<td>87</td>
<td>57</td>
<td>42</td>
<td>33</td>
<td>27</td>
<td>23</td>
<td>20</td>
<td>18</td>
</tr>
<tr>
<td>15,000</td>
<td></td>
<td>438</td>
<td>220</td>
<td>146</td>
<td>109</td>
<td>87</td>
<td>57</td>
<td>42</td>
<td>33</td>
<td>27</td>
<td>23</td>
<td>20</td>
<td>18</td>
</tr>
<tr>
<td>20,000</td>
<td></td>
<td>440</td>
<td>220</td>
<td>146</td>
<td>109</td>
<td>87</td>
<td>57</td>
<td>42</td>
<td>33</td>
<td>27</td>
<td>23</td>
<td>20</td>
<td>18</td>
</tr>
<tr>
<td>25,000</td>
<td></td>
<td>441</td>
<td>221</td>
<td>147</td>
<td>109</td>
<td>87</td>
<td>57</td>
<td>42</td>
<td>33</td>
<td>28</td>
<td>23</td>
<td>20</td>
<td>18</td>
</tr>
<tr>
<td>50,000</td>
<td></td>
<td>443</td>
<td>221</td>
<td>147</td>
<td>110</td>
<td>87</td>
<td>57</td>
<td>42</td>
<td>33</td>
<td>28</td>
<td>23</td>
<td>20</td>
<td>18</td>
</tr>
<tr>
<td>1,000,000</td>
<td></td>
<td>446</td>
<td>222</td>
<td>147</td>
<td>110</td>
<td>87</td>
<td>57</td>
<td>42</td>
<td>34</td>
<td>28</td>
<td>23</td>
<td>20</td>
<td>18</td>
</tr>
</tbody>
</table>

Wild hût statist. quart., 44 (1991)
### TABLE 1C. LOAS SINGLE-STAGE SAMPLE SIZES FOR ALPHA = 0.05 AND d' = 2

<table>
<thead>
<tr>
<th>N</th>
<th>0.01</th>
<th>0.02</th>
<th>0.03</th>
<th>0.04</th>
<th>0.05</th>
<th>0.075</th>
<th>0.1</th>
<th>0.125</th>
<th>0.15</th>
<th>0.175</th>
<th>0.2</th>
<th>0.225</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>#NUM!</td>
<td>100</td>
<td>93</td>
<td>82</td>
<td>73</td>
<td>56</td>
<td>45</td>
<td>38</td>
<td>32</td>
<td>28</td>
<td>25</td>
<td>22</td>
</tr>
<tr>
<td>200</td>
<td>201</td>
<td>164</td>
<td>130</td>
<td>107</td>
<td>91</td>
<td>66</td>
<td>51</td>
<td>41</td>
<td>35</td>
<td>30</td>
<td>26</td>
<td>23</td>
</tr>
<tr>
<td>1,000</td>
<td>459</td>
<td>260</td>
<td>181</td>
<td>138</td>
<td>112</td>
<td>75</td>
<td>56</td>
<td>45</td>
<td>37</td>
<td>32</td>
<td>27</td>
<td>24</td>
</tr>
<tr>
<td>2,000</td>
<td>522</td>
<td>279</td>
<td>190</td>
<td>143</td>
<td>115</td>
<td>77</td>
<td>57</td>
<td>45</td>
<td>37</td>
<td>32</td>
<td>27</td>
<td>24</td>
</tr>
<tr>
<td>2,500</td>
<td>536</td>
<td>283</td>
<td>191</td>
<td>144</td>
<td>116</td>
<td>77</td>
<td>57</td>
<td>45</td>
<td>37</td>
<td>32</td>
<td>28</td>
<td>24</td>
</tr>
<tr>
<td>5,000</td>
<td>568</td>
<td>291</td>
<td>195</td>
<td>146</td>
<td>117</td>
<td>77</td>
<td>58</td>
<td>46</td>
<td>38</td>
<td>32</td>
<td>28</td>
<td>24</td>
</tr>
<tr>
<td>10,000</td>
<td>584</td>
<td>295</td>
<td>197</td>
<td>147</td>
<td>118</td>
<td>78</td>
<td>58</td>
<td>46</td>
<td>38</td>
<td>32</td>
<td>28</td>
<td>24</td>
</tr>
<tr>
<td>15,000</td>
<td>590</td>
<td>297</td>
<td>198</td>
<td>148</td>
<td>118</td>
<td>78</td>
<td>58</td>
<td>46</td>
<td>38</td>
<td>32</td>
<td>28</td>
<td>24</td>
</tr>
<tr>
<td>20,000</td>
<td>593</td>
<td>297</td>
<td>198</td>
<td>148</td>
<td>118</td>
<td>78</td>
<td>58</td>
<td>46</td>
<td>38</td>
<td>32</td>
<td>28</td>
<td>24</td>
</tr>
<tr>
<td>25,000</td>
<td>595</td>
<td>298</td>
<td>198</td>
<td>148</td>
<td>118</td>
<td>78</td>
<td>58</td>
<td>46</td>
<td>38</td>
<td>32</td>
<td>28</td>
<td>24</td>
</tr>
<tr>
<td>50,000</td>
<td>598</td>
<td>299</td>
<td>199</td>
<td>148</td>
<td>118</td>
<td>78</td>
<td>58</td>
<td>46</td>
<td>38</td>
<td>32</td>
<td>28</td>
<td>24</td>
</tr>
<tr>
<td>1,000,000</td>
<td>602</td>
<td>300</td>
<td>199</td>
<td>149</td>
<td>118</td>
<td>78</td>
<td>58</td>
<td>46</td>
<td>38</td>
<td>32</td>
<td>28</td>
<td>24</td>
</tr>
</tbody>
</table>

### TABLE 1D. LOAS SINGLE-STAGE SAMPLE SIZES FOR ALPHA = 0.05 AND d' = 3

<table>
<thead>
<tr>
<th>N</th>
<th>0.01</th>
<th>0.02</th>
<th>0.03</th>
<th>0.04</th>
<th>0.05</th>
<th>0.075</th>
<th>0.1</th>
<th>0.125</th>
<th>0.15</th>
<th>0.175</th>
<th>0.2</th>
<th>0.225</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>#NUM!</td>
<td>#NUM!</td>
<td>101</td>
<td>95</td>
<td>86</td>
<td>68</td>
<td>55</td>
<td>46</td>
<td>40</td>
<td>35</td>
<td>31</td>
<td>27</td>
</tr>
<tr>
<td>200</td>
<td>#NUM!</td>
<td>#NUM!</td>
<td>189</td>
<td>156</td>
<td>130</td>
<td>111</td>
<td>81</td>
<td>63</td>
<td>51</td>
<td>43</td>
<td>37</td>
<td>33</td>
</tr>
<tr>
<td>1,000</td>
<td>538</td>
<td>320</td>
<td>224</td>
<td>171</td>
<td>139</td>
<td>94</td>
<td>70</td>
<td>56</td>
<td>46</td>
<td>40</td>
<td>34</td>
<td>30</td>
</tr>
<tr>
<td>2,000</td>
<td>622</td>
<td>345</td>
<td>235</td>
<td>178</td>
<td>143</td>
<td>95</td>
<td>71</td>
<td>57</td>
<td>47</td>
<td>40</td>
<td>35</td>
<td>30</td>
</tr>
<tr>
<td>2,500</td>
<td>661</td>
<td>350</td>
<td>238</td>
<td>179</td>
<td>144</td>
<td>96</td>
<td>71</td>
<td>57</td>
<td>47</td>
<td>40</td>
<td>35</td>
<td>30</td>
</tr>
<tr>
<td>5,000</td>
<td>703</td>
<td>361</td>
<td>242</td>
<td>182</td>
<td>146</td>
<td>97</td>
<td>72</td>
<td>57</td>
<td>47</td>
<td>40</td>
<td>35</td>
<td>30</td>
</tr>
<tr>
<td>10,000</td>
<td>725</td>
<td>367</td>
<td>245</td>
<td>184</td>
<td>147</td>
<td>97</td>
<td>72</td>
<td>57</td>
<td>47</td>
<td>40</td>
<td>35</td>
<td>30</td>
</tr>
<tr>
<td>15,000</td>
<td>732</td>
<td>369</td>
<td>246</td>
<td>184</td>
<td>147</td>
<td>97</td>
<td>72</td>
<td>57</td>
<td>47</td>
<td>40</td>
<td>35</td>
<td>30</td>
</tr>
<tr>
<td>20,000</td>
<td>737</td>
<td>370</td>
<td>246</td>
<td>184</td>
<td>147</td>
<td>97</td>
<td>72</td>
<td>57</td>
<td>47</td>
<td>40</td>
<td>35</td>
<td>30</td>
</tr>
<tr>
<td>25,000</td>
<td>739</td>
<td>370</td>
<td>246</td>
<td>184</td>
<td>147</td>
<td>97</td>
<td>72</td>
<td>57</td>
<td>47</td>
<td>40</td>
<td>35</td>
<td>30</td>
</tr>
<tr>
<td>50,000</td>
<td>743</td>
<td>371</td>
<td>247</td>
<td>185</td>
<td>147</td>
<td>97</td>
<td>72</td>
<td>57</td>
<td>47</td>
<td>40</td>
<td>35</td>
<td>30</td>
</tr>
<tr>
<td>1,000,000</td>
<td>748</td>
<td>373</td>
<td>247</td>
<td>185</td>
<td>147</td>
<td>97</td>
<td>72</td>
<td>57</td>
<td>47</td>
<td>40</td>
<td>35</td>
<td>30</td>
</tr>
</tbody>
</table>

# NUM! : not possible to calculate sample size — taille de l'échantillon non calculable.
<table>
<thead>
<tr>
<th>N</th>
<th>Hypothesized seroprevalence (P₀) — Séroprévalence supposée (P₀)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.01</td>
</tr>
<tr>
<td>100</td>
<td>#NUM!</td>
</tr>
<tr>
<td>200</td>
<td>#NUM!</td>
</tr>
<tr>
<td>1,000</td>
<td>648</td>
</tr>
<tr>
<td>2,000</td>
<td>755</td>
</tr>
<tr>
<td>2,500</td>
<td>780</td>
</tr>
<tr>
<td>5,000</td>
<td>831</td>
</tr>
<tr>
<td>10,000</td>
<td>859</td>
</tr>
<tr>
<td>15,000</td>
<td>869</td>
</tr>
<tr>
<td>20,000</td>
<td>873</td>
</tr>
<tr>
<td>25,000</td>
<td>876</td>
</tr>
<tr>
<td>50,000</td>
<td>882</td>
</tr>
<tr>
<td>1,000,000</td>
<td>888</td>
</tr>
</tbody>
</table>

# NUM!: not possible to calculate sample size — taille de l'échantillon non calculable.
### TABLE 2A. LOAS SINGLE-STAGE SAMPLE SIZES FOR ALPHA = 0.01 AND $d^* = 0$

<table>
<thead>
<tr>
<th>N</th>
<th>0.01</th>
<th>0.02</th>
<th>0.03</th>
<th>0.04</th>
<th>0.05</th>
<th>0.075</th>
<th>0.1</th>
<th>0.125</th>
<th>0.15</th>
<th>0.175</th>
<th>0.2</th>
<th>0.225</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>85</td>
<td>73</td>
<td>64</td>
<td>57</td>
<td>51</td>
<td>41</td>
<td>33</td>
<td>28</td>
<td>24</td>
<td>21</td>
<td>18</td>
<td>16</td>
</tr>
<tr>
<td>200</td>
<td>146</td>
<td>115</td>
<td>94</td>
<td>79</td>
<td>69</td>
<td>51</td>
<td>40</td>
<td>32</td>
<td>27</td>
<td>23</td>
<td>20</td>
<td>18</td>
</tr>
<tr>
<td>1,000</td>
<td>350</td>
<td>210</td>
<td>150</td>
<td>116</td>
<td>94</td>
<td>63</td>
<td>47</td>
<td>37</td>
<td>30</td>
<td>25</td>
<td>22</td>
<td>19</td>
</tr>
<tr>
<td>2,000</td>
<td>423</td>
<td>235</td>
<td>161</td>
<td>122</td>
<td>98</td>
<td>65</td>
<td>48</td>
<td>38</td>
<td>31</td>
<td>26</td>
<td>22</td>
<td>19</td>
</tr>
<tr>
<td>2,500</td>
<td>442</td>
<td>240</td>
<td>164</td>
<td>124</td>
<td>99</td>
<td>66</td>
<td>48</td>
<td>38</td>
<td>31</td>
<td>26</td>
<td>22</td>
<td>19</td>
</tr>
<tr>
<td>5,000</td>
<td>484</td>
<td>252</td>
<td>170</td>
<td>127</td>
<td>101</td>
<td>66</td>
<td>49</td>
<td>38</td>
<td>31</td>
<td>26</td>
<td>22</td>
<td>19</td>
</tr>
<tr>
<td>10,000</td>
<td>509</td>
<td>259</td>
<td>172</td>
<td>129</td>
<td>102</td>
<td>67</td>
<td>49</td>
<td>38</td>
<td>31</td>
<td>26</td>
<td>22</td>
<td>19</td>
</tr>
<tr>
<td>15,000</td>
<td>518</td>
<td>261</td>
<td>173</td>
<td>129</td>
<td>103</td>
<td>67</td>
<td>49</td>
<td>38</td>
<td>31</td>
<td>26</td>
<td>22</td>
<td>19</td>
</tr>
<tr>
<td>20,000</td>
<td>522</td>
<td>262</td>
<td>174</td>
<td>130</td>
<td>103</td>
<td>67</td>
<td>49</td>
<td>38</td>
<td>31</td>
<td>26</td>
<td>22</td>
<td>19</td>
</tr>
<tr>
<td>25,000</td>
<td>525</td>
<td>263</td>
<td>174</td>
<td>130</td>
<td>103</td>
<td>67</td>
<td>49</td>
<td>38</td>
<td>31</td>
<td>26</td>
<td>22</td>
<td>19</td>
</tr>
<tr>
<td>50,000</td>
<td>530</td>
<td>264</td>
<td>175</td>
<td>130</td>
<td>103</td>
<td>67</td>
<td>49</td>
<td>38</td>
<td>31</td>
<td>26</td>
<td>22</td>
<td>19</td>
</tr>
<tr>
<td>1,000,000</td>
<td>536</td>
<td>266</td>
<td>175</td>
<td>130</td>
<td>103</td>
<td>67</td>
<td>49</td>
<td>38</td>
<td>31</td>
<td>26</td>
<td>22</td>
<td>19</td>
</tr>
</tbody>
</table>

### TABLE 2B. LOAS SINGLE-STAGE SAMPLE SIZES FOR ALPHA = 0.01 AND $d^* = 1$

<table>
<thead>
<tr>
<th>N</th>
<th>0.01</th>
<th>0.02</th>
<th>0.03</th>
<th>0.04</th>
<th>0.05</th>
<th>0.075</th>
<th>0.1</th>
<th>0.125</th>
<th>0.15</th>
<th>0.175</th>
<th>0.2</th>
<th>0.225</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>101</td>
<td>93</td>
<td>84</td>
<td>75</td>
<td>68</td>
<td>55</td>
<td>45</td>
<td>39</td>
<td>33</td>
<td>29</td>
<td>26</td>
<td>23</td>
</tr>
<tr>
<td>200</td>
<td>186</td>
<td>150</td>
<td>125</td>
<td>106</td>
<td>92</td>
<td>69</td>
<td>54</td>
<td>45</td>
<td>38</td>
<td>32</td>
<td>28</td>
<td>25</td>
</tr>
<tr>
<td>1,000</td>
<td>466</td>
<td>282</td>
<td>202</td>
<td>156</td>
<td>127</td>
<td>86</td>
<td>65</td>
<td>51</td>
<td>42</td>
<td>36</td>
<td>31</td>
<td></td>
</tr>
<tr>
<td>2,000</td>
<td>567</td>
<td>316</td>
<td>218</td>
<td>166</td>
<td>133</td>
<td>89</td>
<td>66</td>
<td>52</td>
<td>43</td>
<td>36</td>
<td>31</td>
<td></td>
</tr>
<tr>
<td>2,500</td>
<td>593</td>
<td>324</td>
<td>222</td>
<td>168</td>
<td>135</td>
<td>90</td>
<td>66</td>
<td>52</td>
<td>43</td>
<td>36</td>
<td>31</td>
<td></td>
</tr>
<tr>
<td>5,000</td>
<td>651</td>
<td>340</td>
<td>229</td>
<td>172</td>
<td>138</td>
<td>91</td>
<td>67</td>
<td>53</td>
<td>43</td>
<td>36</td>
<td>31</td>
<td></td>
</tr>
<tr>
<td>10,000</td>
<td>685</td>
<td>349</td>
<td>233</td>
<td>175</td>
<td>139</td>
<td>91</td>
<td>67</td>
<td>53</td>
<td>43</td>
<td>36</td>
<td>31</td>
<td></td>
</tr>
<tr>
<td>15,000</td>
<td>697</td>
<td>352</td>
<td>235</td>
<td>175</td>
<td>139</td>
<td>92</td>
<td>67</td>
<td>53</td>
<td>43</td>
<td>36</td>
<td>31</td>
<td></td>
</tr>
<tr>
<td>20,000</td>
<td>703</td>
<td>354</td>
<td>235</td>
<td>176</td>
<td>140</td>
<td>92</td>
<td>69</td>
<td>53</td>
<td>43</td>
<td>36</td>
<td>31</td>
<td></td>
</tr>
<tr>
<td>25,000</td>
<td>707</td>
<td>355</td>
<td>236</td>
<td>176</td>
<td>140</td>
<td>92</td>
<td>68</td>
<td>53</td>
<td>43</td>
<td>36</td>
<td>31</td>
<td></td>
</tr>
<tr>
<td>50,000</td>
<td>714</td>
<td>357</td>
<td>237</td>
<td>176</td>
<td>140</td>
<td>92</td>
<td>68</td>
<td>53</td>
<td>43</td>
<td>36</td>
<td>31</td>
<td></td>
</tr>
<tr>
<td>1,000,000</td>
<td>722</td>
<td>359</td>
<td>237</td>
<td>177</td>
<td>140</td>
<td>92</td>
<td>68</td>
<td>53</td>
<td>43</td>
<td>37</td>
<td>31</td>
<td>27</td>
</tr>
</tbody>
</table>

*Repr. trimest. statist. sanit. mond., 44 (1991)*
### Table 2C.

**LOAS Single-Stage Sample Sizes for $\alpha = 0.01$ and $d^* = 2$**

<table>
<thead>
<tr>
<th>$\text{N}$</th>
<th>Hypothesized seroprevalence ($R_0$) — Séroprévalence supposée ($R_0$)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.01</td>
</tr>
<tr>
<td>100</td>
<td>96</td>
</tr>
<tr>
<td>200</td>
<td>176</td>
</tr>
<tr>
<td>1,000</td>
<td>564</td>
</tr>
<tr>
<td>2,000</td>
<td>693</td>
</tr>
<tr>
<td>2,500</td>
<td>726</td>
</tr>
<tr>
<td>5,000</td>
<td>801</td>
</tr>
<tr>
<td>10,000</td>
<td>844</td>
</tr>
<tr>
<td>15,000</td>
<td>859</td>
</tr>
<tr>
<td>20,000</td>
<td>867</td>
</tr>
<tr>
<td>25,000</td>
<td>872</td>
</tr>
<tr>
<td>50,000</td>
<td>881</td>
</tr>
<tr>
<td>1,000,000</td>
<td>891</td>
</tr>
</tbody>
</table>

$d^* = 2$
$Z(1-\alpha) = 2.326$

### Table 2D.

**LOAS Single-Stage Sample Sizes for $\alpha = 0.01$ and $d^* = 3$**

<table>
<thead>
<tr>
<th>$\text{N}$</th>
<th>Hypothesized seroprevalence ($R_0$) — Séroprévalence supposée ($R_0$)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.01</td>
</tr>
<tr>
<td>100</td>
<td>#NUM!</td>
</tr>
<tr>
<td>200</td>
<td>#NUM!</td>
</tr>
<tr>
<td>1,000</td>
<td>650</td>
</tr>
<tr>
<td>2,000</td>
<td>809</td>
</tr>
<tr>
<td>2,500</td>
<td>848</td>
</tr>
<tr>
<td>5,000</td>
<td>940</td>
</tr>
<tr>
<td>10,000</td>
<td>992</td>
</tr>
<tr>
<td>15,000</td>
<td>1011</td>
</tr>
<tr>
<td>20,000</td>
<td>1021</td>
</tr>
<tr>
<td>25,000</td>
<td>1026</td>
</tr>
<tr>
<td>50,000</td>
<td>1038</td>
</tr>
<tr>
<td>1,000,000</td>
<td>1050</td>
</tr>
</tbody>
</table>

$d^* = 3$
$Z(1-\alpha) = 2.326$

# NUM!: not possible to calculate sample size — taille de l'échantillon non calculable.
TABLE 2E. LOAS SINGLE-STAGE SAMPLE SIZES FOR ALPHA = 0.01 AND d* = 4

<table>
<thead>
<tr>
<th>N</th>
<th>Hypothesized seroprevalence (P0) — Séroprévalence supposée (P0)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.01</td>
</tr>
<tr>
<td>100</td>
<td>#NUM!</td>
</tr>
<tr>
<td>200</td>
<td>#NUM!</td>
</tr>
<tr>
<td>1,000</td>
<td>727</td>
</tr>
<tr>
<td>2,000</td>
<td>916</td>
</tr>
<tr>
<td>2,500</td>
<td>964</td>
</tr>
<tr>
<td>5,000</td>
<td>1072</td>
</tr>
<tr>
<td>10,000</td>
<td>1134</td>
</tr>
<tr>
<td>15,000</td>
<td>1156</td>
</tr>
<tr>
<td>20,000</td>
<td>1168</td>
</tr>
<tr>
<td>25,000</td>
<td>1175</td>
</tr>
<tr>
<td>50,000</td>
<td>1189</td>
</tr>
<tr>
<td>1,000,000</td>
<td>1202</td>
</tr>
</tbody>
</table>

TABLE 2F. LOAS SINGLE-STAGE SAMPLE SIZES FOR ALPHA = 0.01 AND d* = 5

<table>
<thead>
<tr>
<th>N</th>
<th>Hypothesized seroprevalence (P0) — Séroprévalence supposée (P0)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.01</td>
</tr>
<tr>
<td>100</td>
<td>#NUM!</td>
</tr>
<tr>
<td>200</td>
<td>#NUM!</td>
</tr>
<tr>
<td>1,000</td>
<td>796</td>
</tr>
<tr>
<td>2,000</td>
<td>1018</td>
</tr>
<tr>
<td>2,500</td>
<td>1073</td>
</tr>
<tr>
<td>5,000</td>
<td>1199</td>
</tr>
<tr>
<td>10,000</td>
<td>1271</td>
</tr>
<tr>
<td>15,000</td>
<td>1297</td>
</tr>
<tr>
<td>20,000</td>
<td>1310</td>
</tr>
<tr>
<td>25,000</td>
<td>1318</td>
</tr>
<tr>
<td>50,000</td>
<td>1334</td>
</tr>
<tr>
<td>1,000,000</td>
<td>1350</td>
</tr>
</tbody>
</table>

# NUM!: not possible to calculate sample size — taille de l’échantillon non calculable.
<table>
<thead>
<tr>
<th>Po</th>
<th>Pa</th>
<th>n</th>
<th>d</th>
<th>Po</th>
<th>Pa</th>
<th>n</th>
<th>d</th>
<th>Po</th>
<th>Pa</th>
<th>n</th>
<th>d</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.01</td>
<td>0.001</td>
<td>448</td>
<td>1</td>
<td>0.05</td>
<td>0.005</td>
<td>87</td>
<td>0</td>
<td>0.15</td>
<td>0.015</td>
<td>27</td>
<td>0</td>
</tr>
<tr>
<td>0.02</td>
<td>0.002</td>
<td>634</td>
<td>2</td>
<td>0.10</td>
<td>0.010</td>
<td>123</td>
<td>2</td>
<td>0.30</td>
<td>0.030</td>
<td>38</td>
<td>2</td>
</tr>
<tr>
<td>0.03</td>
<td>0.003</td>
<td>898</td>
<td>4</td>
<td>0.15</td>
<td>0.015</td>
<td>174</td>
<td>3</td>
<td>0.45</td>
<td>0.045</td>
<td>53</td>
<td>3</td>
</tr>
<tr>
<td>0.04</td>
<td>0.004</td>
<td>1306</td>
<td>7</td>
<td>0.20</td>
<td>0.020</td>
<td>253</td>
<td>6</td>
<td>0.60</td>
<td>0.060</td>
<td>77</td>
<td>6</td>
</tr>
<tr>
<td>0.05</td>
<td>0.005</td>
<td>1991</td>
<td>12</td>
<td>0.25</td>
<td>0.025</td>
<td>385</td>
<td>12</td>
<td>0.75</td>
<td>0.075</td>
<td>117</td>
<td>11</td>
</tr>
<tr>
<td>0.06</td>
<td>0.006</td>
<td>3269</td>
<td>23</td>
<td>0.30</td>
<td>0.030</td>
<td>631</td>
<td>22</td>
<td>0.90</td>
<td>0.090</td>
<td>191</td>
<td>20</td>
</tr>
<tr>
<td>0.07</td>
<td>0.007</td>
<td>6078</td>
<td>48</td>
<td>0.35</td>
<td>0.035</td>
<td>1171</td>
<td>46</td>
<td>0.95</td>
<td>0.105</td>
<td>354</td>
<td>41</td>
</tr>
<tr>
<td>0.08</td>
<td>0.008</td>
<td>14243</td>
<td>122</td>
<td>0.40</td>
<td>0.040</td>
<td>2741</td>
<td>118</td>
<td>1.00</td>
<td>0.120</td>
<td>824</td>
<td>106</td>
</tr>
<tr>
<td>0.09</td>
<td>0.009</td>
<td>59144</td>
<td>551</td>
<td>0.45</td>
<td>0.045</td>
<td>11367</td>
<td>530</td>
<td>1.05</td>
<td>0.135</td>
<td>3404</td>
<td>476</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.02</td>
<td>0.002</td>
<td>222</td>
<td>1</td>
<td>0.075</td>
<td>0.008</td>
<td>57</td>
<td>0</td>
<td>0.15</td>
<td>0.018</td>
<td>22</td>
<td>0</td>
</tr>
<tr>
<td>0.04</td>
<td>0.004</td>
<td>314</td>
<td>2</td>
<td>0.10</td>
<td>0.015</td>
<td>80</td>
<td>2</td>
<td>0.30</td>
<td>0.035</td>
<td>32</td>
<td>2</td>
</tr>
<tr>
<td>0.06</td>
<td>0.006</td>
<td>445</td>
<td>4</td>
<td>0.15</td>
<td>0.023</td>
<td>114</td>
<td>3</td>
<td>0.45</td>
<td>0.053</td>
<td>45</td>
<td>3</td>
</tr>
<tr>
<td>0.08</td>
<td>0.008</td>
<td>648</td>
<td>7</td>
<td>0.20</td>
<td>0.030</td>
<td>165</td>
<td>6</td>
<td>0.60</td>
<td>0.070</td>
<td>64</td>
<td>6</td>
</tr>
<tr>
<td>0.10</td>
<td>0.010</td>
<td>987</td>
<td>12</td>
<td>0.25</td>
<td>0.038</td>
<td>251</td>
<td>11</td>
<td>0.75</td>
<td>0.088</td>
<td>98</td>
<td>10</td>
</tr>
<tr>
<td>0.12</td>
<td>0.012</td>
<td>1620</td>
<td>23</td>
<td>0.30</td>
<td>0.045</td>
<td>411</td>
<td>22</td>
<td>0.90</td>
<td>0.105</td>
<td>160</td>
<td>20</td>
</tr>
<tr>
<td>0.14</td>
<td>0.014</td>
<td>3011</td>
<td>47</td>
<td>0.35</td>
<td>0.053</td>
<td>762</td>
<td>45</td>
<td>0.95</td>
<td>0.123</td>
<td>295</td>
<td>40</td>
</tr>
<tr>
<td>0.16</td>
<td>0.016</td>
<td>7054</td>
<td>121</td>
<td>0.40</td>
<td>0.060</td>
<td>1783</td>
<td>115</td>
<td>1.00</td>
<td>0.140</td>
<td>687</td>
<td>103</td>
</tr>
<tr>
<td>0.18</td>
<td>0.018</td>
<td>29283</td>
<td>546</td>
<td>0.45</td>
<td>0.068</td>
<td>7386</td>
<td>516</td>
<td>1.05</td>
<td>0.158</td>
<td>2835</td>
<td>462</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.03</td>
<td>0.003</td>
<td>147</td>
<td>1</td>
<td>0.1</td>
<td>0.010</td>
<td>42</td>
<td>0</td>
<td>0.2</td>
<td>0.020</td>
<td>19</td>
<td>0</td>
</tr>
<tr>
<td>0.06</td>
<td>0.006</td>
<td>208</td>
<td>2</td>
<td>0.2</td>
<td>0.020</td>
<td>59</td>
<td>2</td>
<td>0.4</td>
<td>0.040</td>
<td>27</td>
<td>1</td>
</tr>
<tr>
<td>0.09</td>
<td>0.009</td>
<td>295</td>
<td>4</td>
<td>0.3</td>
<td>0.030</td>
<td>83</td>
<td>3</td>
<td>0.6</td>
<td>0.060</td>
<td>38</td>
<td>3</td>
</tr>
<tr>
<td>0.12</td>
<td>0.012</td>
<td>428</td>
<td>7</td>
<td>0.4</td>
<td>0.040</td>
<td>121</td>
<td>6</td>
<td>0.8</td>
<td>0.080</td>
<td>55</td>
<td>6</td>
</tr>
<tr>
<td>0.15</td>
<td>0.015</td>
<td>652</td>
<td>12</td>
<td>0.5</td>
<td>0.050</td>
<td>184</td>
<td>11</td>
<td>1.0</td>
<td>0.100</td>
<td>83</td>
<td>10</td>
</tr>
<tr>
<td>0.18</td>
<td>0.018</td>
<td>1071</td>
<td>22</td>
<td>0.6</td>
<td>0.060</td>
<td>301</td>
<td>21</td>
<td>1.2</td>
<td>0.120</td>
<td>136</td>
<td>19</td>
</tr>
<tr>
<td>0.21</td>
<td>0.021</td>
<td>1989</td>
<td>47</td>
<td>0.7</td>
<td>0.070</td>
<td>558</td>
<td>44</td>
<td>1.4</td>
<td>0.140</td>
<td>251</td>
<td>39</td>
</tr>
<tr>
<td>0.24</td>
<td>0.024</td>
<td>4658</td>
<td>120</td>
<td>0.8</td>
<td>0.080</td>
<td>1303</td>
<td>112</td>
<td>1.6</td>
<td>0.160</td>
<td>585</td>
<td>101</td>
</tr>
<tr>
<td>0.27</td>
<td>0.027</td>
<td>19330</td>
<td>540</td>
<td>0.9</td>
<td>0.090</td>
<td>5395</td>
<td>503</td>
<td>1.8</td>
<td>0.180</td>
<td>2409</td>
<td>449</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.04</td>
<td>0.004</td>
<td>109</td>
<td>1</td>
<td>0.125</td>
<td>0.013</td>
<td>33</td>
<td>0</td>
<td>0.25</td>
<td>0.023</td>
<td>17</td>
<td>0</td>
</tr>
<tr>
<td>0.08</td>
<td>0.008</td>
<td>155</td>
<td>2</td>
<td>0.15</td>
<td>0.025</td>
<td>46</td>
<td>2</td>
<td>0.45</td>
<td>0.045</td>
<td>23</td>
<td>1</td>
</tr>
<tr>
<td>0.12</td>
<td>0.012</td>
<td>219</td>
<td>3</td>
<td>0.20</td>
<td>0.038</td>
<td>65</td>
<td>3</td>
<td>0.68</td>
<td>0.068</td>
<td>33</td>
<td>3</td>
</tr>
<tr>
<td>0.16</td>
<td>0.016</td>
<td>319</td>
<td>6</td>
<td>0.25</td>
<td>0.050</td>
<td>95</td>
<td>6</td>
<td>0.90</td>
<td>0.090</td>
<td>48</td>
<td>5</td>
</tr>
<tr>
<td>0.20</td>
<td>0.020</td>
<td>485</td>
<td>12</td>
<td>0.30</td>
<td>0.063</td>
<td>144</td>
<td>11</td>
<td>1.13</td>
<td>0.113</td>
<td>72</td>
<td>10</td>
</tr>
<tr>
<td>0.24</td>
<td>0.024</td>
<td>796</td>
<td>22</td>
<td>0.35</td>
<td>0.075</td>
<td>235</td>
<td>21</td>
<td>1.35</td>
<td>0.135</td>
<td>118</td>
<td>19</td>
</tr>
<tr>
<td>0.28</td>
<td>0.028</td>
<td>1478</td>
<td>46</td>
<td>0.40</td>
<td>0.088</td>
<td>435</td>
<td>43</td>
<td>1.58</td>
<td>0.158</td>
<td>217</td>
<td>38</td>
</tr>
<tr>
<td>0.32</td>
<td>0.032</td>
<td>3460</td>
<td>119</td>
<td>0.45</td>
<td>0.100</td>
<td>1016</td>
<td>109</td>
<td>1.80</td>
<td>0.180</td>
<td>505</td>
<td>98</td>
</tr>
<tr>
<td>0.36</td>
<td>0.036</td>
<td>14353</td>
<td>535</td>
<td>0.50</td>
<td>0.113</td>
<td>4200</td>
<td>489</td>
<td>2.03</td>
<td>0.203</td>
<td>2077</td>
<td>436</td>
</tr>
</tbody>
</table>
### TABLE 3B. LQAS SINGLE-STAGE SAMPLE SIZES AND CRITICAL VALUES FOR ALPHA = 0.01 AND BETA = 0.20

**TABLEAU 3B. TAILLES D’ÉCHANTILLONS POUR LQAS À 1 DEGRÉ ET VALEURS CRITIQUES POUR ALPHA = 0.01 ET BETA = 0.20**

| \(Z_{(95)}\) | 2.33 |
| \(Z_{(80)}\) | 0.84 |

<table>
<thead>
<tr>
<th>(P_{o} = 0.01)</th>
<th>(P_{o} = 0.05)</th>
<th>(P_{o} = 0.15)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(P_{a} = n_{d})</td>
<td>(n_{d})</td>
<td>(n_{d})</td>
</tr>
<tr>
<td>0.001</td>
<td>823</td>
<td>1</td>
</tr>
<tr>
<td>0.002</td>
<td>1132</td>
<td>3</td>
</tr>
<tr>
<td>0.003</td>
<td>1572</td>
<td>6</td>
</tr>
<tr>
<td>0.004</td>
<td>2250</td>
<td>11</td>
</tr>
<tr>
<td>0.005</td>
<td>3384</td>
<td>20</td>
</tr>
<tr>
<td>0.006</td>
<td>5494</td>
<td>37</td>
</tr>
<tr>
<td>0.007</td>
<td>10110</td>
<td>77</td>
</tr>
<tr>
<td>0.008</td>
<td>23477</td>
<td>199</td>
</tr>
<tr>
<td>0.009</td>
<td>96692</td>
<td>894</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>(P_{o} = 0.02)</th>
<th>(P_{o} = 0.075)</th>
<th>(P_{o} = 0.175)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(P_{a} = n_{d})</td>
<td>(n_{d})</td>
<td>(n_{d})</td>
</tr>
<tr>
<td>0.002</td>
<td>408</td>
<td>1</td>
</tr>
<tr>
<td>0.004</td>
<td>561</td>
<td>3</td>
</tr>
<tr>
<td>0.006</td>
<td>779</td>
<td>6</td>
</tr>
<tr>
<td>0.008</td>
<td>1115</td>
<td>11</td>
</tr>
<tr>
<td>0.010</td>
<td>1677</td>
<td>20</td>
</tr>
<tr>
<td>0.012</td>
<td>2722</td>
<td>37</td>
</tr>
<tr>
<td>0.014</td>
<td>5008</td>
<td>77</td>
</tr>
<tr>
<td>0.016</td>
<td>11626</td>
<td>197</td>
</tr>
<tr>
<td>0.018</td>
<td>47871</td>
<td>886</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>(P_{o} = 0.03)</th>
<th>(P_{o} = 0.1)</th>
<th>(P_{o} = 0.2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(P_{a} = n_{d})</td>
<td>(n_{d})</td>
<td>(n_{d})</td>
</tr>
<tr>
<td>0.003</td>
<td>270</td>
<td>1</td>
</tr>
<tr>
<td>0.006</td>
<td>371</td>
<td>3</td>
</tr>
<tr>
<td>0.009</td>
<td>515</td>
<td>6</td>
</tr>
<tr>
<td>0.012</td>
<td>737</td>
<td>11</td>
</tr>
<tr>
<td>0.015</td>
<td>1108</td>
<td>20</td>
</tr>
<tr>
<td>0.018</td>
<td>1798</td>
<td>37</td>
</tr>
<tr>
<td>0.021</td>
<td>3307</td>
<td>76</td>
</tr>
<tr>
<td>0.024</td>
<td>7676</td>
<td>195</td>
</tr>
<tr>
<td>0.027</td>
<td>31597</td>
<td>877</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>(P_{o} = 0.04)</th>
<th>(P_{o} = 0.125)</th>
<th>(P_{o} = 0.225)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(P_{a} = n_{d})</td>
<td>(n_{d})</td>
<td>(n_{d})</td>
</tr>
<tr>
<td>0.004</td>
<td>200</td>
<td>1</td>
</tr>
<tr>
<td>0.008</td>
<td>276</td>
<td>3</td>
</tr>
<tr>
<td>0.012</td>
<td>383</td>
<td>6</td>
</tr>
<tr>
<td>0.016</td>
<td>548</td>
<td>11</td>
</tr>
<tr>
<td>0.020</td>
<td>823</td>
<td>19</td>
</tr>
<tr>
<td>0.024</td>
<td>1336</td>
<td>36</td>
</tr>
<tr>
<td>0.028</td>
<td>2457</td>
<td>75</td>
</tr>
<tr>
<td>0.032</td>
<td>5701</td>
<td>193</td>
</tr>
<tr>
<td>0.036</td>
<td>23460</td>
<td>868</td>
</tr>
</tbody>
</table>

*Rapp. trimest. statist. sanit. mond., 44 (1991)*
### Table 4A. Two-stage sample sizes and critical values for \( \alpha = 0.05 \) and \( \beta = 0.20 \)

<table>
<thead>
<tr>
<th>( \beta )</th>
<th>( n_1 )</th>
<th>( n_2 )</th>
<th>( n_1 + n_2 )</th>
<th>( d_1 )</th>
<th>( d_2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.005</td>
<td>155</td>
<td>3</td>
<td>130</td>
<td>2</td>
<td>130</td>
</tr>
<tr>
<td>0.010</td>
<td>1306</td>
<td>12</td>
<td>130</td>
<td>3</td>
<td>130</td>
</tr>
<tr>
<td>0.015</td>
<td>1991</td>
<td>12</td>
<td>130</td>
<td>6</td>
<td>130</td>
</tr>
<tr>
<td>0.020</td>
<td>485</td>
<td>22</td>
<td>199</td>
<td>14</td>
<td>199</td>
</tr>
<tr>
<td>0.025</td>
<td>796</td>
<td>46</td>
<td>370</td>
<td>21</td>
<td>370</td>
</tr>
<tr>
<td>0.030</td>
<td>1476</td>
<td>119</td>
<td>865</td>
<td>21</td>
<td>865</td>
</tr>
<tr>
<td>0.035</td>
<td>3460</td>
<td>553</td>
<td>3388</td>
<td>21</td>
<td>3388</td>
</tr>
<tr>
<td>0.040</td>
<td>14353</td>
<td>535</td>
<td>3588</td>
<td>535</td>
<td>3588</td>
</tr>
<tr>
<td>0.045</td>
<td>631</td>
<td>22</td>
<td>158</td>
<td>1</td>
<td>158</td>
</tr>
<tr>
<td>0.050</td>
<td>235</td>
<td>12</td>
<td>103</td>
<td>0</td>
<td>103</td>
</tr>
<tr>
<td>0.055</td>
<td>191</td>
<td>11</td>
<td>92</td>
<td>1</td>
<td>92</td>
</tr>
<tr>
<td>0.060</td>
<td>121</td>
<td>6</td>
<td>72</td>
<td>0</td>
<td>72</td>
</tr>
<tr>
<td>0.065</td>
<td>88</td>
<td>3</td>
<td>55</td>
<td>0</td>
<td>55</td>
</tr>
<tr>
<td>0.070</td>
<td>19</td>
<td>1</td>
<td>25</td>
<td>0</td>
<td>25</td>
</tr>
<tr>
<td>0.075</td>
<td>5</td>
<td>0</td>
<td>13</td>
<td>0</td>
<td>13</td>
</tr>
<tr>
<td>0.080</td>
<td>2</td>
<td>0</td>
<td>7</td>
<td>0</td>
<td>7</td>
</tr>
</tbody>
</table>

### Tableau 4A. Tailles d'échantillons pour \( \alpha = 0.01 \) et \( \beta = 0.20 \)

<table>
<thead>
<tr>
<th>( \beta )</th>
<th>( n_1 )</th>
<th>( n_2 )</th>
<th>( n_1 + n_2 )</th>
<th>( d_1 )</th>
<th>( d_2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.005</td>
<td>636</td>
<td>2</td>
<td>536</td>
<td>2</td>
<td>536</td>
</tr>
<tr>
<td>0.010</td>
<td>898</td>
<td>4</td>
<td>536</td>
<td>4</td>
<td>536</td>
</tr>
<tr>
<td>0.015</td>
<td>1306</td>
<td>7</td>
<td>536</td>
<td>7</td>
<td>536</td>
</tr>
<tr>
<td>0.020</td>
<td>1991</td>
<td>12</td>
<td>536</td>
<td>12</td>
<td>536</td>
</tr>
<tr>
<td>0.025</td>
<td>485</td>
<td>23</td>
<td>817</td>
<td>23</td>
<td>817</td>
</tr>
<tr>
<td>0.030</td>
<td>796</td>
<td>48</td>
<td>139</td>
<td>48</td>
<td>139</td>
</tr>
<tr>
<td>0.035</td>
<td>1476</td>
<td>122</td>
<td>1520</td>
<td>122</td>
<td>1520</td>
</tr>
<tr>
<td>0.040</td>
<td>3460</td>
<td>553</td>
<td>3388</td>
<td>553</td>
<td>3388</td>
</tr>
<tr>
<td>0.045</td>
<td>14353</td>
<td>535</td>
<td>3588</td>
<td>535</td>
<td>3588</td>
</tr>
<tr>
<td>0.050</td>
<td>631</td>
<td>23</td>
<td>158</td>
<td>23</td>
<td>158</td>
</tr>
<tr>
<td>0.055</td>
<td>235</td>
<td>12</td>
<td>103</td>
<td>12</td>
<td>103</td>
</tr>
<tr>
<td>0.060</td>
<td>191</td>
<td>11</td>
<td>92</td>
<td>11</td>
<td>92</td>
</tr>
<tr>
<td>0.065</td>
<td>88</td>
<td>3</td>
<td>55</td>
<td>3</td>
<td>55</td>
</tr>
<tr>
<td>0.070</td>
<td>19</td>
<td>1</td>
<td>25</td>
<td>1</td>
<td>25</td>
</tr>
<tr>
<td>0.075</td>
<td>5</td>
<td>0</td>
<td>13</td>
<td>0</td>
<td>13</td>
</tr>
<tr>
<td>0.080</td>
<td>2</td>
<td>0</td>
<td>7</td>
<td>0</td>
<td>7</td>
</tr>
</tbody>
</table>

### Table 4A. Tailles d'échantillons pour \( \alpha = 0.05 \) et \( \beta = 0.20 \)

<table>
<thead>
<tr>
<th>( \beta )</th>
<th>( n_1 )</th>
<th>( n_2 )</th>
<th>( n_1 + n_2 )</th>
<th>( d_1 )</th>
<th>( d_2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.005</td>
<td>634</td>
<td>2</td>
<td>536</td>
<td>2</td>
<td>536</td>
</tr>
<tr>
<td>0.010</td>
<td>898</td>
<td>4</td>
<td>536</td>
<td>4</td>
<td>536</td>
</tr>
<tr>
<td>0.015</td>
<td>1306</td>
<td>7</td>
<td>536</td>
<td>7</td>
<td>536</td>
</tr>
<tr>
<td>0.020</td>
<td>1991</td>
<td>12</td>
<td>536</td>
<td>12</td>
<td>536</td>
</tr>
<tr>
<td>0.025</td>
<td>485</td>
<td>23</td>
<td>817</td>
<td>23</td>
<td>817</td>
</tr>
<tr>
<td>0.030</td>
<td>796</td>
<td>48</td>
<td>139</td>
<td>48</td>
<td>139</td>
</tr>
<tr>
<td>0.035</td>
<td>1476</td>
<td>122</td>
<td>1520</td>
<td>122</td>
<td>1520</td>
</tr>
<tr>
<td>0.040</td>
<td>3460</td>
<td>553</td>
<td>3388</td>
<td>553</td>
<td>3388</td>
</tr>
<tr>
<td>0.045</td>
<td>14353</td>
<td>535</td>
<td>3588</td>
<td>535</td>
<td>3588</td>
</tr>
<tr>
<td>0.050</td>
<td>631</td>
<td>23</td>
<td>158</td>
<td>23</td>
<td>158</td>
</tr>
<tr>
<td>0.055</td>
<td>235</td>
<td>12</td>
<td>103</td>
<td>12</td>
<td>103</td>
</tr>
<tr>
<td>0.060</td>
<td>191</td>
<td>11</td>
<td>92</td>
<td>11</td>
<td>92</td>
</tr>
<tr>
<td>0.065</td>
<td>88</td>
<td>3</td>
<td>55</td>
<td>3</td>
<td>55</td>
</tr>
<tr>
<td>0.070</td>
<td>19</td>
<td>1</td>
<td>25</td>
<td>1</td>
<td>25</td>
</tr>
<tr>
<td>0.075</td>
<td>5</td>
<td>0</td>
<td>13</td>
<td>0</td>
<td>13</td>
</tr>
<tr>
<td>0.080</td>
<td>2</td>
<td>0</td>
<td>7</td>
<td>0</td>
<td>7</td>
</tr>
</tbody>
</table>
### TABLE 4B. LOAS TWO-STAGE SAMPLE SIZES AND CRITICAL VALUES FOR ALPHA = 0.05 AND BETA = 0.20

TABLEAU 4B. TAILLES D'ÉCHANTILLONS POUR LOAS À 2 DEGRÉS ET VALEURS CRITIQUES POUR ALPHA = 0.01 ET BETA = 0.20

<table>
<thead>
<tr>
<th>Po = 0.2</th>
<th>Po = 0.4</th>
<th>Po = 0.6</th>
</tr>
</thead>
<tbody>
<tr>
<td>( n_1 + n_2 )</td>
<td>( d_2 )</td>
<td>( n_1 )</td>
</tr>
<tr>
<td>0.040</td>
<td>27</td>
<td>1</td>
</tr>
<tr>
<td>0.060</td>
<td>38</td>
<td>3</td>
</tr>
<tr>
<td>0.080</td>
<td>55</td>
<td>6</td>
</tr>
<tr>
<td>0.100</td>
<td>83</td>
<td>10</td>
</tr>
<tr>
<td>0.120</td>
<td>136</td>
<td>19</td>
</tr>
<tr>
<td>0.140</td>
<td>251</td>
<td>39</td>
</tr>
<tr>
<td>0.160</td>
<td>585</td>
<td>101</td>
</tr>
<tr>
<td>0.180</td>
<td>2409</td>
<td>449</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Po = 0.3</th>
<th>Po = 0.5</th>
<th>Po = 0.7</th>
</tr>
</thead>
<tbody>
<tr>
<td>( n_1 + n_2 )</td>
<td>( d_2 )</td>
<td>( n_1 )</td>
</tr>
<tr>
<td>0.060</td>
<td>16</td>
<td>1</td>
</tr>
<tr>
<td>0.090</td>
<td>23</td>
<td>3</td>
</tr>
<tr>
<td>0.120</td>
<td>33</td>
<td>5</td>
</tr>
<tr>
<td>0.150</td>
<td>50</td>
<td>9</td>
</tr>
<tr>
<td>0.180</td>
<td>81</td>
<td>17</td>
</tr>
<tr>
<td>0.210</td>
<td>149</td>
<td>35</td>
</tr>
<tr>
<td>0.240</td>
<td>345</td>
<td>89</td>
</tr>
<tr>
<td>0.270</td>
<td>1413</td>
<td>395</td>
</tr>
</tbody>
</table>
Single sampling example when only $P_0$ is specified

Fig. 3 presents the normal density function and illustrates the sampling distribution of $d$ and the environment in which we are hypothesis testing. We see that the expected number of infected people in our sample $nP_0$ is at the centre of the null distribution, while the variability of $d$ is implicitly represented by the "spread" of the distribution.

Suppose we are interested in detecting communities with $P \geq 0.20$ and wish to limit the chance of a type I error to $\alpha = 0.05$. We will need to specify $N$ and $d^*$ in order to calculate $n$. Assume that $N = 50,000$ and we want to conclude that the community has $P < 0.20$ if 4 or fewer people in our sample prove to be seropositive. Therefore, we would set $d^* = 4$. Sample size is calculated by solving the following formula for $n$.

This involves tedious calculations with a quadratic formula:

$$Z_{1-\alpha} = \frac{d^* - nP_0}{\sqrt{nP_0(1-P_0)\left(\frac{N-n}{N-1}\right)}}$$

Table 1 presents, for $\alpha = 0.05$, sample size calculations once $N$, $P_0$ and $d^*$ have been specified. (Table 2 presents comparable results for $\alpha = 0.01$.) To use Table 1 in our example, select the table which corresponds to $d^* = 4$. Then locate the column for $P_0 = 0.20$ and the row for $N = 50,000$. The corresponding sample size is 42. If you want to sample fewer people you would choose a smaller value of $d^*$.

Single sampling example when $P_0$, $P_a$, and $\beta$ are specified

A graphical presentation of this scenario is presented in Fig. 4 where the sample size is reflected in both the location of the null and alternative distributions, and the "spread" of the distributions.

The expected number of infected people under the alternative hypothesis $nP_a$ is at the centre of the alternative distribution. There is only $\alpha$ probability of observing $d \leq d^*$, if we are in fact sampling from a population with $P_a$ seroprevalence or greater, while there is $\beta$ probability of observing $d \geq d^*$ if we are sampling from the alternative distribution. Note that all the same parameters are incorporated in the normal formulation as are represented in the exact probability estimation using the hypergeometric distribution.

Let us develop the example mentioned previously. Suppose we are interested in detecting communities with $P \geq 0.20$ and require 80% power to ensure that communities with $P \leq 0.10$ will not be wrongly classified as high HIV seroprevalence areas. We will
set \( \alpha = 0.05 \) and assume the community to have a large value for \( N \).

The values of \( \alpha \) and \( \beta \) are considered to be constraints that our ultimate choice of \( n \) and \( d^* \) will have to satisfy. In this figure \( d^* \) represents the lower \( \alpha \)th percentile of the sampling distribution centered at \( n\beta \):

\[
d^* = n\beta - z_{1-\alpha} \sqrt{n\beta (1-\beta) \left( \frac{N-n}{N-1} \right)}
\]

and represents the upper \( \beta \)th percentile of the sampling distribution centered at \( n\beta^* \):

\[
d^* = n\beta + z_{1-\beta} \sqrt{n\beta (1-\beta) \left( \frac{N-n}{N-1} \right)}
\]

There is only one value for sample size that will satisfy both these expressions. The following formula for sample size is obtained by setting the two equations for \( d^* \) equal to each other, assuming that the finite population correction equals 1, and solving for \( n \):

\[
n = \left( \frac{z_{1-\alpha} \sqrt{\beta (1-\beta) \left( \frac{N-n}{N-1} \right)}}{n\beta - n\beta^*} \right)^2
\]

For our example, the calculations follow:

\[
n = \left( \frac{1.645 \sqrt{(0.20)(1-0.20)}}{0.38} \right)^2 = 82.9
\]

\[
d^* = (83)(0.20) - 1.645 \sqrt{(83)(0.20)(1-0.20)} = 10.6
\]

Therefore, we will choose a sample of size 83 and accept the community as a low HIV seroprevalence area (i.e. reject \( H_0 \)) only if 10 or fewer people in the sample prove to be HIV seropositive. Note that the tabulated values for \( n \) are rounded up while the values for \( d^* \) are rounded down.

These figures can be found in Table 3A (\( \alpha = 0.05 \)). Locate the table that corresponds to \( P_0 = 0.38 \) and move down to the row corresponding to \( P_\beta = 0.1 \) (Table 3B presents results for \( \alpha = 0.01 \)).

**LOAS double sampling – the theory**

Double sampling, or two-stage sampling, is a sampling plan which may be used to decrease sampling costs of a study. In this scenario, two critical values, \( d_1^* \) and \( d_2^* \), where \( d_1^* \leq d_2^* \), are designated and two samples of sizes, \( n_1 \) and \( n_2 \), are specified. At the first stage we study \( n_1 \) individuals. If the observed number of infected people is less than or equal to \( d_1^* \), we would conclude that the actual proportion infected in the population is significantly less than \( P_0 \). If the observed number of infected people in the first stage sample of size \( n_1 \) is greater than \( d_2^* \), we would conclude that the actual proportion infected in the population is not significantly less than \( P_0 \). If the observed number of infected people in the first stage is greater than \( d_1^* \) but less than or equal to \( d_2^* \), we then proceed to the second stage. At the second stage we continue sampling until either \( d_2^* + 1 \) infected people are observed, indicating a high seroprevalence level, or a total of \( n_1 + n_2 \) people have been sampled without exceeding \( d_2^* \) infected people, indicating a low seroprevalence level.

Rather than thinking of double sampling as a single-sampling plan followed by an additional sample stage, it will be important to realize that the total sample size and critical value of the two stages correspond to the single sample plan outlined previously, while the first stage represents a preliminary "reduced" sample. Herein lies the source of increased economy offered by double sampling. When the results are extreme in the first stage, a health planner can make a conclusion based on fewer subjects sampled than if a one-stage sampling plan had been used. Otherwise, the health planner will continue sampling, knowing that the most that will be sampled will equal the sample size from a one-stage plan. In other words, utilizing a two-stage sampling plan will ensure that the number of people sampled will be less than or equal to that in a one-stage sampling plan.

Sample size and critical value determination for the second stage, namely \( n_1 + n_2 \) and \( d_2^* \), is carried out identically to the single-sampling situation developed above. For example, under the scenario in which the health planner wishes to control the size of the \( \beta \) error, \( n_1 + n_2 \) and \( d_2^* \) would be chosen such that the probability of observing \( d_2^* \) or fewer infected people out of our sample of \( n_1 + n_2 \) people is less than a specified \( \alpha \) level, say 5%, while the probability of observing more than \( d_2^* \) infected people is the \( \beta \) level for some \( P_\beta \). Under the scenario in which no attempt is made to control the \( \beta \) error, a fairly high value for \( d_2^* \) would be designated, and the appropriate sample size relative to \( \alpha = 0.05 \) would be calculated as indicated above.

However, in order to determine \( n_1 \) and \( d_1^* \), it is necessary to identify values that correspond to more restrictive criteria. In other words, we need to choose combinations of \( n_1 \) and \( d_1^* \) such that the probability of observing \( d_1^* \) or fewer infected people out of the preliminary sample of \( n_1 \) people is less than some more restrictive \( \alpha \) level, perhaps 1% rather than the 5% used for choosing the second-stage sample size and critical value. The more stringent rejection criteria for the first stage reflect our desire to conclude the community to be a low HIV seroprevalent area only if we observe an exceptionally low number of infected people in our sample relative to that expected under the null hypothesis. Furthermore, if it is necessary to sample at both stages, it assures that the overall chance of making a type I error somewhere in the two hypothesis tests performed is less than or equal to 1% + 5% = 6%.

Determination of \( n_1 \) and \( d_1^* \) is identical to sample size and critical value determination for a single-sampling plan in which \( \beta \) and \( P_\beta \) are ignored, as shown above. Specifically, the health planner will choose a value of \( d_1^* \) which is less than \( d_2^* \) and will imply a value for \( n_1 \) via the same calculation previously mentioned. The only difference will be the more restrictive \( \alpha \) level.

**Double-sampling example when \( P_0 \), \( P_\beta \) and \( \beta \) are specified**

We will continue to develop the example discussed above. However, now we will consider a double-
sampling plan. The calculations for \( n_1 + n_2 \) and \( d_j^* \) are identical to those for \( n \) and \( d \) above. We are choosing a combination of \( n \) and \( d \) such that there is a 5% chance that we will conclude the community to be a low HIV area when it really is high, and a 20% chance that we will conclude the community to be a high HIV area when it really is low. Assuming the health planner is interested in \( P_a = 0.20 \) and \( P_o = 0.10 \), we will arrive at the following results in exactly the same manner as before:

\[
\begin{align*}
\left( n_1 + n_2 \right) & = 83 \\
\left( d_2 \right) & = 10
\end{align*}
\]

In Table 4, values of \( n_1 \) are 25% of \( n_1 + n_2 \), and \( d_1^* \) is calculated as follows:

\[
d_1^* = n_1 P_o - Z_{0.01} \sqrt{n_1 P_o (1-P_o)}.
\]

Notice that \( d_1^* \) is computed using \( \alpha = 0.01 \), reflecting our desire to accept a community as having low HIV seroprevalence only when substantially fewer infected subjects are sampled than would be expected under \( H_o \). However, for many combinations of \( P_a \) and \( P_o \), this rule yields a negative value for \( d_1^* \). In these cases, we set \( d_1^* = 0 \) and calculate \( n_1 \) as shown above. Note that this rule may lead to first-stage sample sizes similar to the combined two-stage sample size. In this case, double sampling may not be advantageous. Some practical considerations are briefly discussed below.

In our example, we need to set \( d_1^* = 0 \) and calculate \( n_1 \) as discussed above. These calculations yield the following results:

\[
\begin{align*}
n_1 & = 22 \\
d_1^* & = 0
\end{align*}
\]

These values can be found in Table 4B by locating the subtable for \( P_a = 0.20 \) and the row for \( P_o = 0.10 \).

**Practical considerations**

LOAS is a powerful and economical procedure when sampling costs per subject are high, and when there is a designated target. A target makes sense when, for example, programme planners are trying to achieve a specified goal (e.g. a specified level of vaccination coverage, a level of utilization of ORT, etc.). LOAS provides a hypothesis-testing strategy to determine whether the specified goal has been achieved. However, when there is no target and interest focuses on estimating disease prevalence in a population, health planners may be better served by using a more traditional sampling technique (e.g. stratified or cluster sampling) to obtain the most precise confidence interval estimates possible. While our example using HIV seroprevalence was set up as a target situation, in general HIV seroprevalence research may well fall into this second category.

The great advantage of LOAS sampling is that in the situation where there is a target, sampling can stop once \( d_1^* + 1 \) cases have been observed. When the actual population differs widely from the target, this could represent a very large reduction in sample size. In actuality, LOAS is not so much a method of sampling (it is really a stratified random sample) as it is a strategy for analysis. Since it is possible to obtain results quickly, it may encourage programme managers to collect data at more frequent intervals and to take remedial action more quickly when there is evidence that programme targets are not being achieved.

Much of our discussion has focused on the issue of hypothesis testing. If there is interest in estimating the underlying rate in a community, samples from an LOAS sampling plan will typically be too small to yield a reasonably precise estimate. Since LOAS is a special case of stratified sampling, the sample results for all communities in a region may be combined to attain a regional seroprevalence estimate. This procedure is discussed elsewhere.

Double sampling within the context of an LOAS scheme may increase the economy of a health survey under certain conditions. However, if the medical test results cannot be analysed on the spot, then a double-sampling plan may require that a health team return to a community for further sampling if the statistical results are not extreme/conclusive in the first sample. The cost in terms of time and travel may offset the potential gain in economy attained through double sampling. However, if the cost per subject of administering the medical test is high, the health planner may feel that the potential gain from possibly having to carry out only the first stage of sampling makes double sampling preferable.

**SUMMARY**

Lot quality assurance sampling (LOAS) was originally developed for industrial applications. This article discusses the extension of its application to a public health setting. When resources must be allocated to several communities based on whether or not each exceeds a target proportion of members with a characteristic, a sampling scenario arises which is different to the standard parameter estimation scenario. LOAS is an efficient procedure for sampling in this instance. Under certain conditions, double sampling may further enhance the efficiency of an LOAS plan. LOAS is discussed and illustrated by examples, and the theoretical foundations of single and double lot quality assurance sampling are presented. Examples of sample size and critical value tables are also presented.
RÉSUMÉ

Echantillonnage par lots pour l'assurance de la qualité: plans de sondage à un et à deux degrés

L'échantillonnage par lots pour l'assurance de la qualité (lot quality assurance sampling - LQAS) a été initialement mis au point pour des applications industrielles. Cet article étudie la possibilité de l'étendre au domaine de la santé publique. Lorsqu'il faut affecter des ressources à différentes collectivités selon que dans chacune d'entre elles il existe ou non une certaine proportion d'individus présentant telle ou telle caractéristique, il se pose un problème de sondage distinct du problème habituel d'estimation paramétrique. En pareil cas le LQAS constitue une méthode de sondage efficace. Dans certaines conditions, un double sondage peut encore renforcer l'efficacité d'un plan de sondage LQAS. L'article étudie la méthode en l'illustrant par un exemple et en expose les fondements théoriques dans le cas d'un sondage à un degré et d'un sondage à deux degrés. Il donne également des exemples de taille d'échantillon et des tables de valeurs critiques.

REFERENCES — RÉFÉRENCES

Health surveys have been widely utilized in developing countries to provide information to health planners for programme evaluation or monitoring. Traditional survey methods are costly in terms of personnel and time. The results are usually available a considerable time after the survey is completed, making it difficult for health planners to use when making decisions. Moreover, most health surveys only provide information at the regional or national level, not at the level of small health units or population groups, where information is needed in order to direct supervisory activities to those small areas with the poorest health-programme performance. Individual districts do not usually have the resources to carry out health surveys in large and even small areas. Hence, the growing interest in developing methods that are particularly suited to small areas.

The utilization of survey methodology developed in industry for quality control provides tools that could potentially overcome these limitations of traditional health survey methods (1). Among these, lot quality assurance sampling (LQAS) has been proposed as a useful methodology for monitoring health programme performance at both health centre or community level (2, 3). In this article the practical applications of LQAS techniques in health surveys in developing countries will be reviewed, describing their advantages and the current constraints that still limit their use in health monitoring. A detailed description of the methodology and its statistical principles can be obtained elsewhere (3, 4, 5).

After a brief description of the method, each aspect of the LQAS methodology in its application to health surveys will be discussed, addressing current difficulties. The following areas will be reviewed:

- the different ways of defining a lot and a sample unit for its use in LQAS;
- the need to have a precise sampling frame and how this can be created;
- the available methods for calculating the sample size for different sampling schemes;
- a detailed description of how to conduct the sampling, in particular in community surveys, and the key role of the survey team.

Finally, the interpretation of survey results and their use when monitoring health programmes will be discussed. A concluding section will consider the current limitation and potential uses of LQAS in health surveys.

DEFINING THE LOT

The method

LOAS methodology was developed to help manufacturers determine (at a minimum cost) whether their products met a set of quality standards. This was achieved by utilizing small sample sizes. Since industrial production is usually done in batches or lots, the sampling strategy was developed to classify sampling units (lots) into acceptable or unacceptable according to preset quality levels, minimizing the risk of misjudgement—considering the production lot as acceptable when it is not (consumer's risk) or unacceptable when it is (producer's risk). The method selects a sample size, and the maximum number of permissible defects to be found in a sample to consider the lot acceptable. To satisfy the statistical assumptions of the method, the selection of each individual unit to be sampled in each lot has to be done following a random sampling process. It is important to mention that the method does not provide an estimate for the lot sampled; it only classifies it as acceptable or unacceptable.

LOAS methodology utilizes small sample sizes and can be used as frequently as needed to provide information on each lot sampled. It is also possible to combine lots in order to obtain relatively precise estimates of the quality level for the entire area sampled. Because of these advantages, the use of LOAS methodology has been advocated for health monitoring.

In this issue of the World health statistics quarterly, Lemeshow & Taber give a complete description of the LOAS methodology and its application in industry and health areas in general (6). This article will concentrate on the use of LOAS in health surveys to monitor health programmes. Because of the need for a detailed sampling frame, as described later, the application of LOAS to monitor a health programme at the community level would only be justified if repeat surveys are expected to be done in the same area to monitor the programme. This will only occur when a health programme under active implementation, and when multiple surveys can be considered to monitor the progress of the programme, concentrating supervisory skills in the areas (lots) most at need. LOAS should not be considered as the methodology of choice to be used when a single survey is planned at the community level, unless a detailed sampling frame can be easily obtained; in this case, traditional survey methods would be more convenient and efficient.

DEFINING THE LOT

For its application in health surveys, a lot has been defined as a population area assigned to a health unit (2, 5), a health centre, or even health records within a health centre (6). An ideal lot is the smallest unit that could provide meaningful information to the health planner when evaluating a health pro-
programme. For a given lot, it is assumed that the sampling units within the lot have had similar exposure to the health programme under study. If an area assigned to a health unit has had different exposures to a health programme, because of distance to the health centre for example, it would have to be divided into subareas, each called a lot, for sampling purposes. For example, a health unit in charge of an immunization programme in a large population area, involving both urban and peri-urban or rural areas, could be divided into two or more areas or lots, some for urban and others for peri-urban or rural areas.

In this way, programme performance in these different areas could be evaluated independently, allowing the health planner to take specific remedial measures. Lots can be redefined when evaluating a different health programme or when an area has achieved a uniform performance level for the programme being evaluated.

This ideal lot definition, however, needs to be contrasted with practical and financial constraints of the health programme to be evaluated. In the case of a nationwide health programme, it is not practical to operate with a lot structure based on small population groups of health units that would need to be independently evaluated. It is possible to utilize traditional survey methods, like stratified cluster surveys, for sampling large geographical or population-based areas to first identify which region or population-sectors (each requiring the conduct of a complete traditional survey) should be required. Another sampling utilizing LOAS, thus concentrating supervisory activities on those areas only. This combined survey strategy has not yet been tested and will need to be evaluated before it can be recommended.

Experience in the use of LOAS in health surveys is still limited. Small population bases in both urban (2) and rural (5) areas have been used as lots in experimental evaluations of LOAS in monitoring immunization programmes in Peru. In a large evaluation of LOAS in Costa Rica, an initial random sample of 60 health units at the national level was taken and the population assigned to those health units was divided into one lot and sampled using LOAS methodology (7). Since only one round of surveys was conducted, a complete evaluation of the LOAS methodology was not properly carried out.

Sampling frame

When sampling, one needs to make sure that each sampling unit has been drawn randomly, ensuring that the probability of selecting the sampling unit has been equal for all similar units within the lot. In order to satisfy that requirement, a sampling frame with a complete listing of all sampling units or an acceptable equivalent is needed. Such an equivalent could be households. In the case of community surveys of children, for example, it is important that the sampling frame be as complete as possible, in order not to select a biased sample.

When sampling from a population, a complete census of all households or city blocks (with a similar number of households per block) could be used as a sampling frame to select each sampling unit. In many countries, detailed maps are available for small population areas used as census sectors when conducting national censuses. Although they often require some updating (2, 8), these census maps are very useful as a sampling frame. They may not be available for many areas of developing countries, however. When a census can be conducted in a particular area, it will provide the ideal sampling frame. Alternatively, a listing of communities with their estimated population size could be used to select those communities from which one or more sampling units will be required.

Independently of the sampling frame used, it is important to have an estimate of the number of sampling units available in each lot to be sampled, in order to select an appropriate sample size and to obtain weighted estimates of performance level for several lots combined.

When sampling health records, a sampling frame could be selected within the health centre at the time of sampling, eliminating the need for a pre-existing sampling frame.

Systematic random sampling

LOAS requires that each individual sampling unit be selected following a random procedure. When conducting community surveys, it is very useful to have the sampling frame constructed with some geographical order. Then, when sampling, a systematic random sampling could be used to assure that the random selection of sampling units will cover all areas within the lot. If the sampling frame is a list of urban blocks, a number of blocks is then selected. In the field, once the selected block has been identified, the survey team needs to select, following a random procedure, an initial household to look for the sampling unit used in the survey. A complete enumeration of households in the block can be done and one household selected using a table of random numbers. Once the initial household has been identified, the surveyor contacts the family members to determine whether the sampling unit used in the survey is available in the selected household. If not, the next household to its right (or left, following a consistent pattern) is selected, until a sampling unit is identified and classified.

This procedure is not applicable when the only sampling frame available is a list of small communities, which is often the case in poor rural areas. Once the communities have been selected by simple random sampling proportional to the total population, the survey team needs again to select the initial household in which to look for the sampling unit. When conducting cluster surveys in rural communities of Peru (C. Lanata et al., unpublished), it was found practical either to do a rapid enumeration of households in the community with the help of a local leader, or to divide the community into two equal parts, and then to select one randomly. That selected sector was then divided into two similar sectors and one selected randomly. This procedure was repeated until the initial household was identified and the search for a sampling unit initiated. Much simpler approaches could be used when a complete census exists. In this case, the exact unit to be sampled is identified from the sampling frame.

Sampling units

Depending on the purpose of the survey, a sampling unit could be a child 12-23 months for immunization coverage, a child <5 for oral rehydration usage in
cases of diarrhoea, a pregnant woman for use of prenatal services, a woman aged 12-49 for use of family planning methods, etc. The sampling unit has to be the one which will provide the most useful information to evaluate a particular health programme.

The sampling unit can be changed once an initial quality level has been achieved by the programme, so that the programme can focus on another quality level. For instance, if the initial target of an immunization programme was to achieve at least 80% coverage of children aged 12-23 months, a subsequent target could be selected once the initial one has been achieved, for example to provide timely immunization to children <1. This age group could then be sampled to evaluate the percentage of children who have received the appropriate vaccines at the correct age, a target more difficult to achieve.

Different sampling units can also be combined into a single multipurpose survey. In this case, the quality levels to be evaluated, the sample size and the sampling strategy need to be defined independently for each particular sampling unit. This multipurpose use of LQAS was evaluated in a pilot study conducted in Lima (Peru) (8). This evaluation related to the performance of the diarrhoeal disease control programme, the prenatal care programme and the immunization programme. A single sample size of 9 was established with different numbers of permissible defects per sample (2 or 3), according to the indicator used. Within a week, a team of four surveyors evaluated 12 lots in a peri-urban community of Lima with a total population of about 86 000. 5 lots were rejected for their level of ORS use but none for ORS knowledge; 1 lot was rejected for its lack of use of prenatal care programmes; and 4 lots were rejected for their level of immunization coverage and the absence of immunization cards (8). This experience demonstrated the potential value of using LQAS for evaluating multiple health programmes in a single survey.

In the Costa Rica study, several primary health care programme activities were also evaluated at the same time (7). The Household Register Form, a form that stays in the household and is signed and dated by the health worker at every household visit, was used as a sampling unit to evaluate the health workers' programme. Children aged 1.5-35 months were used as a sampling unit to evaluate the immunization programme. Pregnant women were selected to evaluate the prenatal care programme, and infants <60 days were selected to evaluate the newborn referral programme to a physician. Children <6 years were selected as a sampling unit to evaluate the diarrhoeal disease control programme. All these survey experiences illustrate how sampling units are selected depending on the type of health programme to be evaluated.

For practical purposes, alternative sampling units which have a close relationship to the sampling unit to be evaluated are used. For example, when sampling at the community level to evaluate immunization of diarrhoeal disease control programme, one could use the household as an alternative sampling unit, when at the time of the analysis children <5 will be used as sampling units. The use of households is needed to build the sampling frame and for the selection of the households to be included in the survey. Other equivalent sampling units could be selected depending on the health programme to be evaluated.

Selecting the sample size

To select a sample size the method requires that an upper and a lower level of performance be determined for each health programme to be evaluated. The sample size will then be selected to ensure that lots with a real performance level above the upper level of performance will have a good probability of being classified as acceptable, and lots with a real performance level below the lower level of performance will have good probability of being classified as unacceptable. The wider the gap between the upper and lower performance levels, the smaller the sample size.

An immediate concern when carrying out the first LQAS is to estimate the real performance level of the lots to be sampled. Available information can sometimes provide this estimate, making it possible to select the upper or lower performance levels based on that estimate. Frequently, however, these estimates are very different from the real performance level, and the first round of samples may not provide useful information. In our pilot trial in Lima (8), we used estimated performance levels for each programme indicator based on the information provided by the local health officers. For ORS knowledge, for example, the upper level of performance was set at 50% and the lower level at 10%. None of the 12 lots sampled were rejected and the weighted performance level of all lots combined for this indicator was 78% (95% confidence interval of 70-86%). To solve this problem, in our next evaluation of LQAS in a mountainous region of Peru (5), we conducted a baseline survey in the lots selected to identify the performance level, so that in the next rounds, samples using LQAS could be selected more appropriately.

LQAS should be used in health surveys in order to identify those lots that have the lower performance level. Therefore, the selection of the sample size should be done with the intention of separating lots into two groups: acceptable and nonacceptable. The selection of upper and lower performance levels determines the sample size required. There is an “art” to selecting those upper and lower performance levels: they should make sense to health planners, be reasonable from an epidemiological point of view, result in small sample sizes, and produce a reasonable number of lots accepted and rejected. This process is facilitated if there is great variation in the performance level of the health programme among the lots to be sampled. On the contrary, if all lots to be sampled have very similar performance levels for the parameter to be evaluated, LQAS will not be of much help because it will require a very large sample size to identify a narrow variation between the upper and lower performance levels, otherwise all lots will be either accepted or rejected, depending on their performance level, and the information will be of little value. In those cases where the performance level for the indicator selected to evaluate a health programme has little variation, another aspect of the health programme (indicator) with more variation in performance level could be investigated.

The process of selecting a sample size when monitoring a health programme is therefore an active one, in which careful consideration should be given to the indicator, the sampling units and the upper and lower thresholds, in order to make it possible to separate lots into two parts with acceptable misclassification risks, and to obtain small, thus practical and feasible, sample sizes. This active process represents the most
difficult part of the LOAS methodology, because of its complexity, limiting its application by health planners. The process should be simplified in order to encourage individuals in charge of health programmes to use this methodology.

To select the exact sample size and the maximum number of permissible "defects" per lot, there are two options: (i) the use of a series of tables, where different upper and lower performance levels are shown for given type I and type II errors (4, 6, 9, 10); and (ii) the use of a series of curves (called operative characteristic (OC) curves), where the probability of acceptance of the lot is plotted according to the prevalence of defects in the lot (10). The tables are easier to use and understand, but the series of tables currently published have not taken into consideration the finite population correction for lots that have a relatively small number of sampling units. Moreover, the tables do not allow minor variations of type I and II errors, variations that can result in a mild but important reduction of the sample size per lot, a factor that is crucial when sampling a large number of lots in a community survey, some of which may be located in distant geographical areas. The published series of tables take these considerations into account, which are difficult to understand. A spreadsheet microcomputer program, like the one available from Lemeshow & Taber, will enable the user to produce tables for any levels of type I and type II errors. This may help overcome some limitations inherent in any set of published tables. However, using tables or OC curves for selecting sample sizes is awkward at best. What is needed is a friendly interactive microcomputer-based "expert" system which would request the desired upper and lower performance levels, the lot size, and then draw OC curves or prepare tables showing sample sizes and number of permissible defects for different sampling strategies, giving alternative suggestions to the health officer using the system, so that the best sample size could be selected. This potential use of microcomputers would facilitate the application of this powerful method by health personnel with a minimum of training.

Another aspect of the selection of a sample size is what to call a defect. If the immunization programme is used as an example, a typical defect would be a child not vaccinated. However, in the case of an immunization programme, to call a defect a "defect" a child who is vaccinated, but has not received the desired number of doses is the basis to the two critical values: \(d_1\) and \(d_2\). For the first sample, the lot is accepted if \(d_1\) or fewer defects are found. The lot is rejected if more than \(d_2\) defects are found. If the number of defects found in the first stage is greater than \(d_1\) but less than or equal to \(d_2\), a second sample is chosen. If the second sample is completed without exceeding \(d_2\) defects (from the first and second stages of sampling combined), the lot is classified as unacceptable. This two-step sampling scheme has the advantage that the second sample is only conducted in those lots that require an additional sample to classify them as acceptable or unacceptable, limiting the sample to those lots that could be classified as the first sample, therefore utilizing a minimum sample.

When evaluating a mass immunization campaign in urban and rural communities of a mountainous region in Peru (5), we decided to assess the usefulness of a two-stage sampling scheme. The sample for the first stage was 7, and required 0 defects to classify the lot as acceptable, and 3 or more defects to classify the lot as unacceptable. If either 1 or 2 defects were found in the first stage, a second sample of 8 was required. The surveyors found it very difficult to conduct this sampling scheme, especially in rural areas. The three investigators had to get together at the end of the first sample to determine whether a second sample was required. They was needed, they had to return to hard-of-access areas where they were before for the first sampling. Finally, one surveyor did not know if the maximum number of defects was already passed since the other surveyors were in other areas, with the result that most second samples were not stopped when the lot was already classified as unacceptable. This experience has demonstrated that the two-stage sampling scheme may be impractical in community surveys when several surveyors are sampling distant areas.

Another sampling scheme used in industry is sequential sampling (1). It is based on the same principle of double sampling, but extending it to triple and multiple sampling, providing continuous assessment of the need for further information before making the final decision of accepting or rejecting the lot. The method provides two upward-sloping parallel lines in a graph having the number of defects on the y-axis and the number of samples taken on the x-axis. The results of each sample are plotted and the sample is stopped when either line is crossed or the total sample size is completed. This powerful sampling scheme has not yet been reported as having been used in health surveys, although a pilot trial yielded very positive results, when a team was sampling mosquitoes in a malaria-endemic area, looking for drug-resistant malaria strains (G. Stroh, personal communication). Since data on the prevalence of resistant strains in the area, and the cost of the laboratory procedures was high, a sequential sampling scheme was chosen to minimize the laboratory testing. Sampling was stopped in all lots tested because the upper line was crossed quite soon, classifying the areas as drug resistant without the need to exhaust the complete sample.

Many other industrial sampling schemes that have a high potential for use in health surveys have not yet been tested, but it is conceivable that with more experience with LOAS, other industrial sampling
schemes (such as reduced and tightened inspections) will be tested in field conditions, and their role in health surveys identified.

Survey team

It is important to single out the need to select the survey team with great care. Appropriate personnel is important for any health survey, but more particularly in the case of community-based LQAS. We have found it convenient to have the supervisor of the individual who will use the information provided by the survey select the sampling units from the sampling frame. However in the field there will be no supervision of the surveyor to ensure that field procedures are correctly carried out. Although the small sample size utilized in LOAS may seem attractive, the need to select each individual unit randomly makes field activities difficult, frequently requiring travel over long distances just to identify one child. It is therefore very tempting for surveyors to use any child within easier access instead of going to the households or areas selected when sampling. The dedicated personality of surveyors is a key factor for the successful implementation of this methodology. It is important, as in other cases, that the supervisor should revisit some of the sampling units to make sure that procedures were carried out properly. Although we did not experience this problem when conducting the evaluation of LOAS in rural areas of Peru (5), we saw the potential for this occurring in cases where close supervision is not available.

Interpreting results

This is the critical step in LQAS. The lots with a lower performance level should be investigated and the reasons for this ascertained. This should lead to remedial measures designed to improve the programme’s performance. These measures should be evaluated with repeat LOAS surveys on the same lots, until the results justify directing supervisory activities to other lots. The principle is that improving the lots with the lower performance levels should make it more likely that a particular health programme will reach its stated goals faster.

The results of the first round of LOAS surveys also give an idea of the variation that could exist between lots, which is useful when selecting the sample size for the next cycle of surveys. If several lots are sampled at the same time, the performance level of the health programme evaluated could be estimated with a precise confidence interval by combining the results of individual lots in a weighted estimate for the overall region (as stratified random sampling). When analysing a double-sampling survey, only the first sample (which is the only one that is always completely used when estimating for several lots are combined). This mean estimate can then be used to guide the selection of the upper and lower levels of performance for the next cycle of surveys, and a sample size selected. As programmes progress, the gap between the upper and lower performance levels narrows and the sample size needs to be increased accordingly, until it is no longer practical to use that programme’s performance indicator. A different sampling unit should then be selected to evaluate another aspect of the programme, one which is more difficult to achieve.

To test the feasibility of applying LOAS in rural and urban areas in Peru, we decided to evaluate the implementation of three nationwide mass-immunization campaigns organized by the Peruvian government (5, 8) in a particular region. The results of the LOAS surveys were reported to health officials in charge of the immunization campaign at the local level, who after an initial hesitation took the results of our surveys seriously and carried out a series of changes, like changing personnel, redirecting volunteer workers to rural areas, providing travel facilities to health teams to enable them to reach distant areas before the local population left their homes to go to work, etc. Because of the specific changes introduced, we felt that LOAS had contributed to an improvement of the overall immunization coverage from 72% on the baseline survey, to 88% after the third campaign, although it could not be proved in the absence of a control area. It was interesting to see that the results of the LOAS surveys not only helped to correct programme strategies, but also provided feedback to the immunization teams who felt rewarded when they saw that their efforts were given recognition.

Although we used Ministry of Health personnel in our study with the idea that at the end of our project the LQAS methodology could be maintained in the area, this did not happen. The health officials found the methodology too confusing, especially when planning the next round of surveys and selecting an appropriate sample size. The use of OC curves was not clear to them. The public health nurses who were trained in the LQAS methodology resumed their previous tasks.

Conclusions

LOAS has great potential for monitoring health programmes, although its use has not extended further than in experimental evaluations of the methodology. There are several reasons for this. On one hand, because the methodology was developed in industry, it has not been disseminated much outside industrial circles, maybe in an effort not to facilitate the work of competitors. This may explain in part the fact that although the methodology has been available for more than 60 years, the first prevalence evaluations of LQAS in health surveys have only been carried out over the last decade. Despite successful trials of LOAS in health surveys, its routine use has not been established yet. Health planners might be reluctant to use this method frequently because of the difficulties inherent in its application. There is a need to simplify the methodology, perhaps with the help of microcomputers, in order to make it more accessible to health planners. Other limitations include the need to continuously update the sampling frame, which may require frequent censuses. In passing it can be mentioned that it is not rare for local health personnel to take censuses when planning their health actions; very often the information estimator for several lots are combined. This mean estimate can then be used to guide the selection of the upper and lower levels of performance for the next cycle of surveys, and a sample size selected. As programmes progress, the gap between the upper and lower performance levels narrows and the sample size needs to be increased accordingly, until it is no longer practical to use that programme’s performance indicator. A different sampling unit should then be selected to evaluate another aspect of the programme, one which is more difficult to achieve.

To test the feasibility of applying LOAS in rural and urban areas in Peru, we decided to evaluate the im-
are visiting an area in conjunction with other supervisory activities could take advantage of the visit to carry out an LOQAS survey. This has not yet been field-tested.

Despite their limitations, LOQAS and other industrial sampling schemes present several advantages for application in health surveys. The small sample sizes, the fact that they can provide information on small areas (lots) as well as precise estimates on large areas when several lots are combined, the feasibility of repeating surveys as frequently as needed in the population areas requiring close monitoring, the possibility of using impact indicators instead of process indicators to monitor a health programme—these are some of the advantages that make industrial sampling schemes very attractive for use in health surveys. LOQAS is an ideal technique for surveying areas in which considerable variation exists among the lots to be surveyed. More experience is needed to apply LOQAS in large geographical areas, either alone or combined with conventional survey techniques.

It is hoped that in the near future more attention will be given to industrial sampling plans in general and to LOQAS in particular. Its potential to monitor health programmes and provide useful information to re-allocate limited resources warrants further efforts towards its dissemination among health professionals. We are convinced that when the current constraints mentioned in this article are resolved, health planners will find this methodology quite useful as a powerful tool to help them achieve health programme targets in developing countries.

Acknowledgements

The authors would like to express their appreciation to George Stroh for his guidance in understanding and applying LOQAS in Peru.

SUMMARY

Traditional survey methods, which are generally costly and time-consuming, usually provide information at the regional or national level only. The utilization of lot quality assurance sampling (LOQAS) methodology, developed in industry for quality control, makes it possible to use small sample sizes when conducting surveys in small geographical or population-based areas (lots).

This article describes the practical use of LOQAS for conducting health surveys to monitor health programmes in developing countries. Following a brief description of the method, the article explains how to build a sample frame and conduct the sampling to apply LOQAS under field conditions. A detailed description of the procedure for selecting a sampling unit to monitor the health programme and a sample size is given. The sampling schemes utilizing LOQAS applicable to health surveys, such as simple- and double-sampling schemes, are discussed. The interpretation of the survey results and the planning of subsequent rounds of LOQAS surveys are also discussed.

When describing the applicability of LOQAS in health surveys in developing countries, the article considers current limitations for its use by health planners in charge of health programmes, and suggests ways to overcome these limitations through future research. It is hoped that with increasing attention being given to industrial sampling plans in general, and LOQAS in particular, their utilization to monitor health programmes will provide health planners in developing countries with powerful techniques to help them achieve their health programme targets.

RÉSUMÉ

Techniques d’échantillonnage par lots pour l’assurance de la qualité dans les enquêtes sanitaires dans les pays en développement: avantages et obstacles actuels

Les méthodes d’enquête traditionnelles sont généralement coûteuses et longues et ne fournissent des informations qu’au niveau régional ou national. L’utilisation des méthodes de l’échantillonnage par lots pour l’assurance de la qualité, qui ont été mises au point dans l’industrie pour le contrôle de la qualité, permet d’employer des échantillons de taille réduite lors des enquêtes menées dans des petits secteurs géographiques ou démographiques (lots).

L’article décrit l’utilisation pratique de ces méthodes dans les enquêtes sanitaires effectuées pour surveiller les programmes de santé dans les pays en développement. Après une description succincte de la méthode, l’article indique comment élaborer un plan de sondage et effectuer l’échantillonnage en vue d’appliquer cette méthode sur le terrain. La façon de choisir une unité de sondage pour surveiller le programme de santé et de choisir une taille d’échantillon est décrite en détail. Les systèmes de sondage utilisant cette méthode et qui sont applicables aux enquêtes sanitaires, comme les systèmes à un et à deux degrés, sont passés en revue. Enfin, l’article indique comment interpréter les résultats des enquêtes et préparer la série suivante.

La description de l’applicabilité des techniques d’échantillonnage par lots pour l’assurance de la qualité dans les enquêtes sanitaires dans les pays en développement indique les obstacles qui en limitent actuellement l’emploi par les planificateurs sanitaires chargés des programmes de santé. Une dernière section contient des suggestions sur la manière de surmonter ces obstacles grâce aux recherches entreprises à l’avenir. On peut espérer qu’en accordant plus d’attention aux plans de sondage industriels en général et aux techniques d’échantillonnage par lots pour l’assurance de la qualité en particulier, les planificateurs sanitaires disposeront, pour la surveillance des programmes de santé, de techniques très performantes qui leur permettront d’atteindre les buts de ces programmes dans les pays en développement.
REFERENCES — RÉFÉRENCES


THE POTENTIAL OF THE CASE-CONTROL METHOD FOR RAPID EPIDEMIOLOGICAL ASSESSMENT

Jane C. Baltazar *

Case-control design has gained popularity and importance over the past few decades, primarily for studying the causality of noninfectious diseases. The emergence of chronic disease problems in industrialized countries has stimulated the development of the modern case-control study.

At first, case-control studies were considered as preliminary and rather unreliable exercises with the burden of proof resting firmly on subsequent cohort studies (1). However, from the 1950s onwards, great advances were made in clarifying the methodological problems that used to plague such studies, and solutions to these problems were identified. Likewise, analytic techniques had been developed. At the same time, experience with the use of the design, mostly for risk-factor studies of chronic diseases, had accumulated, making it possible to test its reliability in practice. Thus, after a long time, the case-control study was accepted as a valid design.

With this development and with the recognition of its advantages (particularly in terms of duration of the study and resources required), the case-control method found a new application in policy-related epidemiological research. Examples include the evaluation of the effectiveness of health interventions such as vaccine (2), the Pap test in preventing invasive cervical cancer through early case detection (3), aspirin in reducing the risk of myocardial infarction (4), etc. Similarly, the method was used to determine the health impact of family planning practice. *

More recently, Briscoe et al. proposed its use in health-impact evaluation wherein the effect of improvement in water supply and sanitation on diarrhoeal disease (as the health impact measure) can be assessed (5, 6). Based on extensive theoretical considerations of major design issues, the potential of the method as a rapid, relatively inexpensive yet valid tool for this type of investigation, was recognized. However, because in 1975 a World Bank expert panel (6) reported that only very expensive, time-consuming prospective studies would be scientifically valid for assessing the health impact of water supply and sanitation projects, this new application of the method is still in the experimental stage, and needs to be substantiated by field studies.

This article focuses on key methodological considerations of the design when applied to the assessment of the health impact of water and sanitation interventions, using as an illustrative example a clinic-based case-control study in Cebu, a peri-urban area in the Philippines (7). In general, the study followed the procedures outlined in the earlier theoretical study: The strengths, limitations and potential usefulness of the method for rapid epidemiological assessment are briefly discussed.

Methodological issues

The approach

A case-control study starts with the identification of persons with the disease (or other outcome variable) of interest and a suitable control group of persons without the disease. The two groups are compared as to characteristics (exposure factors) which might predict or cause the disease.

In the Cebu study, the case group included children <2 who were brought to the clinic because of diarrhoea, and the control group, children of the same age who were brought to the clinic because of an acute respiratory infection, and who did not have diarrhoea in the preceding 24 hours.

The effect of the exposure factor (in this case, improved water supply and sanitation facilities) on the disease (diarrhoeal disease seen in clinics) is measured by the odds ratio (OR). After accounting for the effect of confounding variables, the OR is calculated as the ratio of the odds of exposure among cases, divided by the odds of exposure among controls. The OR is a direct estimate of the incidence density ratio or the relative rate (i.e. the average rate of diarrhoea in those with adequate facilities relative to those without adequate facilities) when the incident cases and controls are selected concurrently, with the controls being selected from those still at risk at the time of the onset of the case (8-11). The rare disease assumption is not necessary.

Study site

In the choice of study site, an important consideration is a fairly even distribution of the population according to the categories of the exposure factor under investigation. If one level of exposure is too low, the required sample size will become very large. The study area in Cebu had a fairly balanced distribution of population between those who were considered to have "adequate" water supply and excreta-disposal facilities and practices, and those considered to have "inadequate" facilities and practices. For example, in this area half the population used public or private boreholes as their primary source of water; 30% used the municipal water system; 10% used water from open dug wells; and about 10% used water from unprotected springs. Adequate excreta-disposal facilities (flush toilets, 

---

* Professor of Epidemiology, College of Public Health, University of the Philippines Manila, Ermita, Manila, Philippines.
water-sealed latrines or pit latrines) were available to two-thirds of the population, while the remaining one-third had inadequate facilities and practices.

Another consideration is the availability of health clinics. Identification and recruitment of study subjects are more advantageous in health facilities than in the community, both in terms of implementation and accuracy of the outcome information (diagnosis of cases and controls).

Finally, cooperation by the staff of the health facilities and by the community (particularly if the exposure information is partly determined through on-the-spot observation of the home environment) is essential for a successful implementation of the study.

In the present example, cases and controls were recruited from 16 government clinics distributed in the study area. Data were collected both at the clinic (clinical, anthropometric and identification information) and at home (water sources, uses and quantities of water, excreta-disposal practices, and socioeconomic information). Samples of drinking-water were collected from both the source and the home and examined for faecal coliforms.

Study period

Seasonal variation in incidence as well as variation in the predominating etiological agents of diarrhoeal disease are important considerations when determining the study period. Bacterial rather than viral pathogens are more susceptible to improvements in water supply and sanitation (12) and in most developing countries, bacterial pathogens have been shown to predominate in diarrhoea cases during the warm and rainy season. It would be most efficient therefore to conduct a case-control study of the effect of water and sanitation on diarrhoea in the season of peak bacterial diarrhoea. The Cebu study was conducted between July and October, the peak period for bacterial diarrhoea in the area.

Sample size

The method for determining sample size for the purpose of testing a specific hypothesis using the case-control method is described by Schlesselman (13) and by Lemeshow et al. (14). Readers are referred to these for details of sample-size estimation.

The size of the Cebu study was based on a planned analysis that would examine two levels of the exposure factor, and with the following assumptions: that the frequency of exposure to adequate water supply and sanitation facilities in the population is between 40% and 60%, and that there is a 90% chance of detecting a 33% reduction in diarrhoeal disease (the amount that is of public health importance) at the 5% significance level. The desired sample size included 460 cases and 460 controls; the number of subjects recruited, however, was lower (281 cases and 384 controls). The inability to recruit the desired number of study subjects was explained by the low level of service provided by the clinics during the period of the study, and a reduction in diarrhoea incidence during the study year as compared to previous years.

Potential sources of bias

The estimate of the computed odds ratio may deviate systematically from the true figure because of selection bias, misclassification bias and bias due to confounding.

Selection bias. Selection bias is the "truly large problem of the case-control study" (15). The manner of selecting study subjects may lead to a distortion in the estimate of the effect measure of interest. This could result from a violation of the key assumption in the case-control design, i.e. under the null hypothesis of no association between the exposure and the outcome, cases and controls would have an equal chance of being exposed to the factor under investigation.

To satisfy the above-mentioned assumption, several issues have to be addressed in the study design. The first issue pertains to the choice of the control disease. It is a recognized fact that house-to-clinic distance affects the utilization of clinic services, particularly in developing countries, and that the effect of distance on the reporting of the disease varies by severity of the condition. Hence, the control disease should have similar severity to the disease under investigation (diarrhoea) for both conditions to have similar propensity for reporting at the clinic, with similar house-to-clinic distance. This consideration is essential because the level of water and sanitation service differs between areas.

Acute respiratory infection is among the group of candidate control diseases enumerated in the theoretical paper, and was actually used in Cebu for the following reasons: (i) its severity is similar to that of the disease under investigation; (ii) it is not a water- or sanitation-related disease; and (iii) owing to practical considerations, i.e. the disease is a relatively common cause of medical consultations, hence the availability of a large pool of potential controls at the clinic where recruitment takes place.

In the Cebu study, the outcome that was measured was diarrhoea seen in the clinic. This implies that the condition was perceived as serious enough for the child to be brought to the clinic for medical attention. Cases were recruited from government clinics, thereby excluding from the case series those who utilized private clinics. These restrictions in the choice of cases were compensated however by restricting controls only to children with acute respiratory infection who had sought medical consultation from government clinics.

Thus, by carefully defining the eligibility criteria for controls, it was likely that cases and controls would be drawn from the same geographical and socioeconomic groups.

Another issue in the selection of cases and controls is whether individuals can be recruited into the study more than once. Because diarrhoea or acute respiratory infections occur in episodes, it is possible that a child may be seen in the clinic more than once during the time of the study. It is important, however, to make sure that the subsequent visits do not represent the same episode as the first one. Hence, a child who is first recruited as a case can subsequently be recruited either as a case again or as a control; and vice versa. In this situation, the relevant sampling procedure is the "incidence density" method in which a person can be sampled more than once during the time of the study. Con-
trols are selected from among those who are at risk of the disease under investigation at the time of the onset of the case. With such a short recruitment period no case nor control visited the clinic a second time in the Philippine study.

Misclassification. Misclassification of disease or exposure may arise from imperfect information on these variables. In a clinic-based case-control study, the classification of subjects into cases and controls is dependent on symptomatology as reported by the mother and confirmed by health personnel. Since the diagnosis of the disease is relatively easy, misclassification at the level of the outcome variable is quite unlikely.

Information on exposure status can be obtained by interview in the clinic or at home or by on-site observations. While it is more convenient to collect all data at the clinic, it has been reported that information on sanitary practices as collected by interviewing mothers does not correlate well with observed practices (16).

The Philippine study used a combination of interviews at the clinic and at home, and direct observation. Thus information on water quality (defined by source of water, method used to remove water from container and treatment method) was most likely to be of high quality except for water treatment which relied only on the reports of the mothers. With regard to excreta disposal, in addition to questions on latrine availability and utilization asked of the mother, observations were made by the field workers on the sanitation conditions around the house (based on the presence or absence of faeces, flies and odours). While an analysis of these data failed to reveal any individual interviewer bias, the validity of the data remains uncertain.

Confounding. Confounding refers to the effect of an extraneous variable (such as socioeconomic status) that wholly or partially accounts for the observed effect of the study exposure (such as water quality) on the disease (such as diarrhoea) under investigation.

In the Cebu study detailed information was collected on all potential confounders and taken into account in the analysis of their possible effects. Potential confounders not only included risk factors of the disease but also selection confounders, namely house-to-clinic distance and clinic of recruitment. As to the clinic variable, since controls were recruited from the same clinic as the cases, in effect there was matching on this variable. Because the clinic of recruitment is related to the exposure variable, clinic became a selection confounder which had to be controlled in the analysis. House-to-clinic distance had an influence on the propensity of the case or control to utilize the clinic services and could be associated with the quality of water supply and environmental facilities.

In controlling for potential confounders in the analysis, care should be taken to exclude intervening variables in the causal pathway under investigation. Two risk factors for diarrhoea — family history of diarrhoea in the past week and nutritional status — are influenced by water supply and sanitation and thus occupy an intervening position in the causal pathway; their control in the analysis, therefore, could lead to a biased odds ratio.

Comparison of findings from field studies

A case-control study in a rural African setting (Zomba in Malawi) (17, 18) used the same methodology as the Philippine study, except for some minor variations. The principal results of the two studies were very similar.

Another comparison can be made between the results derived from two different methodologies addressing the same research question. Cebu, aside from being the site of a case-control study, was also the site of a large prospective study which examined, inter alia, the effect of water supply, sanitation and feeding conditions on diarrhoeal diseases in young children. Details of the comparison of the two studies have been described elsewhere, but the main findings include the following: both studies show that improved water supply and excreta disposal are associated with substantial reductions in diarrhoea, with the effects statistically significant in the prospective study. The results differ in that the case-control study suggests that the joint effect of improved water supply and sanitation is about the same as the sum of the individual effects, whereas the prospective study suggests that the effect of the two improvements is synergistic. However, in neither study was the interaction statistically significant.

As explained in this article, for a variety of practical and analytical reasons there were a number of differences between the two studies which might be expected to affect the precision and produce bias in the estimates of the effect. Considering the cost, time requirement and results of the two studies, it is believed that the case-control method can provide policy-makers with broadly consistent and meaningful answers to specific questions about the effects of particular health interventions much faster and at a much lower cost than a prospective study. On the other hand, a prospective study is essential for exploring more complex and fundamental questions of child health.

Strengths and limitations

In general, there are a number of recognized advantages of the case-control study over other designs. With respect to its application to evaluation studies of the effect of improvements in water supply and sanitation facilities on diarrhoea incidence, the theoretical paper referred to earlier listed a number of additional advantages for the case-control design. The first one pertains to the timing of data collection. It would require the collection of only a single round of data in a relatively short period. Data could be collected once the improved facilities are functioning adequately and being used appropriately. The second advantage relates to the quality of collected data. When the study is based entirely on health facilities, aside from being relatively quick and easy to implement, it also provides information of a higher quality on disease status than with other designs. There is relatively less misclassification of disease status when cases are recruited from clinics than when the information is collected through recall in a home interview. Also, the more severe diarrhoea cases are the ones brought to the clinic, and thus are more likely to be included in the study.


d

Briscoe, J. et al. Quick and dirty or long and clean? Case-control and prospective studies of environmental sanitation and diarrhoea. (Submitted for publication).
the likelihood that they are caused by enteric pathogens is higher than cases in the community. As to the exposure information, it is likewise expected to be of high quality because, all things being equal, the smaller sample-size requirement would allow more attention to be paid to the quality of exposure data. Finally, the method does not give rise to the serious ethical problems often associated with quasi-experimental designs.

However, like other study designs, it has certain limitations. Firstly, the design will not allow for examination of the effect of exposure on more than one outcome. Secondly, it is not a suitable method when the exposure under investigation is rare. Finally, it is highly susceptible to bias. While this latter limitation is the most serious problem of case-control studies, at present there are recommended methods that minimize or adequately control the potential sources of bias both in the design and analysis stages of the study. In applying this design to health impact evaluation of water supply and sanitation programmes, procedures have been outlined which, based on extensive theoretical considerations, are expected to enhance the validity of the study. So far, experience from a few field studies has supported this expectation.

Potential uses in rapid assessment

The increasing interest in the prospects of the case-control study for rapid epidemiological assessment represents a new application of the approach, since it has traditionally been used mostly in risk-factor studies.

This new application has been examined in the context of assessing the effect of improvements in environmental sanitation and water supply on diarrhoea. From the few field studies cited, it could be discerned that in comparison with the prospective approach, the case-control design has the advantage in terms of time and cost without compromising the validity of the study. On the other hand, it was likewise noted that the successful application of the design was largely dependent on the availability of specialized manpower (advanced training in epidemiology and biostatistics), trained field staff, substantial logistical support and good computer capacity, all pointing to the fact that the case-control study is not a simple design. Consequently, the usefulness of the design in rapid assessment would be limited to places where the relevant specialized skills and the needed logistical support are available, usually in urban centres. Unfortunately, the peripheral areas, where there are generally only basic research capabilities, have a greater need for rapid assessment. The major challenge, therefore, is the development of a standard protocol that is valid, yet simple enough or whose degree of complexity is compatible with the research capabilities in the settings where the investigation will be undertaken. Again, in the area of environmental sanitation and diarrhoeal disease, the issues in considering a simpler protocol have been identified and relevant questions raised (18). Nonetheless, it is still essential that more field studies be conducted before conclusions can be drawn on the future of utilizing the case-control approach as a rapid, inexpensive and relatively simple yet valid tool for assessing the effects of improvements in environmental sanitation.

In addition, the design has been employed in studies of other policy-related questions. It would be worthwhile therefore to explore the usefulness of the design as a rapid assessment tool for health impact evaluation of different programmes in various settings.

In conclusion, while it is evident from the few available studies that the case-control design has potential in rapid assessment, there is still need for further work on the development of standard protocols that are simple enough to be carried out in settings where only basic research capabilities exist, but without compromising the validity of the study. Likewise more field studies are necessary to substantiate this new application of the design.

SUMMARY

Over the past few decades, the case-control method has been mostly applied to risk-factor studies of chronic diseases. Recently, among its new applications is the use of the method to study the health effect of improvements in sanitation and water supply. The methodological considerations, prospects and constraints of the method for rapid assessment are reviewed.

RÉSUMÉ

Le potentiel de la méthode cas-témoins pour une évaluation épidémiologique rapide

Au cours des dernières décennies, la méthode cas-témoins a été appliquée surtout aux études sur les facteurs de risque des maladies chroniques. Parmi les applications nouvelles figure depuis peu son emploi pour étudier les conséquences pour la santé d'une amélioration de l'assainissement et de l'approvisionnement en eau. L'article passe en revue les considérations d'ordre méthodologique, les perspectives et les contraintes de cette méthode pour une évaluation rapide.
REFERENCES — RÉFÉRENCES


During recent years, the importance of qualitative approaches in understanding social realities has been increasingly recognized by social scientists as well as by programme managers. Many researchers have started questioning the adequacy of an exclusively quantitative approach in explaining changes in the social and demographic situation. Among the various qualitative methods, "focus-group discussion" has become very popular and is being extensively used in social and behavioural research. While focus group is an established method in market research, its use in social science, demography or other related disciplines is rather new.

Following a brief description of the focus-group methodology, this article outlines the potential use of focus-group discussions, their strengths and weaknesses as well as methodological issues that still need to be investigated in order to make use of the full potential of this method. The conclusions call for more investigation into the factors that influence the outcome of a focus-group discussion, and warn against using focus groups as a stand-alone, rapid assessment method.

What is a focus group?

A focus-group session is an in-depth discussion in which a small number of people (usually 8-12) from the target population, under the guidance of a facilitator (moderator) discuss topics of importance for a particular study/project. It is basically a qualitative method in which the moderator, with the help of predetermined guidelines, stimulates free discussion among the participants on the subject of inquiry. The order in which the topics are covered is flexible, but generally the discussion starts with more general issues and slowly flows into more specific ones. At the end, a few probing questions are sometimes asked to reveal more in-depth information or to clarify earlier statements or responses.

Generally the participants are chosen purposively and it is recommended that they should be homogeneous with respect to characteristics which might otherwise impede the free flow of discussion. It is also considered desirable that the participants should not know each other or the subject of the discussion in advance.

The focus-group session should be held in a natural setting and be conducted in a relaxed manner. The full discussion is tape-recorded. Apart from the participants and moderator, a note taker also sits in the session but does not participate in the discussion. The note taker knows about the objectives and subject of inquiry, and is expected to be well trained in observing and noting nonverbal group feedback, such as facial expressions. Later the note taker also transcribes the complete discussion based on notes and tapes. These transcripts then serve as basic data for analysis.

It is expected that the informal homogeneous group setting, and the open-ended nature of questions, will encourage the participants to feel free from various constraints to which they are subject during individual interviews. Thus it is believed that they express their views openly and spontaneously. The moderator helps the participants to interact and this interaction stimulates memories and feelings and thus leads to a full in-depth discussion of the topic at hand. These group dynamics distinguish focus-group sessions from individual in-depth interviews typical of ethnographic research.

How can focus groups be used?

Available literature shows that the focus-group approach, like some other qualitative methods, could effectively be used as follows.

- As an idea-generation tool
  Focus-group discussions could, for example, be used by a health programme to find out what motivates people to use a specific health product or health-service facility, or to adopt better health-related practices. Such background information can be critical to health planners who need to know how the population views various health issues. In addition, focus-group discussions with health-care providers can be useful in pinpointing problems and in generating ideas for improvements in services.

- In conjunction with a quantitative study
  Focus-group discussions are often used as a complement to a quantitative study, helping to answer such questions as "why?" or "how?", rather than "how many?". They can also be used as a preliminary step, providing background information, and to generate hypotheses for field-testing. They can also be used to refine a questionnaire, and to ensure that the words and concepts correspond to those commonly used by the target group.

Focus groups have also been used as a follow-up to a quantitative study, to explain, expand and illum-

---

*A* Associate, Population Council, Baroda, India.
*§* Statistician, Division of Epidemiological Surveillance and Health Situation and Trend Assessment, World Health Organization, Geneva.
*®* Research Executive, Centre for Operations Research and Training, Baroda, India.
*©* Project Manager, Centre for Operations Research and Training, Baroda, India.
*©* Research Associate, Operations Research Group, Baroda, India.
*©* Knodel, J. & Pramualratana, A. Focus-group research as a means of demographic inquiry. Population Studies Center, University of Michigan, April 1987. (No. 87-106).
inlicate quantitative data, in order to gain some understanding about the reasons for certain findings.

In short, focus-group discussions, when used alongside quantitative studies, can result in a much greater understanding than either method used alone.

- **As a primary data-collection method**

Focus-group discussions can be used as a primary data-collection method for some topics that cannot easily be studied through quantitative methods. Focus-group discussions are particularly suited to subjects that are of a sensitive or personal nature; for example, Suyono et al. (1) covered abortion, and Kowaleski (2) covered sexual behaviour among gay men. Neither of the studies reported problems in discussing these rather sensitive topics. In fact, the former study in Indonesia found that participants were much more willing to discuss abortion in the focus-group discussion than they were in survey interviews. Group discussions suggested high awareness of abortion and different techniques for abortion, whereas sample-survey results indicated low awareness of abortion. The researchers concluded that survey interviews, which are usually watched by neighbours, are probably much less conducive to eliciting information about sensitive topics than are focus-group discussions which are away from home, among anonymous participants and in a supportive setting.

However, care must be taken to treat the results of focus-group interviews with some caution, since they can only suggest plausible answers, and cannot be indicative of the distribution of attitudes or beliefs in the population.

**Strengths and limitations of focus-group interviews**

The advantages of focus-group discussions have been discussed extensively in the literature (3-5, 10) These advantages could briefly be summarized as follows.

Focus-group discussions offer many of the advantages of qualitative studies without requiring full-scale anthropological investigations. They provide a wealth of insight into motivation, attitudes, feelings and behaviour that cannot easily be obtained by quantitative methods alone. This is probably the reason why the focus-group discussion approach has become so popular in recent times.

The group setting is believed to be beneficial in many situations. An informal, supportive group of people with similar backgrounds can often put people at ease, and encourage them to express their views freely and frankly. It enables participants to elaborate on ideas, and the group interaction can stimulate memories and feelings. Because each participant is relating to a group of people with similar backgrounds, the likelihood of participants giving answers they think will please the interviewer (a common problem of surveys) is reduced. In addition, because of the interaction during focus-group discussion, the moderator has more of a chance to clarify the questions, and there is less likelihood of questions being misunderstood. Finally, the relatively free format of the focus-group discussion allows the moderator to pursue unexpected avenues which are relevant to the topic at hand, but could not have been foreseen beforehand.

There are a number of limitations to focus-group discussions. Firstly, a group setting is not always ideal for encouraging free expression. Sometimes the group can inhibit discussion. For example, Vlassoff described a focus-group discussion amongst adolescent girls in India, during which the girls were painfully shy, not wishing to discuss their opinions in front of other people, despite extensive efforts to create a relaxed setting conducive to discussion.9

In addition, care must be taken in preparing transcripts from taped discussion. Chances of introducing error are particularly high if the interview has to be translated from the native language to the language of the investigator (a problem which is significant in multilingual environments).

Focus-group discussions also have many of the limitations of other qualitative methods. Their samples are small and purposively selected and therefore do not allow generalization to larger populations. In addition, as with other qualitative methods, the chances of introducing bias and subjectivity into the interpretation of the data are high. Because of this, it is not appropriate to treat the findings from focus-group discussions as though they were findings from quantitative research. While the focus-group discussion can provide plausible insights and explanations, one should not extrapolate from focus-group discussions to the distribution of responses in a population. This tenet is not always followed. In fact Merton, one of the founders of focus-group discussions, recently expressed his concern that "focus-group research is being mercilessly misused as quick-and-easy claims for the validity of the research not subjected to further, quantitative test" (6).

**Methodological issues**

Much has been written about the way to conduct focus-group interviews. However, this discussion has tended to be superficial, with little empirical backup and many basic questions remaining unanswered. This section discusses some key issues that still need to be addressed in order to further develop the focus-group method.

**The number of focus-group discussion sessions**

Little is known about how many discussion sessions are needed to be reasonably sure that all/most aspects related to the subject of inquiry have been explored.

Two aspects of this question need to be explored: the number of target groups needed and the number of discussion sessions with each target group. As discussed earlier, the population should be divided into homogeneous subgroups according to characteristics relevant to the research, such as users and non-users, males and females, working women and housewives, and geographical areas. This will enable the researcher to do a separate analysis for groups whose behaviour is different from one another. It will also help to create a supportive ambia-

---


Rapp. trimest. statist. sanit. mond., 44 (1991)
ance for the discussion, as the group members will have some characteristics in common. In addition, participants should have similar socioeconomic status, and possibly educational backgrounds as well, so that they all feel on an equal footing in the discussion. Other criteria for creating separate focus groups might be cultural or religious differences, gender or age, or any other characteristic which is likely to stand in the way of free discussion.

Little is known about the number of focus-group sessions needed for each subgroup. Debus et al. recommend doing at least two focus-group interviews with each subgroup, to compare the results. Textbooks in market research advocate forming additional groups until no new information comes to light. Even if only two interviews per target group are carried out, the number of interviews required may be large. For example, a study of attitudes towards contraceptives in India might easily require 24 interviews per geographical area since separate interviews would likely be required for males and females, for younger and older age groups, and for different caste groups (say three different castes). If more than one geographical area is included in the study, the number of sessions is multiplied accordingly.

For the time being, common sense and financial resources are the only guiding principles. However, this situation is far from ideal. What is needed is a methodological experiment as part of a larger study.

**Analysis of focus-group discussions**

Focus-group discussions provide a great deal of data, including interview notes and often transcripts of the session. This information needs to be analysed and organized in an understandable fashion. Content coding is often mentioned as the method of choice. This consists of listening or watching the tapes and reading the transcripts (if available) to generate a list of key ideas for each topic under discussion. Quotations and ideas are then placed under subcategories or combined into larger themes.

In order to see whether the transcripts are really necessary, and whether different researchers would come to the same conclusions from the same interviews, one of the authors (M. E. Khan) carried out a methodological experiment as part of a larger study. The larger study conducted in collaboration with the International Labour Organisation (ILO) Geneva, was to evaluate family-planning programmes at the workplace, using focus-group interviews as one of its tools. In this study all the sessions (7 in all) were both video- and audio-taped. Subsequently the videotapes along with written transcripts and objectives of the study were provided to seven different professionals, to go through the materials and analyse them to draw their own conclusions about the study.

Although the analysis is not yet final, initial observations suggest the following:

- Drawing conclusions from the videotape alone is difficult and time-consuming.
- Going through the transcripts is faster and easier.
- If conclusions are drawn only on the basis of seeing or listening to the tapes, there seems to be a certain amount of variation in drawing conclusions or picking up their expression.
- If transcripts are used for analysis, and content analysis is done properly (i.e. care is taken to note which views are expressed how many times and by how many participants), the answers or the conclusions drawn are fairly stable. Videotapes or observations of note takers add further stability to the interpretations.

**Practical applications in rural areas or urban slums**

It must be remembered that focus-group discussions were originally developed for market research in developed countries, where transport and communication are relatively advanced. Therefore, it is not surprising that some of the methods for conducting focus-group discussions will need modification in the face of realities in developing country environments.

This section is based upon recent experience in applying focus-group interviews in rural areas and in urban slums of India.

In many instances it was not practical to follow the usual guidelines, and a number of methodological issues and problems came to light.

**Should focus-group members know one another beforehand?** It is usually recommended that focus groups consist of individuals who are not acquainted with one another. It is believed that this increases the likelihood that group members express themselves frankly. However, this is not a practical option in many rural villages or urban slums where it is generally very difficult to find people who are not acquainted with one another. Based on experience in carrying out focus-group discussions in India, we feel that for topics which are not sensitive, the performance of informant does not make much difference, and the usual rule of anonymity can be relaxed. However, in the case of sensitive issues, participants who do not know one another provide better information than acquainted ones.

**Logistical problems in conducting focus-group interviews with women.** Focus-group discussions usually last for one or two hours. Group members are expected to concentrate on the topics being discussed. Experience suggests that this is difficult for women in the Indian context (and perhaps in other societies where free movement of women is socially restricted). Often, the women selected for the session feel it necessary to bring someone with them, especially the younger women who are frequently accompanied by their mother-in-law or younger sister-in-law. In addition, mothers are often required to leave the room to attend to some urgent work (e.g. to take care of crying children) and subsequently come back. This interrupts the discussion, and makes it harder for respondents to follow. Availability of space where a focus group could be privately conducted is a serious problem in some rural areas and urban slums (unless the respondents are ready to come to a community centre such as a school or Panchayat hall). If the sessions are conducted in a private home, getting enough space and privacy might be problematic.
Objections to tape recorder. Another problem encountered was that some participants objected to the use of tape recorders. In such instances it was not possible to record the session, and therefore it was necessary to rely only on notes. This can impede the analysis of the interview.

Homogeneity of the group. Although at present intragroup homogeneity is emphasized, our own experience reveals that in some cases heterogeneity may also be useful. For example, in a focus-group session consisting of women of lower-middle reproductive age, initially we found it very difficult to stimulate discussion on the problems related to reproductive health. But after a while an older woman (mother-in-law of one of the participants) present in the group started talking. This stimulated the younger women who then came forward with very useful information.

The role of the moderator. The moderator is crucial in focus-group research. It is the job of the moderator to keep the group focused on the topic at hand, to encourage group members to speak freely, to ensure that no group member dominates the conversation and that all opinions are heard, to create a supportive atmosphere, to probe when necessary, and to listen well. However, not much is known about the effect of the moderator's style on the results of interviews. For example, does an active moderator get more and better information than a quiet, laid-back moderator? Does a challenging argumentative moderator evoke more or better responses than a polite friendly moderator? More experimentation is required with moderator style, in order to be able to make informed choices on this important issue.

The ideal number of respondents within a focus group. Usually a focus group has anywhere between 8 and 12 respondents. However, lately there has been an increasing trend to use minigroups with 4-6 respondents. It would be useful to compare results from both types of group. Some work on this topic has been done by Fern (7) who observed that the number of ideas generated did not double as group size increased from 4 to 8, and that the ideas produced in a group were not necessarily superior in quality to those produced in individual interviews.

Conclusions

This article has shown that focus-group discussions have considerable potential to be used as a complementary approach to enrich social and behavioural research. However, its limitations need to be appreciated and its indiscriminate use should be discouraged. This article has also demonstrated that a number of methodological issues remain unanswered. There are hardly any methodological studies evaluating the trustworthiness and usefulness of the procedures. It is suggested that experimental studies should be undertaken to evaluate qualitative approaches, particularly how focus-groups fare against other qualitative methods, and how the findings of focus-group research are influenced by the various procedural differences raised above. It is also important to experiment in using the focus-group approach for rapid appraisal of health-promoting behaviours related to selected diseases. Unless attention is paid to strengthening the methodology by undertaking evaluative experimental studies, it is feared that the indiscriminate use of focus groups may cause more harm than benefit.

SUMMARY

The use of focus groups as a qualitative method for rapid assessment is discussed. A focus-group session is an in-depth discussion in which a small number of people (usually 8-12) from the target population discuss topics that are of importance for a particular study or project. Generally the participants are chosen purposively, and it is recommended that they should be homogeneous with respect to characteristics which might otherwise impede the free flow of discussion.

Focus groups can be used for idea generation, in conjunction with a quantitative method, or as a primary data-collection method. However, if focus groups are used as a primary data-collection method, their results must be treated with caution.

The main advantage of using focus-group discussions during rapid assessment is that they provide in-depth information without requiring full-scale anthropological investigations. The informal group setting is believed to make people feel at ease, encouraging them to express their views freely. However, there are a number of limitations to focus-group discussions. The samples are small and purposefully selected, and therefore do not allow generalization to larger populations. In addition, as with other qualitative methods, the chances of introducing bias and subjectivity into the interpretation of the data are high.

There are a number of methodological issues which still need to be addressed in order to further develop the method. Little is known about how many discussion sessions are needed to be reasonably sure that most aspects related to the subject of inquiry have been explored. The best way to analyse focus-group discussions, and the extent to which the results reflect the opinions and biases of the analyser, are not well understood.

Focus-group discussions have been used very successfully by market researchers in developed countries. Naturally, adapting this technique to rural and slum areas in developing countries will involve some changes in methodology. Methods of adapting focus-group discussions to rural and urban slum settings are still not fully worked out, and more exploratory work in this important area is required.

The conclusions indicate that focus-group discussions have considerable potential to enrich social and behavioural research, and suggest that more experimental methodological studies in using the focus-group approach for rapid assessment should be undertaken.

Rapp. trimest. statist. sanit. mond., 44 (1991)
L'utilisation des groupes focaux en recherche sociale et comportementale: problèmes méthodologiques

L'une des méthodes qualitatives d'évaluation rapide est celle des groupes focaux. Le groupe focal est composé d'un petit nombre de personnes (en général 8-12) provenant de la population cible, qui se réunissent pour discuter de sujets présentant de l'importance pour une étude ou un projet. En général, les participants ne sont pas choisis au hasard et il est recommandé qu'ils constituent un ensemble homogène pour éviter que certaines de leurs caractéristiques individuelles ne gênent le libre cours de la discussion.

Les groupes focaux peuvent servir à faire naître les idées ou être utilisés conjointement avec une méthode quantitative, ou encore être un moyen de collecter des données primaires. Toutefois, dans le cas de cette dernière utilisation, il conviendra de traiter les résultats obtenus avec prudence.

Le principal avantage de la méthode des groupes focaux pour l'évaluation rapide est qu'elle apporte une information en profondeur sans qu'il soit besoin de se livrer à des recherches anthropologiques en vraie grandeur. Par son côté informel, le groupe focal est censé mettre à l'aise ceux qui en font partie et les encourager à s'exprimer en toute liberté. La méthode comporte cependant un certain nombre de limitations. Le groupe focal représente un petit échantillon, dont la sélection est orientée et qui, par conséquent, ne permet pas la généralisation à des populations plus importantes. En outre, comme c'est le cas des autres méthodes qualitatives, le risque est considérable d'introduire un biais et de la subjectivité dans l'interprétation des données.

La méthode n'atteindra son plein développement que si l'on résout un certain nombre de questions méthodologiques. On ne sait guère combien de séances de discussion sont nécessaires pour être raisonnablement certain que la plupart des aspects du sujet traité ont été explorés. On ne voit pas non plus clairement quels sont les meilleurs moyens d'analyser les débats d'un groupe focal ni dans quelle mesure les résultats subissent l'influence des opinions et des préjugés de l'analyste.

La méthode des groupes focaux a été utilisée avec grand succès pour les études de marché dans les pays développés. Il va de soi que cette technique ne peut être adaptée aux zones rurales et aux quartiers urbains déséchus des pays en développement sans quelques changements de méthodologie. Les méthodes permettant d'assurer cette adaptation ne sont pas encore entièrement au point et il est nécessaire dans ce domaine important de procéder à des travaux exploratoires supplémentaires.

En conclusion, la méthode des groupes focaux peut apporter un enrichissement considérable à la recherche sociale et comportementale, mais il y aurait lieu de poursuivre les études méthodologiques expérimentales concernant son utilisation pour l'évaluation rapide.

REFERENCES — RÉFÉRENCES

The use of key informants to provide relevant infor-
mation on the health status of a population can
be a useful approach to the rapid assessment of
health problems in communities. Several studies on
community diagnosis (1, 2), community- and hos-
pital-based morbidity (3-6), malaria (7), schistoso-
miasis (8, 9), diarrhoea (10), coronary heart disease
(11) and child disabilities (12) have shown that dis-
eases and their main signs and symptoms were
often well recognized and perceived by community
members, and that this knowledge offers an import-
ant source of information.

Two major approaches can be considered for
assessing community health status on a large scale
(Fig. 1). The biomedical approach collects in-
formation on normative needs, either by the analysis
of routine health statistics, and/or by costly and
time-consuming epidemiological surveys. In con-
trast, the use of community-based questionnaires
allows the collection of information on demand pat-
terns, by making use of community illness and
disease conceptualization and prioritization. This is
an all-too-often neglected element, although match-
ing actions with the felt needs of the population
might be a key to success and sustainability in a
primary health care approach to improving health
(13). Both channels contribute to the health planning
process, which in turn should have a positive in-
fluence on the community health status (and modify
the perception patterns of the population).

Information from community informants is usually
collected by scientists or their field staff in a limited
number of communities, for an in-depth assessment
of socioeconomic, attitudinal and/or anthropological
aspects of health. A few studies have also shown the
value of direct interviews as a means of individual
diagnosis in a hospital or community setting (3-6,
11, 12). For community health assessment on a
larger scale, however, the distribution and com-
pletion of simple (self-administered) questionnaires
through an existing administrative system offers a
very cost-effective alternative, if operationally
feasible. This represents an "indirect" approach, in
contrast to the classical "direct" interview approach.

However, using the subjective perception of a
limited group of informants or even of a single
person in a community can lead to a number of
pitfalls. Respondents' motivation and personal ex-
erience can significantly affect their answers. The
risk of a selection bias is unavoidable, since no
informant is representative. Therefore, a proper val-
idation of the interview approach with standardized
epidemiological markers is needed to assess the
relevance and efficacy of the approach.

This article discusses the use that can be made of
key informants, with particular reference to two
studies carried out in the United Republic of
Tanzania (14, 15), which used school-based ques-
tionnaires to screen for communities at high risk for
urinary schistosomiasis. Urinary schistosomiasis
offers a good example of a disease with focal dis-
tribution, for which comprehensive epidemiological
information is required for planning control pro-
grammes (16).

Description of the Tanzanian studies

The two studies were undertaken in the Kilombero
and Kilosa districts of Morogoro Region, south-
eastern Tanzania. The Kilombero District has an area
of 15,000 km², with a population of 187,000 living in
49 villages. The Kilosa District has the same area,
with a population of 347,000. In both districts, the
population is mainly engaged in subsistence farm-
ing.

The questionnaires

The head-teacher questionnaire. The questionnaire
aimed at head teachers had three main questions:
(i) the ranking of the diseases most prevalent
among schoolchildren (to be chosen from a list
of 10 diseases);
(ii) the ranking of the signs and symptoms most
prevalent among schoolchildren (to be chosen
from a list of 11 symptoms);
(iii) the ranking of priority diseases for control (no
proposed list).

Two additional questions probed health among
other community problems, and the availability of
drinking-water.

The schoolchildren's questionnaire. A teacher asked
the children individually and separately if they had
experienced during the last month the listed
8 symptoms (coughing, itching, headache, fever,
abdominal pain, blood in urine, blood in stool, diar-
hoea = question 1) and 8 diseases (malaria, diar-
hoea, skin diseases, eye diseases, schistosomiasis,
respiratory infections, worms, abdominal problems =
question 2). The teachers wrote down the children's
answers as "yes", "no" or "don't know" (counted as
"no" in the evaluation). The answers of the whole
class (up to 60 children) were written down on a
single sheet of paper (recto-verso). The prevalence
rate of positive answers for "blood in urine" and
"schistosomiasis" was later computed by our team,

Rapp. trimest. statist. sanit. mond., 44 (1991)
There are two alternatives for collecting information on the health status of communities: biomedical measurements (assessing normative needs) vs. community perception/prioritization (assessing demand patterns). Health information collected by both channels contributes to planning and therefore to an improved community health status (up arrow).

* Il y a deux méthodes au choix pour recueillir des informations sur l'état de santé des communautés: déterminations biomédicales (évaluation des besoins normatifs) ou perception/priorités de la communauté (évaluation des schémas de la demande). Les informations sanitaires recueillies selon ces deux méthodes contribuent à la planification et, partant, à une meilleure appréciation de l'état de santé d'une communauté (flèche dirigée vers le haut).
per class and per school. Both questionnaires are found elsewhere.\textsuperscript{9}

Study 1 (Kilombero District)

Questionnaire distribution. After being designed and pretested, the head-teacher and student questionnaires were printed locally and distributed to all 77 primary schools by the district education office, along with routine administrative mail (Fig. 2A). The input of our team was limited to the design and production of the forms and to an extensive briefing of the District Education Officer. He was the only person who knew about the project's interest in schistosomiasis, since the exercise was presented to the teachers as community-problem-oriented, and to the children as overall health-oriented. Within 6 weeks, 77/77 (100%) head-teacher and 75/77 (97.4%) student questionnaires were returned to the district education office. 6,772 children were interviewed by their class teachers. A slightly modified questionnaire was also sent through the political party to all village chairmen.

Validation. After reception of all questionnaires, an extensive parasitological screening survey was conducted by a mobile field-laboratory team. Urine filtration according to a standard procedure was performed in a sample of 56 out of 77 primary schools. In total, 4,469 children were examined (mean = 85 ± 15 per school, range = 53–127). All egg-positive children were treated with a single dose of Praziquantel (40 mg/kg).

Study 2 (Kilosa District)

In the second study, the design was modified by handing over the questionnaire validation to the head teachers, instead of having a specialized team going around in all schools. In order to make this possible, another biomedical validation method was chosen: reagent sticks detecting blood in urine, which are fast, practical and efficient for diagnosing urinary schistosomiasis cases, especially in an epidemiological context (17). The role of our team was restricted to planning and crosschecking the quality of the head teachers' testing.

Questionnaire distribution. In a first step (Fig. 2B), school questionnaires similar to those used in the Kilombero District and aimed at head teachers and schoolchildren were sent to all 168 primary schools of the district. Within 4 weeks, 164 forms (97.6%) were returned to the district education office, all filled in properly, with 15,073 children interviewed. Based on the questionnaire data from the Kilombero District, schools were classified as positive (high risk) if at least 35% of the children declared they had had schistosomiasis during the last month.

Validation (teacher reagent-stick testing). For the second step of the study we selected all 49 high-risk ("positive") schools, as well as 26 randomly-selected low-risk ("negative") schools as controls. The head teachers of these 75 schools were asked by the district education office to come to a one-day seminar where they were given a short lecture on schistosomiasis, and then instructed in the use of reagent sticks for detecting microhaematuria. At the end of the seminar, they were given enough reagent sticks to test 80 children of their school. Six weeks later, 73 out of 75 schools (97.3%) had returned the result sheets to the district education office, and 5,740 children had been tested.

Crosschecking. For crosschecking the quality of the head teachers' testing, our team and a representative of the district medical office selected randomly 18 schools among the 73, and retested with reagent sticks all the children already screened by the teachers, who were in school on the day of our visit. We could confirm that the head teachers performed reagent-stick testing in a very satisfactory way, as there was no significant difference between their testing and ours. At the end of the study, all haematuria-positive children were treated with a single dose of Praziquantel, either by our team or by the district medical office staff.

Results

Only the main results, useful for illustrating the concept of key informants, are given in this article. Detailed results can be found elsewhere (14, 15).

Diagnostic performance of the questionnaires

The correlation between the biomedical prevalence rates and the questionnaire results were significant at p<0.0001 for all questions in both studies. Subsequently, the predictive values of the questions could be calculated using Baye's theorem (18). Especially question 2 of the children's questionnaire ("Did you have schistosomiasis during the last month?") showed an excellent diagnostic performance in both districts (sensitivity and specificity >90%). Of special interest were the high negative predictive values of the questions in both studies (>92%), since it meant that low-risk ("negative") schools could safely be identified and excluded from further testing. Resources could then be concentrated on the high-risk schools.

Priority thresholds

The priority rank of schistosomiasis as a disease to be controlled was found to be clearly related to its prevalence in the community. This showed that the disease was well-recognized by the respondents, and that a process of prioritization was taking place. Furthermore, there seemed to be a sort of "high-priority limit", a disease prevalence rate above which schistosomiasis was consistently a "top 5" disease, while below this limit it was often not even cited. This information might prove very relevant for setting an intervention threshold in view of a specific control programme.

Comparative costing of the approaches

A comparison of the financial screening costs showed that questionnaires aimed at key informants, and distributed through an existing administrative system, are inexpensive when compared to parasitological screening. In the Kilosa District, the cost per school of the questionnaire approach (US$6.3) was 24 times below that estimated for a specialized team performing urine filtration (US$153.8).

FIG. 2

DESIGN OF TWO STUDIES ASSESSING THE VALUE OF QUESTIONNAIRES FOR THE IDENTIFICATION OF COMMUNITIES AT A HIGH RISK FOR URINARY SCHISTOSOMIASIS, UNITED REPUBLIC OF TANZANIA

A. STUDY 1 – KILOMBERO DISTRICT

- 77 schools (class 1, 3 and 5)
- 51 party branches

56 schools
Mean: 85 children/school

- Map of S. haematobium distribution
- Positive children treated

STEP 1
Questionnaires
sent via District Education Office and party

STEP 2
Validation
Urine filtration by mobile laboratory

Outcome

B. STUDY 2 – KILOSA DISTRICT

168 Schools
All children of class 1, 3 and 5

164 schools responding (97.6%)

115 negative schools
49 positive schools

26 randomly selected schools
all schools

75 schools
80 children/school

Crosscheck in 18 randomly selected schools

- Map of S. haematobium distribution
- Positive children treated

STEP 1
School questionnaires
sent via District Education Office

Classification
Threshold = 35% of children replied "I had schistosomiasis"

STEP 2
Validation
Teacher reagent-stick testing

Crosschecking

Outcome
FIG. 2

PLAN DE DEUX ÉTUDES POUR ÉVALUER L'UTILITÉ DES QUESTIONNAIRES EN VUE D'IDENTIFIER LES COMMUNAUTÉS À HAUT RISQUE DE SCHISTOSOMIASIS URINAIRE, RÉPUBLIQUE-UNIE DE TANZANIE

A. ÉTUDE 1 – DISTRICT DE KILOMBERO

- 77 écoles (classes 1, 3 et 5)
- 51 sections du parti

<table>
<thead>
<tr>
<th>56 écoles</th>
</tr>
</thead>
<tbody>
<tr>
<td>Moyenne: 85 enfants par école</td>
</tr>
</tbody>
</table>

- Carte de la distribution de *S. haematobium*
- Enfants positifs soignés

ÉTAPE 1

*Questionnaires* expédiés par l'intermédiaire du bureau d'éducation du district et du parti

ÉTAPE 2

*Validation*

Filtration de l'urine par un laboratoire mobile

Résultat

B. ÉTUDE 2 – DISTRICT DE KILOSA

168 écoles

Tous les enfants des classes 1, 3 et 5

Réponses de 164 écoles (97,6%)

<table>
<thead>
<tr>
<th>115 écoles négatives</th>
</tr>
</thead>
<tbody>
<tr>
<td>49 écoles positives</td>
</tr>
</tbody>
</table>

26 écoles choisies de façon aléatoire

Toutes les écoles

<table>
<thead>
<tr>
<th>75 écoles</th>
</tr>
</thead>
<tbody>
<tr>
<td>80 enfants par école</td>
</tr>
</tbody>
</table>

Vérification dans 18 écoles choisies de façon aléatoire

- Carte de la distribution de *S. haematobium*
- Enfants positifs soignés

ÉTAPE 1

*Questionnaires pour les écoles* expédiés par l'intermédiaire du bureau d'éducation du district

Classification

Seuil = 35% de réponses des écoliers “J'ai eu la schistosomiase”

ÉTAPE 2

*Validation*

Epreuve à bandelette réactive pratiquée par l'instituteur

Vérification

Résultat

*WHO 91715*

*Rapp. trimest. statist. sanit. mond., 44 (1991)*
Methodological considerations for the “indirect” questionnaire approach

Having shown an application, we should review methodological considerations more systematically. The indirect key-informant approach can be divided into three main components: (i) designing and pretesting; (ii) distribution, completion and returning; (iii) validation (Fig. 3).

Questionnaire designing and pretesting

Since the questionnaires are to be distributed and administered by non-health professionals, in the absence of the investigators, a simple and concise design is essential. A covering letter or an explanatory text at the beginning of the forms should give all necessary instructions for dispatching and completion. Questions should be short and clearly recorded, using words that are known to the respondents. Both closed-ended and open-ended questions should be tried. Closed-ended questions (e.g. “Please rank among the following diseases the 6 most important ones”) aim to elicit answers that are structured and easy to analyse. The list of proposed answers to choose from should be restricted to a maximum of 10 items, and to terms/concepts that are familiar to the respondents. In the United Republic of Tanzania, the proposed lists of symptoms and diseases that the teachers and party chairmen were requested to choose from were established following extensive community health surveys in the area [19]. But this information could also be obtained from routine statistics. Open-ended questions (e.g. “What are in your regard the most important health problems”) should allow more detailed investigation of prioritization patterns and would provide additional qualitative information.

Time should be taken to carefully pretest the questionnaires, since no corrections can be made once the forms are distributed. The usual rules of pretesting in social science field work apply: initial comprehension-testing with a few selected people, then limited testing with respondents similar in socioeconomic and ethnic characteristics to the target population. In case of translation into another language, comprehension of the back translation should also be assessed. The investigator should carefully record any design or comprehension problems and correct the forms accordingly. After each correction, the new form should be retested.

Distribution

Distribution via a working administrative system simplifies the administration of the questionnaires and can greatly increase cost-efficiency. The two administrative systems that were chosen in the United Republic of Tanzania (schools and political party) proved to be very reliable and efficient for distributing and administering the questionnaires. Return rates were respectively 239 schools out of 245 within 6 weeks (97.6%) and 44 party branches out of 51 within 12 weeks (86.3%). But other systems, such as cooperatives, postal administration, traditional chiefs, churches and others, able (ii) to provide literate and articulate informants and (iii) to ensure good coverage, might also be considered for the purpose. Distribution of the questionnaires should go along with routine administrative mail, in order to minimize workload and transport costs.

Two different interview situations may be found:
1. Self-administered. The questionnaire is forwarded to the key informant, who fills in the form (e.g. teacher, politician).
2. Multiple interviews by a community member. The questionnaire is forwarded to a community member, who is requested to interview a special group of key informants (e.g. teacher interviewing schoolchildren, religious leaders interviewing community members). This situation allows the “interview” of a large number of individuals in each community, reducing thereby the problem of subjectivity associated with a single person’s answer. Both approaches worked well in schools in the Tanzanian studies, giving interesting and complementary information.

An important point to avoid a bias is that the respondents should not be aware of the problem or disease of interest to the investigator. Questionnaires should never be disease-specific, but always overall-health or even community-problem oriented. This is also desirable in view of replacing the disease of interest in the context of other health determinants, and to obtain therefore elements of prioritization. Experience with schistosomiasis has shown that if interviews are performed in the frame of disease-specific surveys, the answers are highly unreliable [20].

Validation

After receipt of all questionnaires, a biomedical screening survey should be conducted in an adequate sample of the target population, in order to provide the data base needed for evaluating the diagnostic performance of the questionnaire approach (for the disease/health problem of interest). For some studies, the required biomedical data might be readily available (existing statistics, death certificates, etc.).

Questionnaires provide mainly a qualitative type of information [14, 15]. The units to be screened (e.g. individuals, villages, traditional chiefdoms, schools), have therefore to be classified before the analysis into “positive” or “negative”, according to preset criteria. In schistosomiasis for example, high-risk (“positive”) communities can be defined as having a parasitological prevalence rate in schoolchildren of at least 50% [21]. In another study, self-administered questionnaires were tried regarding their ability to predict death from coronary heart disease [11]. Once this classification is set up, the ability of the questionnaires to identify correctly the “positive” and the “negative” units can be calculated and quantified in terms of sensitivity, specificity, and predictive values [18].

Evaluation is of course easier for chronic diseases (e.g. filariasis, skin diseases, cardiovascular diseases) than for acute diseases (e.g. fever/malaria, diarrhoea), which will require a retrospective or longitudinal validation methodology. Validation might be impossible if no practical and/or reliable field test is available for the disease of interest (e.g. meningitis) or when this disease is rare (e.g. yellow fever). But basically, the same restriction for these diseases applies to community-based biomedical surveys.

Financial costs such as transport, equipment, salaries, travel and other costs should always be systematically recorded during validation, so that a cost-efficiency analysis can be done.

Wld hlth statist. quart., 44 (1991)
FIG. 3
LES TROIS PHASES DU DÉPISTAGE «INDIRECT» PAR QUESTIONNAIRE, AVEC LES PRINCIPAUX PARAMÈTRES D'ÉVALUATION *

PHASES

1. Designing and pretesting

2. Distribution
   - Completion
   - Returning

3. Validation

EVALUATION

Design, comprehension

Feasibility
Return rate
Return time

Diagnostic performance
Cost-efficiency

* The circle indicates the phase executed by an existing administrative system, while squares indicate work done by the biomedical team.

FIG. 3
THE THREE PHASES OF "INDIRECT" QUESTIONNAIRE SCREENING, WITH THE MAIN EVALUATION PARAMETERS *

PHASES

1. Plan de l'enquête et épreuve préalable

2. Distribution
   - Achèvement
   - Renvoi

3. Validation

EVALUATION

Conception, compréhension

Faisabilité
Taux de renvoi
Délai de renvoi

Résultat du diagnostic
Rentabilité

* Le cercle indique la phase exécutée par un système administratif existant, tandis que les rectangles indiquent le travail effectué par l'équipe biomédicale.
<table>
<thead>
<tr>
<th>Mobile laboratory (biomedical surveys)</th>
<th>Key informants (questionnaires)</th>
</tr>
</thead>
</table>

### Prerequisites

- Requires a qualified team
- Requires a diagnostic field test

### Advantages

- Quantitative results
- Public relations
- Health education/information can be given within survey
- Case treatment

### Disadvantages

- High screening cost
- High resources input (staff, vehicle, equipment)
- Time-consuming
- Disease-focused

---

### Prerequisites

- Nécessite une équipe qualifiée
- Nécessite une épreuve diagnostique sur le terrain

### Conditions préalables

- Les résultats quantitatifs
- Relations publiques
- Une éducation/information sanitaire peut être dispensée pendant l'enquête
- Traitement des cas

### Avantages

- Faible coût du dépistage
- Rapidité
- Basée sur la communauté et orientée vers des problèmes précis
- Les ressources pour les interventions ultérieures sont centrées sur les communautés à haut risque

### Inconvénients

- Coût élevé du dépistage
- Apport important de ressources (personnel, véhicules, matériel)
- Durée prolongée
- Axé sur une maladie

---

*Wid hith statist. quart., 44 (1991)*
Extension and limitations of the key-informant approach

More experience is needed on the operational and epidemiological determinants that are crucial for the success of this approach to health assessment. This is the general objective of a standardized multicentre study for the identification of communities at a high risk for urinary schistosomiasis, initiated by WHO in 7 African countries (Cameroon, Congo, Ethiopia, Malawi, Zaire, Zambia, Zimbabwe), and modelled on the Kilosa study. Preliminary results from the first phase (questionnaire distribution) showed that this approach was successfully carried out in all countries, suggesting that the operational requirements for the distribution and collection of questionnaires were fulfilled in a number of African countries.

Besides the operational constraints and the problems related to validation, the major restriction for the extension of this approach to other health problems is that the disease/health problem of interest to the investigator must be selectively perceived by the population. This perception can relate to the general picture of the disease or health problem or it can relate to a single sign or symptom, e.g. blood in urine for urinary schistosomiasis (14, 15), clinical signs or symptoms indicative of morbidity (5, 6, 11, 12). More than sensitivity, the specificity of an indicator might be problematic to achieve, as illustrated by the complex problem of the etiologies of “fever” or “headache” in relation to malaria diagnosis.

Unfortunately, the confidentiality of the collected information cannot be fully guaranteed and this limits the use of the indirect questionnaire approach for certain diseases or health problems (AIDS, other sexually transmitted diseases, child mortality, etc.).

The most important points of the discussion are summarized in Table 1, where the mobile laboratory and the indirect key-informant approaches are compared for area-wide health assessment. However, these methodological considerations should not deter from the fact that the crucial issue in health status assessment remains to know who determines what information is required and who will make use of it.

Acknowledgements

The authors wish to thank the National Institute of Medical Research (NIMR), United Republic of Tanzania (Director-General: Professor W. L. Kilama) for permission to publish. This work was partly supported by the Swiss Directorate for Technical Cooperation and Humanitarian Aid (SDC). One of the authors was supported by the Rudolf Geigy Stiftung zu Gunsten des Schweizerischen Tropeninstituts.

SUMMARY

The requirement for a careful interview set-up in the classical key-informant approach considerably limits its extension to a larger scale. However, the administration of simplified questionnaires entirely through an existing administrative system can provide an alternative approach and become a valuable tool for fast and cost-efficient community health assessment. This methodology relies on a well-structured and working administrative system, and is restricted to health problems that are well and distinctly perceived by community members. It gives mainly qualitative results, and can therefore be used to single out rapidly and inexpensively high-risk and low-risk units. The biomedical screening of a large number of negative units can thus be avoided, and available resources can be concentrated on the positive ones. An extensive validation of this “indirect” approach by standardized biomedical measurements is a crucial requirement for its usefulness.

In addition to biomedical information, questionnaires also provide information on the perception and prioritization of community health problems, which is of interest for health planning.

RÉSUMÉ

L’utilité des questionnaires destinés aux répondants clés et distribués dans le cadre d’un système administratif existant pour une appréciation rapide et rentable de la situation sanitaire

Dans l’approche classique basée sur des répondants clés, la nécessité d’une entrevue détaillée limite considérablement l’élargissement de cette méthode sur une plus grande échelle. Cependant, l’emploi de questionnaires simplifiés, diffusés uniquement par le biais d’un système administratif existant, peut constituer une solution de rechange qui se révèlera très utile pour apprécier rapidement et de façon rentable l’état de santé d’une communauté. Cette méthodologie repose sur un système administratif bien structuré et en bon état de marche et elle se limite aux problèmes de santé que les membres de la communauté perçoivent aisément et clairement. Elle fournit principalement des résultats qualitatifs et peut donc être utilisée pour distinguer rapidement et à peu de frais les unités à haut risque et les unités à faible risque. On évitera ainsi de procéder au dépistage biomédical chez un grand nombre d’unités négatives, les ressources disponibles pouvant dès lors être concentrées sur les unités positives. L’une des conditions capitales régissant l’utilité de cette approche «indirecte» est sa validation poussée au moyen de déterminations biomédicales normalisées.

Outre les informations d’ordre biomédical, les questionnaires fournissent aussi des renseignements sur la perception des problèmes de santé par une communauté et la priorité qu’elle leur assigne, ce qui est intéressant pour la planification sanitaire.
REFERENCES


OMS Série de Rapports techniques No 728, 1985 (Lutte contre la schistosomiasi: rapport d’un comité OMS d’experts).
THE BENEFITS OF THE APPLICATION OF GEOGRAPHICAL INFORMATION SYSTEMS IN PUBLIC AND ENVIRONMENTAL HEALTH

Henk J. Scholten a & Marion J. C. de Lepper b

Measures to improve health and the quality of life are now receiving greater attention, together with the need to protect and improve the environment. Environmental health problems — which are diverse and complex — cross national boundaries and in many cases need to be dealt with internationally. It is therefore not surprising that large organizations try to pool their efforts in the context of environmental and health policies and research.

One of the most important questions arising in public and environmental health today concerns the type of instrument that can be used to devise quick, reliable and scientifically valid methods of rapid assessment, to assist in health research and in the planning, monitoring and evaluation of health programmes. New techniques must enable public authorities to gain insights into the consequences of decisions relating to investment in environmental management and public health, and make it possible to review the current situation in which new developments are proposed. Research institutes must have access to various data sources. In order to achieve these objectives, information is vital for effective guidance in this rapidly-changing context. All relevant information must be stored, managed, made available and presented in a suitable form for use at different stages in the research and planning process. The information may be of various types, extensive in quantity, variable in quality and referring to area units of different size.

This article describes how geographical information systems (GIS) can provide excellent frameworks within which these activities can be undertaken. GIS can be defined as automated information systems based on the methodology of geographical data-base management and query. It begins by clarifying the objectives of GIS and by outlining the basic types of data and data storage. Subsequently, hardware and software developments and the variety of application fields, organizations and users are discussed. By describing some case studies, emphasis will in addition be on the role GIS have played, and could have played, thus far in research on health issues. In the final section, a broader perspective is presented which outlines the application of GIS in health research and policies in the European context.

Health and ill-health are affected by a variety of lifestyle and environmental factors, including where people live. Characteristics of these locations (including sociodemographic and environmental exposure) offer a valuable source for epidemiological research studies on health and the environment. Health and ill-health always have a spatial dimension therefore. More than a century ago, epidemiologists and other medical scientists began to explore the potential of maps for understanding the spatial dynamics of disease. One of the most famous early users of maps in medical science was John Snow (1813-1858) — a London anaesthesist and Queen Victoria’s obstetrician. Snow had the idea that cholera — the classic epidemic disease of the XIXth century — might be spread by contaminated water supplies. By using maps showing the geographical distribution of cholera deaths in the Soho area of London in 1854, he demonstrated that the association between cholera deaths and contaminated water supplies resulted in a striking geographical distribution.

Nowadays, as AIDS is projected to become of critical importance in the remainder of this century, attempts to forecast and predict its developments are urgently needed. Although a vast amount of literature (describing many different aspects of the disease) has already been produced, the important geographical aspect seems to have been largely ignored. As Kabel (2) points out: “Rarely can one find an attempt to model the spread of AIDS incorporating the basic spatial dimensions of human existence. Most modelling seems to be focused completely within the temporal domain”. One of Kabel’s main lines of argument is that modelling the geographical distribution of AIDS can contribute to both educational intervention and the planning of health care delivery systems. Mapping can play an important role in both areas as it is an excellent means of communication. In order to be of use to resource planners, predictions of AIDS should include a spatial component.

The above examples describe the use of spatial information in health research. For most people, using spatial information is the same as producing maps, whereby the map is used to show the geographical distribution of some medical condition. Cliff & Haggett (3) give an excellent review of the use of maps in health research, and its problems and bene­fits. Based on their work, a number of aspects related to the use of spatial information can be formulated. Their most important argument is that map-making constitutes much more than presenting the geographical distribution of various diseases. Rather, it is a methodology that enables the handling of the spatial dimension of information. Another aspect related to map-making they bring forward is that “While maps are sometimes used to good effect, in other instances they tend to obscure evidence, to suggest false concentrations and to start false trails... The spatial distribution of diseases remains one of the oldest of puzzles and yet one of the most contemporary... They need to be handled with as much care and critical attention as any other source of evidence”. The authors describe how, by combining cartographic and statistical arguments, the value of maps can be enhanced and errors of interpretation reduced.

1 Free University, Department of Regional Economics, Amsterdam, The Netherlands; and National Institute of Public Health and Environmental Protection, Bilthoven, The Netherlands.
3 The general introduction of this article is based on reference (1).
Given that health and ill-health contain a spatial dimension, the science geography has always played a role in health research. Medical geography embraces both the study of geographical variations in the provision of health care and the distribution of diseases and the environment (e.g. climate, vegetation, water and air quality). However, it is almost impossible for a geographer to link cause and effect, knowledge about spatial variation in the incidence of disease is of significant value to the broader field of etiology and to medical science in general (4).

From this it clearly follows that the handling of spatial information can be defined as a methodology with strong links to geography, cartography and statistical science. It must be stressed that this methodology has to be developed by specialists in the field of application, in this case epidemiologists.

The objectives of GIS

Geographical information systems represent a technology designed to achieve particular objectives. In recent years, a variety of GIS products to assist the management and manipulation of spatial and non-spatial data have arrived on the market, and users worldwide have begun to gain familiarity with these new systems. Experience suggests that the application of GIS is definitely making significant contributions in facilitating the availability, integration and presentation of information.

A GIS can be defined as a set of tools for collecting, storing, retrieving, analysing and displaying spatial data. As a technology, a GIS is not necessarily limited to a single independent system, but may well have several components, each with a particular objective. A GIS must therefore accomplish the following four main tasks:

(i) To store, manage and integrate large amounts of spatially-referenced data. A spatially-referenced data base contains two types of information: locational and attribute data. Locational or spatial data are two- or three-dimensional coordinates of points (nodes), lines (segments) or areas (polygons). Nonlocational or descriptive data, on the other hand, refer to the features or attributes of points, lines or areas. Data are obtained from a wide variety of sources and one of the most important features of GIS is the facility to integrate data, e.g. converting data values to a common spatial framework.

(ii) To enable spatial retrievals. These spatial query possibilities are very flexible and still very powerful. It is possible to retrieve data for defined areas (polygons) like a municipality or postal code area, but there are other possibilities that would appear to be extremely relevant for public and environmental health purposes (5). For example, radial retrieval on a single point, radial retrieval for a search region defined as a buffer around a site’s perimeter, retrieval of data lying in a corridor focused on a linear map feature (e.g. an overhead wire), retrieval of data for an area defined by the user, retrieval of data for search areas defined by manipulating existing digital map data bases (e.g. relationship of cancer rates with geological patterns).

(iii) To provide methods of analysis which relate specifically to the geographical component of the data. The analysis techniques may be simple or sophisticated. At the simplest level, for example, data about different spatial entities such as soil type (per square kilometre) and land use (by local administrative area) can be combined by overlay analysis. At an intermediate level, GIS may allow statistical calculations of the relationship between data sets to be computed, or distances between entities may be used to determine the route to be followed to move as quickly as possible from one location to another. It must be mentioned that the implementation of GIS can make available an enormous amount of (point) data, which calls for new, explorative methods (6). The most sophisticated analysis occurs when modelling is introduced. In this context, there are a variety of analytical opportunities. For example, it is possible to use atmospheric modelling techniques to discover which areas might be affected by pollution resulting from an explosion at a particular hazardous installation (e.g. Chernobyl), given certain wind and weather conditions. Alternatively, modelling methods can be used to determine the impact of locating a large public facility (e.g. a hospital) at different sites in a city or region.

(iv) To display data on map forms of high quality. Maps no longer have to be drawn by hand; they are an implicit product of all the work carried out within a GIS. However, for many different purposes, other forms of display (e.g. graphs and tables) may also be required, often for use in combination with maps. The Ruimtelijke Informatie via Automatisering (RIA) system, developed by the National Physical Planning Agency in the Netherlands (7) is a good example of alternative methods of displaying spatial information in an interactive and user-friendly way.

Types of data storage

The distinction has already been drawn between locational and attribute data. It is important to distinguish further between three forms in which locational data can be incorporated within a GIS: raster, vector and quadtree storage.

Raster or grid storage.

This form of storage for locational data involves a regular grid of cells being laid over an area. Attribute data are collected for each grid cell which may measure, for example, 500 x 500 m². This means that the whole area is covered by a group of cells, each of which has an attribute value. Within a grid-oriented system of this type, only limited use is generally made of the attribute data. Satellite photographs, in which a considerably smaller grid size is used, provide raster information. In a satellite photograph, a single value is attached to each cell. In this way, factual data can be collected in a very efficient way. Curran (8) indicates how remote sensing data are frequently contained in raster-based GIS.

Vector storage

The storage of locational data in vectors gives a very precise representation of reality. In this way points, lines and areas are incorporated as Cartesian coor-
The initial GIS products date from the 1970s, a decade in which the typical hardware configuration comprised a central computer surrounded by memory and storage disks and a number of peripheral devices. Time-share systems enabled a large number of terminals with lines attached to the mainframe to be used at the same time. At the beginning of the 1980s, this centralized approach was extended by connecting minicomputers to the central mainframe in order to carry out certain processes. This was the period in which GIS came of age. Very large data bases began to be assembled, and the need for processor capacity increased enormously. In the middle of the decade, the personal computer (PC) arrived, although its impact for GIS meant little more than the provision of access to information by a large cell size is used whereas towards the edges of the area, the cell size diminishes to form a precise picture. An area is therefore covered by considerably fewer quadtree cells of varying size than is the case in the regular grid storage method. The speed with which analyses can be carried out with a quadtree structure is high, whilst the original precision of the data is retained rather well.

Quadtree storage

The quadtree storage method falls between the grid and vector storage methods. In this method, the data are stored in grid cells of variable size. Within larger homogeneous areas, a large cell size is used whereas towards the edges of the area, the cell size diminishes to form a precise picture. An area is therefore covered by considerably fewer quadtree cells of varying size than is the case in the regular grid storage method. The speed with which analyses can be carried out with a quadtree structure is high, whilst the original precision of the data is retained rather well.

Hardware and software in a GIS environment

Hardware

The initial GIS products date from the 1970s, a decade in which the typical hardware configuration comprised a central computer surrounded by memory and storage disks and a number of peripheral devices. Time-share systems enabled a large number of terminals with lines attached to the mainframe to be used at the same time. At the beginning of the 1980s, this centralized approach was extended by connecting minicomputers to the central mainframe in order to carry out certain processes. This was the period in which GIS came of age. Very large data bases began to be assembled, and the need for processor capacity increased enormously. In the middle of the decade, the personal computer (PC) arrived, although its impact for GIS meant little more than an extra terminal in many cases. Nevertheless, the PC has become central to the popularization of GIS.

In the second half of the 1980s, the PC played an important role in simple GIS tasks such as the automatic production of maps. It is exactly this function which has brought GIS to the attention of many organizations, whereas the basic concept of a common central data base tends to be forgotten. Attempts to transfer fully-fledged GIS from the mainframe onto the PC have been commercially successful, even though in many cases their performance leaves something to be desired. Of much greater importance in the mid-1980s was the further development of minicomputers and workstations connected to a network. At this time, the larger organizations making use of GIS realized that the central mainframe option was not the solution to a number of GIS tasks. It was recognized that each separate task required its own processor capacity or its own working environment. Throughout the 1980s, it became clear that hardware vendors were meeting these demands perfectly well by means of further optimization of minicomputers and the increased processor capacity of workstations, and that hardware costs were decreasing significantly.

Software

Identification of the main tasks of a GIS environment allows us to specify a corresponding set of software demands.

Data-base management and query. One of the most significant advances in GIS software development came at the end of the 1970s with the introduction of the concept of the relational data base. In the relational model, different data sets are linked together by the use of common key fields. For example, attribute data available for two different sets of spatial units (areas) will require a third set of information to show how the two spatial bases fit together. This type of structure can also be used to construct spatial data bases in which lines are linked together to represent polygons. The creation, maintenance and accessing of a data base requires a data-base management system (DBMS). In order to handle very large quantities of information, a relational DBMS is a necessity for most GIS applications (9). It facilitates data manipulation and analysis. Many of the proprietary GIS have their own systems to handle basic storage, management and analysis operations (e.g. INFO is the relational DBMS of ARC/INFO; ORACLE is the relational DBMS of ARGIS).

Analysis. The software required to perform certain analytical tasks varies according to the nature of the problem, the quantity of information available and the objectives of the organizations involved. A variety of analytical tools is now available within GIS. The overlay procedure, for example, has been widely used for combining different data sets in order to identify areas or sites with required characteristics (10). Buffering, addressing mapping, and network analysis are additional tools adopted in planning applications of GIS. However, the development of analytical functionality within GIS has tended to neglect the important benefits that modelling procedures can contribute through data transformation, integration, updating, simulation, optimization, impact assessment, and forecasting (8-9). The specialized and complex nature of modelling algorithms and the restrictions imposed by processor capacity have both been influential in keeping the modelling component separate. One of the challenges confronting the next generation of GIS is to improve the integration of modelling and GIS so as to provide researchers, decision makers and planners with enhanced model-based decision-support systems.

Access and presentation software. One of the most important functions of GIS development is frequently the provision of access to information by a variety of different users. In the context of planning, individuals in different departments of the same organization (public works, social services, transport, parks, etc.) may require access to the same data base. Similarly, users in different national, regional and local organizations may wish to access the same data base simultaneously. The National On-line Manpower Information System (NOMIS) in the United Kingdom is a good example of a GIS which stores up-to-date information about employment, un-
The technological advances in computer hardware which have been occurring over the past 20 years have had a direct impact on the presentation of information in GIS. There have been striking advances in automated mapping, i.e. the development of programs enabling maps to be produced automatically (12). The experience gained over this period has meant that very professional products are now the norm, and micro-based mapping software (ATLAS*Graphics, GI1MMS, MICROMAP, for example) provides high-quality output.

Applications and users of GIS

Applications of GIS

Traditionally the use of GIS started in two fields: urban and regional planning, and the construction and maintenance of utilities (electricity, water, etc.). As with other computer applications, only a few specialists knew how to handle computerized geographical data. Nowadays, applications of GIS are many and varied (13). Associated with the emerging awareness of the fact that many types of information contain a locational or spatial component, there is a growing utilization of geographically-referenced data. Some of the fields of application of GIS are the following:

- traffic and transport planning
- agricultural planning
- environment and natural resource management
- recreation planning
- location/allocation decisions
- spatial planning (land use)
- service planning (education, police, health, social, etc.)
- marketing
- public and environmental health research and policy.

Type of user and kind of need

When acquiring a GIS, one of the first steps to be taken is analysis of the needs of users and the decision-making process of which the GIS is intended to be part and to influence. This step is critical to GIS success because it requires the potential users of the system to identify the information they need in order to perform their work. Such an analysis will result in an identification of different user types and their differentiated needs. A number of categories of users can be identified on the basis of the objectives of the organization to which these users belong; similarly organizations can be classified according to the type of activity that each performs. The type of GIS that is adopted and applied therefore varies between categories of organizations and between organizations of different size and function within the same category. However, it is possible to identify particular groups of individuals across the spectrum of organizations whose occupational characteristics with respect to GIS are distinctive.

Different types of organizations. Four main types of organization may be distinguished. Firstly, there is the research institute where research may be carried out to find solutions to problems or answers to questions posed by external paymasters. Data collection and manipulation take place and descriptive, explanatory and predictive analyses are undertaken. Secondly, there are administrative institutions such as public utilities or property-registration agencies. Here, the objective is to manage information in such a way that the process of acquiring and manipulating data is made as simple as possible. The management of the waste-disposal system of sewerage pipes for a local administrative area is one example where accurate and quick answers are required on the basis of the information stored in the GIS to questions such as: where are the oldest parts of the system?; what is the total length of pipe involved?; at what depth are the pipes buried?; and how many houses are connected to these pipes? The third type of organization is the government agency whose objective is to formulate policy recommendations. For this purpose, concept design and evaluation take place. Commercial enterprises are the fourth type of organization. Their aim is to maximize their profits by selling goods and services. Information is collected and manipulated within an integrated modelling/GIS environment, for example to establish optimum locations for new retail outlets.

Different types of user groups. Within every organization, there is a differentiation in functions. Different functions and different needs for, and use of, information are interrelated. It is clear that because of these differences, the user demand on GIS varies. The following groups of users can be distinguished: information specialists, researchers, research coordinators, policy planners, decision makers and third parties.

In each of the above four categories of organization, information specialists are required to acquire and manage data, computer hardware and software. The information specialist usually works with the raw data and requires a "large" GIS (e.g. ARC/INFO, ARGIS, SYSTEM9), in which the data base is fundamental and flexible, and can be connected to other systems. Researchers, on the other hand, tend to be confined to their own institutes or to commercial companies. They work either with raw data or data that have been partially processed or transformed. They demand user-friendliness from a GIS, analytical features (as with SPANS, for example) and appropriate interfaces that allow transfer of information to other packages for modelling and other purposes. Research coordinators are concerned with the inter-relationship between the different products of the organization and therefore work with manipulated data and require a simple, user-friendly GIS. Policy formulation is usually the responsibility of a government agency, and the main requirement of a GIS in this context is that it should be easy to use. The same applies to the decision makers in administrative, government and commercial organizations, whose job is to translate information into policy statements, and to third parties, who simply utilize information provided by government agencies or research institutes.

From low-end to high-end GIS

The distinction between different types of user of spatial information has important consequences in practical terms, in the application of GIS software. In general, a GIS enables four main functions: input, management, analysis and presentation of geographical data. Data input includes such functions as data collection, digitizing point data, and editing. Management is the handling of permanent alphanumeric and geographical data. The analysis function examines
data in order to create new data, the purpose of which is to produce information. All operations which produce graphic output fall under the presentation function. From the foregoing it clearly follows that not all users wish or need to make similar use of these main functions.

The tools for handling geographical data differ widely in level of sophistication. Meijer (14) refers to this as a continuum from low-end to high-end GIS. At the one end are simple microcomputer mapping packages solely for presentation of data, at the other end the mainframe implementation of complex tools for all kinds of geographical analysis. The emergence of new hardware and software systems has greatly contributed to the possibilities for different users to choose the right GIS tools.

High-end GIS refers to a full integration of all tools in one system which is, inter alia, capable of handling raster (grid) and vector information, and to perform analysis, such as overlay techniques, buffering etc. Low-end GIS consists of tools that have only part of this functionality. Notwithstanding the fact that a high-end GIS has a much larger functionality than a low-end GIS, not all users are professionally trained GIS users and would probably be better off with the right GIS tools at the lower end of the GIS spectrum. From the foregoing it follows that in most cases three characteristics are important: user-friendliness, interfaces and output quality.

User-friendliness is most important for the decision maker who generally has little knowledge of computers. Even the reading of a manual should not be necessary for completely self-explanatory programs. The decision maker should be able to use the GIS tools almost without any specialist training. A transparent menu-driven program is most suitable. User-friendliness is far less important for the information specialist who is used to working with commands and will probably only use menus if this would speed up the work.

Interface with other software — sometimes high-end GIS — is however of utmost importance for the information specialist. Both the locational and attribute information must be easily transferable to and from other systems for a proper functioning of low-end GIS. Those engaged in research, policy preparation and policy making, are less interested in the graphics interface; in a true GIS working environment the information specialist takes care of the conversion. These user groups will not create the locational information they need themselves. The digitizing will often be obtained from third parties. The locational information they work with will in general not be derived from the data, but be a priori given, like administrative boundaries. The data-base interface is critical to these users as they will have all kinds of attribute information in packages like dBaseII+, Lotus123 or a statistical package such as SPSS/PC+. Finally, transferring data to the geographical GIS tools should be very easy.

The third aspect of low-end GIS, the quality of the output, refers to the end products of the GIS process which are used in reports and must therefore be of high quality. High-speed media-like electrostatic plotters are required by information specialists for producing complex maps. For all the other users, the relatively low-priced A4/A3 pen-plotters and laser-printers will do.

In conclusion, the successful implementation of a GIS depends on a careful preliminary assessment of the type required to meet the demands of the various users of the system. Ideally, there should be detailed functional specifications for each of the needs to be met by a GIS. Whilst GIS packages are now available for the range of GIS users that we have identified (Fig. 1 (15) shows the different user groups and their varying demands), there remain considerable difficulties, particularly in large organizations, in establishing the configuration of hardware and software, and in enabling the easy exchange of information between the machines and packages involved. It is important to ensure that there is a GIS management framework which takes into account the separation of functions between user departments and the central data store.

The application of GIS in health research

A large number of available applications may fall under a GIS umbrella, but in most cases it is questionable whether or not these applications constitute GIS, or "old-fashioned" hand-made maps. Automated mapping is the most well-known field of GIS in health research at present. Map 1 shows an example of an automated map (16).

![FIG. 1](image)

**FIG. 1**

**USER GROUPS OF GEOGRAPHICAL INFORMATION SYSTEMS (GIS) AND THEIR DEMANDS**

<table>
<thead>
<tr>
<th>Type of user</th>
<th>Information demand</th>
<th>User demand</th>
<th>Type of GIS</th>
<th>Development</th>
</tr>
</thead>
<tbody>
<tr>
<td>Information specialist</td>
<td>Raw data</td>
<td>Management</td>
<td>Large</td>
<td>Links to other packages</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Analysis</td>
<td>Flexible</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Flexibility</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Researcher</td>
<td>Raw data and pretreated data (= information)</td>
<td>Analysis</td>
<td>Compact manageable GIS software</td>
<td>Macrolanguages Problem-oriented</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Good accessibility</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Specific</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Manager/decision maker</td>
<td>Strategic information</td>
<td>Good accessibility</td>
<td>&quot;Small and beautiful&quot; Key information</td>
<td>User-friendly interfaces</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Weighting and optimization models</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Target group/third parties</td>
<td>Information</td>
<td>Good accessibility to users</td>
<td>&quot;Small and beautiful&quot;</td>
<td>User-friendly interfaces</td>
</tr>
</tbody>
</table>

Source: Based on reference (15).
### Fig. 1

**GROUPES D'UTILISATEURS DES SYSTÈMES D'INFORMATION GÉOGRAPHIQUE (SIG) ET LEURS EXIGENCES**

<table>
<thead>
<tr>
<th>Type d'utilisateur</th>
<th>Besoins en information</th>
<th>Exigences de l'utilisateur</th>
<th>Type de SIG</th>
<th>Développements</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spécialiste de l'information</td>
<td>Données brutes</td>
<td>Gestion</td>
<td>Etendu</td>
<td>Liaisons avec d'autres progiciels</td>
</tr>
<tr>
<td>Chercheur</td>
<td>Données brutes et données pré-traitées (= information)</td>
<td>Analyse, souplesse</td>
<td>Souple</td>
<td>Logiciel SIG compact et souple</td>
</tr>
<tr>
<td>Décideur/gestionnaire</td>
<td>Information stratégique</td>
<td>Bonne accessibilité, spécificité</td>
<td>«Petit et beau»</td>
<td>Interfaces conviviales</td>
</tr>
<tr>
<td>Tiers/groupe cible</td>
<td>Information</td>
<td>Facilité d'accès</td>
<td>«Petit et beau»</td>
<td>Interfaces conviviales</td>
</tr>
</tbody>
</table>

Source: D'après la référence (15).

### MAP 1. EXAMPLE OF AN AUTOMATED MAP

**Stomach Males 1979-82**

In recent years a large number of automated atlases has been produced; a complete, geographical review is given by Cliff & Haggett (3). These atlases show how interesting patterns of spatial distribution of health and ill-health can often be revealed simply by mapping. Once the characteristics of a mapped distribution of a disease are described, however, questions will be raised about the reasons for a particular geographical distribution. In this sense therefore, maps may provide an indication of those areas on the map in which further research may be useful. Interpreting spatial patterns and looking for relationships lies within the domain of spatial analysis and in this case more specifically in the field of epidemiology.

Spatial analysis encompasses GIS functions with which computer mapping is not equipped to deal, as has become clear in the case study on childhood cancer in Northern England (17), which was set up as a result of the current considerable interest with regard to cancer clusters and their possible association with environmental factors.

In an attempt to search for the relationships between geographical correlates of children with cancer in...
Northern England, a set of spatial analysis methods has recently been integrated and automated — the Geographical Analysis Machine (GAM) (17). One impetus to the development of GAM has been that previous epidemiological attempts at finding evidence of cancer clusters provided uncertain and spurious results of the data, by virtue of the arbitrary geographical units used for analysis, such as districts or wards. With the development of GAM, a boundary-free method is provided and thereby an attempt is made to move away from arbitrary partitions of geographical space. The method is based on both point and area information. Area information, based on enumeration districts (ED) — the smallest building blocks for which census data are released (e.g. small area health statistics in the United Kingdom) is used to ascertain the number of children at risk, and random selections are made from these data in order to determine expected point distribution of cases. The link from areas to points is made via the ED centroids, grid references accurate to 100 m, provided for each areal unit. The method then counts the observed number of cases falling within a circle of given size and location and compares this with the expected value. If the count significantly exceeds the expectation, the circle is plotted on the map. Multiple overlapping circles arise because the potential centres of circles are locations on a fine grid and because a wide range of circle sizes is considered. Clearly, this method provides the possibility to distinguish "real" from "illusory" clusters. Although it has not yet been undertaken, a follow-up on the testing of the existence of localized aggregations of disease by means of GAM might be to extend the research by using a GIS framework. This would allow for a more intensive search for possible multivariate associations via spatial analysis.

That a GIS approach certainly offers benefits in epidemiological research may be further illustrated by a case study on cancer of the larynx in North-West England (18). GIS turned out to improve the possibilities of description and analysis of the disease, especially since large data sets such as those from national and regional cancer registries were used. The ability to execute multiple queries on a data set, together with the use of high resolution graphics, had greatly facilitated manipulation of the data, such that subsets of cases by cancer site, sex and age could be readily extracted and analysed. In addition, GIS seemed very helpful in connection with the assessment of the relationship of disease to environmental variables. The facilities offered by GIS appeared particularly valuable, enabling users to define their own areas of interest (areas around incinerators).

Public and political concern about possible links between the emission of dioxin from commercial and municipal incinerators and ill-health has been the main incentive to this epidemiological study. Earlier studies by the same North-West Regional Research Laboratory in Lancaster (United Kingdom) had suggested a possible link between cancer of the larynx and proximity to an industrial waste incinerator. The empirical study taken up consequently tried to assess whether there is any association between the geographical distribution of laryngeal cancer and proximity to an industrial waste incinerator; it included testing the hypothesis that cancer incidence is significantly higher around incinerators than elsewhere. In the first instance, the study was restricted to one of the 19 district health authorities that constitute the regional health authority — where an industrial incinerator that had operated in the dis-
district for several years had been closed. Data on cancers registered between 1974 and 1983 were used. First of all, some simple descriptive mapping was undertaken to see if there was any visual evidence of localized clustering of cases in the vicinity of the incinerator. Such mapping was made possible by converting the unit post codes of cases to ordnance survey grids, accurate to 100 m. The computerized central post-code directory (a directory wherein unit post codes are associated with an ordnance survey grid reference) was used for this purpose. Point distribution maps showed that the distribution of most cancers mirrored the distribution of the population as a whole. The exception to this was laryngeal cancer. Here, out of only 58 cases found, 5 were within about 2 km of the closed industrial incinerator. Clearly, the mapping provided a first indication on where further investigation was needed.

Consequently, the study was extended to a larger geographical area. Now, much larger data sets were used, namely on cancer of the larynx within the entire area covered by the North-West Regional Health Authority. The cancer data, including age, sex and cancer site of notification, unit post code, ordnance survey grid reference, were imported into a GIS which uses both vector and raster types of data storage as well as various spatial analysis techniques, together with digitized electoral ward boundaries and data from the 1981 population census. The latter included the age-sex data needed to compute expected incidence of cancers. GIS made it possible to attach a ward code to each case. The observed number of cases, by sex, cancer site and ward, were then counted. Expected occurrences per ward were calculated by applying lung and laryngeal cancer rates for the North-West Regional Health Authority, based on sex and ten-year age groups, to the age-sex population at risk. Standardized registration ratios (SRRs) — i.e. measures of relative risk — of cancer of the larynx were determined, and those wards with significant SRRs were displayed. The GIS was also used to define areas possibly at risk from pollution by hospital incinerators.

The results of the GIS operations were consequently put into a statistical model of cancer incidence based on the theory of spatial point processes. Briefly, it entailed setting up a formal statistical model which states that the intensity of cases of the larynx is a function firstly of background intensity (population at risk), and secondly of distance from the putative source of radiation. The model tests whether proximity to the pollution source is a significant influence on the distribution of this cancer. This method clearly enables separating out “point hazard” effects from the expected “natural” clustering due to population distribution.

Although the number of applications of GIS in health research is growing very fast, at present GIS is still an unfamiliar working environment. To give an example of a geographical study which could have benefited from GIS, we will briefly elaborate on a study on Chernobyl fall-out and perinatal mortality in England & Wales (19).

Earlier studies had concluded that radioactive fall-out from Chernobyl (26 April 1986) may have caused an increase in perinatal mortality in Western Germany (20) and the United States of America (21, 22), and marked geographical variations in contaminations from Chernobyl seemed to exist in England & Wales. This prompted researchers in the United Kingdom to launch a geographical study further investigating this question. The search for evidence of the existence of a relationship between an increase in perinatal mortality and radiation was the focus of attention. Whether increased perinatal mortality rates in the period following Chernobyl were relatively higher in regions with higher radiation doses than elsewhere was the main research question. The highest doses of radioactive fall-out from Chernobyl in England & Wales were found in the counties of Cumbria, Clwyd and Gwynedd where there was heavy rainfall during the passage of the radioactive cloud. Observed and expected numbers of stillbirths, neonatal deaths and perinatal deaths in the three counties in the year following Chernobyl were calculated. Expected numbers were calculated from national rates multiplied by the ratio of rates in the three counties to national rates during the years 1981-1985. Mortality data for these counties were adjusted to national rates so that the hypothesized increase in mortality rates would not be falsely accepted due to the fact that the regions studied reflected any general tendency for mortality rates to be different from national rates.

Evidence for an effect of Chernobyl would be greater increases in perinatal mortality in regions with higher radiation doses and less elsewhere, but no such evidence was found. In fact, it was observed that perinatal mortality in these areas did not rise relative to the national average in the year following Chernobyl. The significant deficit of perinatal mortality in these three counties raised the question of whether this was strictly a local event or whether it had wider significance. Consequently, a geographical study of a wider range of areas was undertaken. Since deposition may be a poor guide to radiation doses received by the population, consumption of contaminated milk was taken into account to investigate the pattern of perinatal mortality in the year following Chernobyl for any relationship with geographical radioactive deposition. Estimates of contaminated milk were provided for, and observed and expected numbers of perinatal deaths were derived by the same methods used in the analysis of Cumbria, Clwyd and Gwynedd. But the negative finding of the earlier study was confirmed: there is no evidence that radiation from Chernobyl caused a rise in perinatal mortality in England & Wales.

It is important to stress here that although the Chernobyl study did not actually make use of GIS, it provides a good example of how GIS can be helpful in such a study. Several aspects of the study favour using a GIS methodology. Some examples of the operations that GIS could have performed in this case are the following:

- Large sets of spatially-referenced data were used which within a GIS can be handled, integrated and manipulated quickly and easily.
- Some spatial analysis functions that a GIS can perform could have been used. At a very simple level, data about different spatial entities such as perinatal mortality rates by local administrative area and fall-out doses per user-defined area, can be combined by overlay analysis.
- Lastly, map-making might have been useful because a map is worth a thousand words and constituted therefore an excellent displaying device.

The most important conclusion to be drawn and lesson to be learned from the above case studies is that applying a GIS approach to health research can be of great value. Not only does GIS offer possibilities to perform various tasks fundamental in such research far more quickly and with less effort, it
also provides health researchers with new, reliable and scientifically valid methods for handling their spatial information. By using GIS, functions can be performed with which automated statistical and/or computer mapping packages are not equipped to deal. In this sense, GIS may lead the way to new insights into "old" information and thereby even contribute to a better understanding of the health problems faced nowadays. However, caution should be exercised since GIS itself will never prove causal relationships, and it is up to analytical epidemiologists to adopt the new technology.

The application of GIS in the European context

The starting point for European cooperation with regard to policy and research on health and the environment is recognition of the international character of environment and health. Many environmental and health problems need to be dealt with internationally. "The programmes that protect and enhance health in relation to the environment face many of the same issues in different European countries, even if priorities vary. Much of the necessary research, assessment and codes of practice can therefore be accomplished most effectively internationally" (23). It is therefore not surprising that large organizations try to pool their efforts when dealing with environmental and health policies and research.

One of the first examples of European initiatives to pool efforts was the CORINE project, which was set up by the European Community in 1985. The main impetus to the CORINE programme was the need to gather, coordinate and ensure the consistency of information on the state of the environment and natural resources. The programme had two aims, namely:

(i) to verify the usefulness of a permanent information system on the state of the environment for Community environmental policy, to check the technical feasibility of creating such a system, and to identify the conditions required for its installation and functioning;

(ii) to supply information useful for Community environmental policy.

Concerning the first aim, the CORINE programme results show that a permanent information system on the state of the Community environment is necessary and technically feasible.

The second aim has also been successfully attained. Data on priority topics were collected, supplemented by a series of basic data, and organized in an operational GIS. Therefore the Council of Europe ministers took the decision to transform the CORINE prototype into a permanent information system.

The CORINE experiences and results have encouraged other European initiatives aimed at joint efforts with regard to environmental and health problems. At the First European Conference on Environment and Health, held in the Federal Republic of Germany (7-8 December 1989) the ministers of the environment and of health of the Member States of the European Region of the World Health Organization (WHO) adopted the European Charter on Environment and Health, and accordingly agreed on principles and strategies laid down therein as a firm commitment to action. Target 19 of the Charter reads:

"By 1990, all Member States should have adequate machinery for the monitoring, assessment and control of environmental hazards which pose a threat to human health, including potentially toxic chemicals, radiation, harmful consumer goods and biological agents. The achievement of this target will require the establishment of well-coordinated monitoring programmes with clearly defined objectives; the development of methodologies and health criteria for the assessment of data in relation to control procedures; the investment of adequate levels of funding for control measures, and their introduction and maintenance; and the training and utilization of sufficient numbers of competent personnel for all aspects of environmental health protection" (23).

Major developments in the implementation of the Charter have recently taken place. The WHO Regional Office for Europe decided in 1989 to start the Health and Environment Geographical Information System (HEGIS) project, a permanent geographical information system for public health and environment for Europe, based on the cooperation of national focal points, WHO collaborating centres and research institutes. The National Institute of Public Health and Environmental Protection, Bilthoven, the Netherlands, was asked to take the lead in the development and application of HEGIS. It is planned that the HEGIS project to be one of the cornerstone activities of the WHO European Centre for Environment and Health in the Netherlands established in 1990.

Conclusion

As demonstrated in this article, the potential of a GIS lies in the areas of collecting, storing, retrieving, analysing and displaying spatial data. Associated with recognition of the fact that, in principle, almost every aspect of the environment potentially affects health for good or ill, and that health and ill-health almost always contain a spatial dimension, it is not surprising that the potential for applying GIS in health research and policy at both national and international levels is beginning to be realized.

The number of applications of GIS in the field of health policy and health research has been limited however. At this early stage, there is an emerging and growing awareness among health professionals of the possibilities of GIS. These professionals are unskilled with regard to the implementation of such systems, hence the need to adopt a multidisciplinary approach to explore the possibilities offered by GIS to improve our knowledge of public and environmental health.
SUMMARY

One of the most important issues in public and environmental health today concerns the type of instruments that can be used to devise quick, reliable and scientifically valid methods of rapid assessment which, in turn, can be utilized in health research and in the planning, monitoring and evaluation of health programmes. As the applications of geographical information systems (GIS) relate to the collection, storage, integration, management, retrieval, analysis and display of spatial data, it is not surprising that the potential usefulness of this new technology in the fields of health research and policy is beginning to be realized.

This article seeks to demonstrate the opportunities which the use of geographical information systems can offer to research and policy on health issues. The article first describes the principles and objectives of GIS before going on to discuss hardware and software developments as well as the variety of application fields, organizations and users. Some examples of current applications are provided to illustrate the type of work being undertaken. The final sections address issues specifically related to the application of GIS in health research and policies in the European context.

RÉSUMÉ

Les avantages de l’utilisation des systèmes d’information géographique en santé publique et en hygiène de l’environnement

Aujourd’hui, l’une des tâches les plus importantes, en santé publique et en hygiène de l’environnement, consiste à déterminer quel est le type d’instruments à utiliser pour élaborer des méthodes rapides, fiables et scientifiquement valables d’évaluation immédiate qui, à leur tour, pourront être appliquées à la recherche en santé ainsi qu’à la planification, la surveillance permanente et l’évaluation des programmes de santé. Étant donné que les systèmes d’information géographique (SIG) englobent la réunion, le stockage, l’intégration, la gestion, la recherche, l’analyse et l’affichage de données géographiques, il n’est pas étonnant que l’on commence à s’intéresser aux possibilités d’application de cette nouvelle technologie dans les domaines de la recherche et des politiques sanitaires.

Cet article se propose de montrer quelles sont les voies ouvertes par les systèmes d’information géographique en matière de recherche et de politique sanitaires. L’article donne tout d’abord une description des principes et des objectifs des SIG, avant d’étudier les développements intervenus dans le domaine du matériel et du logiciel, de même que la variété des applications et des structures, ainsi que la diversité des utilisateurs. On y trouvera divers exemples d’applications courantes illustrant les types de travaux présentement exécutés. Les dernières sections traitent de questions concernant plus particulièrement l’application des SIG aux politiques et à la recherche en santé dans le contexte européen.

REFERENCES — RÉFÉRENCES


Disasters requiring international action are occurring with increasing frequency and expenditures on emergency relief are absorbing significant proportions of development aid. Although no accurate records are maintained on emergency relief, incomplete reports from national and international agencies estimate relief disbursements at over US$ 1 000 million each year. The Sahelian famines of the early 1970s, the earthquakes in Tangshan (China) (1976), Mexico (1985) as well as the Armero volcanic eruption demonstrated the need for a professional approach to mass emergency responses and the importance of preparedness in developing countries. In addition, there has been a significant increase in mortality in nearly all types of disasters between the periods 1960-1969 and 1970-1979. The mortality rate increased from 750 per event in the previous decade to 4 871. (Table 1). This increase is probably largely due to rampant urbanization, environmental degradation and population pressures on land (1). Finally, the long-term consequences of acute nutritional stress resulting from famines or harvest loss in floods and cyclones, waterborne diseases and disability are rarely assessed and addressed by emergency programmes.

As a result of current inefficiencies in the approach to health management in disasters, the ever-increasing expenditures have not had any visible results in attenuating or preventing the ravages of natural disasters in the developing world. The handling of relief as purely charitable exercises, providing whatever aid is readily available regardless of needs, has characterized emergency aid. Health response in emergencies has been typically ad hoc action that is generally inappropriate and usually late. Today, field agencies, donors and national governments recognize the need for rationalizing the response to emergencies so that the needs are assessed correctly and on time (2). The importance of preparedness, especially regarding the availability of information necessary for planning rapid response and serving as early warning signals, has been identified as a key element in the improvement of health management in disasters (2).

**Differential impact of natural disasters and the relevance of an epidemiological approach**

The impact of disasters on populations varies according to the type of disaster, but specific population subgroups also differ in their vulnerability to disaster impact. A greater increase in the lethality impact of some disasters has been observed. For example, earthquakes have shown the greatest increase in mortality in the last two decades. When aggregated over countries, mortality and morbidity from disasters seem to be significantly higher for countries with a low GNP than for more affluent countries, even when controlled for population density (Table 2).

Table 3 presents data from two different earthquakes, one in Managua (Nicaragua) and the other in San Fernando Valley, California. The earthquake in California resulted in 60 deaths, while the smaller quake in Managua left over 5 000 dead. Although many other factors besides seismic intensity and socioeconomic factors influence the number of fatalities and injuries from an earthquake, this table gives some indication of the significant influence of socioeconomic conditions on the mortality and morbidity resulting from earthquakes.

The risk of disaster-related mortality and morbidity is also determined by demographic characteristics. Figs 1 & 2 present data from selected disasters where the vulnerability of older children to mortality was substantially higher than other age groups. De Bruycker et al. (4) in their study of the earthquake in Campania (Italy) also observed that children between 5 and 9 years were at higher risk of injury and death than smaller children. This could be explained by the fact that parents take care of small children in crisis...
TABLE 2. DISASTER MORTALITY BY LEVEL OF ECONOMIC DEVELOPMENT
TABLEAU 2. MORTALITÉ PAR CATASTROPHE EN FONCTION DU NIVEAU DE DÉVELOPPEMENT ÉCONOMIQUE

<table>
<thead>
<tr>
<th>Mortality - Mortalité</th>
<th>Economy - Économie</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low income - Faible revenu</td>
<td>Middle income - Revenu moyen</td>
</tr>
<tr>
<td>Per event - Par événement</td>
<td>3 300</td>
</tr>
<tr>
<td>Per 1 000 population - Pour 1 000 habitants</td>
<td>69</td>
</tr>
<tr>
<td>Per 1 000 km² - Pour 1 000 km²</td>
<td>48</td>
</tr>
</tbody>
</table>

Source: Adapted from reference (1) – Adapte de la référence (1).

TABLE 3. COMPARISON OF CHARACTERISTICS OF EARTHQUAKES IN MANAGUA (1972) AND SAN FERNANDO VALLEY, CALIFORNIA (1971)
TABLEAU 3. COMPARAISON DES CARACTÉRISTIQUES DES TREMBLEMENTS DE TERRE SURVENUS À MANAGUA (1972) ET DANS LA VALLÉE DE SAN FERNANDO, CALIFORNIE (1971)

<table>
<thead>
<tr>
<th>Disaster characteristics - Caractéristiques de la catastrophe</th>
<th>Managua</th>
<th>San Fernando Valley</th>
</tr>
</thead>
<tbody>
<tr>
<td>Richter scale reading - Degré de magnitude sur l'échelle de Richter</td>
<td>5.6</td>
<td>6.6</td>
</tr>
<tr>
<td>Extent of destruction (Mercalli Intensity Range VI-VII) - Etendue des destructions (VI-VII sur l'échelle d'intensité de Mercalli)</td>
<td>100 km²</td>
<td>1 500 km²</td>
</tr>
<tr>
<td>Population in affected area - Population des zones touchées</td>
<td>420 000</td>
<td>7 000 000</td>
</tr>
<tr>
<td>Dead - Morts</td>
<td>5 000</td>
<td>60</td>
</tr>
<tr>
<td>Injured - Blessés</td>
<td>20 000</td>
<td>2 540</td>
</tr>
</tbody>
</table>

* The Richter scale for expressing the intensity of an earthquake ranges from 0 to 8 - L'échelle de Richter, qui exprime la magnitude d'un tremblement de terre, va de 0 à 8.

Source: Reference (8) – Référence (8).

FIG. 1
AGE-SPECIFIC MORTALITY, SUMPANGO EARTHQUAKE, GUATEMALA, 1976
MORTALITÉ PAR ÂGE, TREMBLEMENT DE TERRE DE SUMPANGO, GUATEMALA, 1976

Studies on risk for disaster-related mortality and morbidity have identified factors linked to population density (3), structural quality (6), time of strike (4) and intensity of seismic activity (7). However, the risk of mortality and morbidity in disasters is clearly not only a function of physical characteristics of the event but is also determined by the prevailing socioeconomic and health conditions of the affected community (8). The differential health impact of disasters on a community indicates that the potential for efficient and accurate rapid assessment techniques can conditions while they expect the older children to take care of themselves. Guha-Sapir et al. (5) noted similar vulnerability patterns in a survey of affected communities in Chad during the 1985 famine. Mortality was higher among children >2. They concluded that infants were protected against the decrease in food intake by being breastfed while older children were mistakenly expected to be able to secure their own food. The study also noted the increased vulnerability to famine of specific occupational groups within the community.
FIG. 2
AGE-SPECIFIC MORTALITY DURING TWO EARTHQUAKES (NICARAGUA 1972 AND GUATEMALA 1976)
TAUX DE MORTALITÉ PAR ÂGE LORS DE DEUX TREMBLEMENTS DE TERRE (NICARAGUA 1972 ET GUATEMALA 1976)

<table>
<thead>
<tr>
<th>Town/Ville</th>
<th>Number of deaths</th>
</tr>
</thead>
<tbody>
<tr>
<td>Managua 1972</td>
<td>N = 158</td>
</tr>
<tr>
<td>Patzicia 1976</td>
<td>N = 377</td>
</tr>
<tr>
<td>Sumpango 1976</td>
<td>N = 244</td>
</tr>
<tr>
<td>Santa Maria Cauque</td>
<td>N = 78</td>
</tr>
</tbody>
</table>

be developed using indicators related to death and injury and robust sentinel surveys.

Inefficiency of health relief and role of better assessment

Recently, the repeated disasters in sub-Saharan Africa and in South Asia (in particular, Bangladesh) and the massive international operations that they have entailed, have raised the issue of preparedness of local communities and efficiency of response. As a result of this attention, many different aspects of disaster management have come to light and information has become a major issue. In the last decade, the need for information has crystallized in two main areas: (i) early warning systems for famines became a central issue for famine control and prevention; (ii) methods for rapid assessment came to the forefront for all emergency interventions. This interest in rapid assessment subsequently revealed the inappropriateness of relief in terms of time delay and content. The neglect of proper assessment of needs tended to produce health relief founded more on rumour than on fact and therefore led to inefficient and inappropriate use of limited health resources.

Until recently emergency response, despite sufficient funds and goodwill, has consistently been late and inappropriate, mainly because of inadequacies in rapid assessment. Relief commonly arrives well after the first crisis has passed and continues to arrive beyond the period of need. As a consequence of this continuously lagged response, relief remains largely irrelevant. Finally, when the long-term effects begin to manifest themselves, relief dries up altogether, since these needs were not assessed during the emergency period.

The arrival of large quantities of inappropriate medications, standard relief articles (such as blankets, clothes), surgical and anaesthetic teams is a direct result of political or humanitarian pressures for action. Decision makers are unable to wait for long-drawn-out assessment results, and any available materials are sent. Medical teams and relief supplies that are not required result in logistical backlogs and take time and resources for storage, inventory and sorting. Drugs require classification, storage and occasionally, destruction. Relief personnel require housing and supervision. Volunteers require coordination. This can take valuable time and resources away from the main objectives of preventing and mitigating further deaths and damage. For example, on average 60 agencies arrived each week in Bangladesh over a period of four months following the 1988 floods. While these cases sound anecdotal, they unfortunately characterize the profile of the majority of relief aid today. The urge to send medi-

*Blankets sent to India following the 1976 floods were donated to Nepal in 1982 and redonated to India in 1987; relief for the victims of the Mexico earthquake included contraceptives and high-heeled shoes.
cines, personnel, surgical items and any other handy material has been rooted in the international image of disaster relief. For example, the standard response of assistance agencies to an earthquake or other acute emergency has been to send teams of trauma surgeons or anaesthesiologists. In Mexico (1985), two emergency surgical teams arrived to find that the city of Mexico had sufficient facilities and skilled personnel to treat victims who needed immediate care. The teams returned home 11 days later never having dismantled their field hospitals or treated anyone. (G. D'Allemagne, personal communication, 1985). More recently, an unspecified request in December 1989 for blood and plasma for Romania generated US$ 5 million in donations from several European countries by mid-December. No agency considered the need for an assessment of the real requirements. At the end of December, products worth US$ 4 million remained unused. Similarly, 20% of the medication sent to Armenia had passed its expiry dates on arrival, was soon to expire or was of no use in an emergency (9).

Information needs and assessment procedures in disasters

Information needs and assessment methods in disasters depend largely on the type of disaster and the time at which the assessment is made. Natural disasters which have a significant impact on humans may be classified into two categories for this purpose:

(i) rapid onset (little or no advance warning): earthquakes, flash floods, cyclones, high winds;
(ii) slow onset (at least some advance warning): floods, famines, epidemics, civil strife, refugees, displaced persons.

Information needs may also be categorized into phases corresponding to the evolution of the process:

- baseline information phase;
- postimpact information phase:
  - immediate relief information
  - secondary relief information;
- rehabilitation information phase;
- evaluation information phase.

The time phase defines the detail and scope of the assessment.

The need for information and the choice of assessment techniques are determined by the rapidity with which the results are required and the time phase of the disaster process in which the assessment is being done. Ideally, baseline information should reflect normal circumstances, i.e. the predisaster period. This body of information should include those items which although labour-intensive, time-consuming or painstaking to collect provide essential input to needs assessment and the planning of relief operations. Examples of such information are: demographic characteristics, agricultural or meteorological data and prevalence data of diseases that could directly or indirectly be affected by the disaster (10). Often these data are not available.

A needs assessment undertaken in the immediate postimpact period will focus on life-line needs and prevention of impact-related mortality. In the secondary phase, the assessment will address longer-term shelter, food and health care. Subsequently, the development of sentinel surveillance systems and immunization programmes will also be added to the aims of the assessment.

The success of an emergency assessment can be greatly enhanced by the availability of baseline information. For example, identification of nutritional priority groups in postflood conditions requires some knowledge of the normal nutritional status of the population.

Similarly, the assessment of increased incidence of malaria or diarrhoeal disease after cyclones would be facilitated if routine surveillance data were available (Figs 3 & 4). Such baseline data, together with population characteristics, would enable the assessors to estimate the epicurve and propose appropriate action. Therefore, areas which are prone to recurrent emergency situations need to be encouraged to develop sentinel surveillance systems which will allow the required baseline information to be collected without overburdening health workers.

The endemic disease profile of the affected community, its predisese health status and infrastructure will determine to a large extent the content and methodology of an assessment. The predisaster nutritional situation and cropping pattern will determine actual and potential food needs. Loss of harvest (actual or potential), salination of arable soil, loss of tools of trade could all eventually lead to a severe food crisis in poor, agricultural populations. Emergency assessment would have to evaluate these risks and propose provisions against these eventualities.

The range of what can be achieved by rapid assessment in emergencies is best appreciated by examining the type of damages they create. Box 1 shows the common effects on environmental health services caused by the four most frequent natural disasters.

Rapid assessment in emergencies: current practices

The rationale for rapid assessment is the need for a rapid response. The inevitable loss in accuracy, completeness and reliability in rapid methodologies can only be justified for this reason. In most natural disasters, the time frame for assessing immediate health needs is a matter of days. Some, such as earthquakes or cyclones, require the most immediate assessment within 24-48 hours. This stringent time constraint has been justified by studies such as those by De Bruycker et al. for Italy (4) and Glass et al. for Guatemala (6), in which the majority of the deaths due to the earthquake occurred within the first 48 hours (Fig. 5).

Purposes and common use of data collected by rapid assessment in emergencies

The purposes of undertaking a rapid assessment in any emergency situation are:

- determination of the magnitude of the disaster;
- measurement of present and potential impact;
- assessment of resources needed, including local response capacity;
- planning of appropriate response.

Ideally, as described above, an initial survey is first undertaken for immediate needs. The time required for an initial assessment varies with the type of dis-
Among the different emergency situations, two specific areas have been relatively well examined. These are rapid health assessment among refugees and rapid nutritional assessment. Assessment techniques for acute natural disasters still remain fragmentary and ad hoc. Response to epidemics tends to be extreme, involving either launching a full-scale epidemiological investigation, at the termination of which the epidemic has burnt out, or providing unsuitable supplies which happen to be available at the time.

Rapid assessment in food-crisis situations

Rapid assessment of nutritional status has been relatively well studied by many researchers. In particular, Trowbridge et al. (11) and more recently, Manley et al. (12) have published methodologies particularly adapted to rapid assessment. In the
Box 1. Common effects of natural disasters on environmental health services

<table>
<thead>
<tr>
<th>Service</th>
<th>Most common effects</th>
<th>Earthquake</th>
<th>Hurricane/tornado</th>
<th>Flood</th>
<th>Tsunamis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Water supply and wastewater disposal</td>
<td>Damage to civil engineering structures</td>
<td>+++</td>
<td>+++</td>
<td>+++</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>Broken mains</td>
<td>+++</td>
<td>++</td>
<td>++</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>Power outages</td>
<td>+++</td>
<td>+++</td>
<td>++</td>
<td>++</td>
</tr>
<tr>
<td></td>
<td>Contamination (biological or chemical)</td>
<td>++</td>
<td>+++</td>
<td>+++</td>
<td>+++</td>
</tr>
<tr>
<td></td>
<td>Transportation failures</td>
<td>+++</td>
<td>+++</td>
<td>+++</td>
<td>+++</td>
</tr>
<tr>
<td></td>
<td>Personnel shortages</td>
<td>+++</td>
<td>++</td>
<td>++</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>System overloading (due to shifts in population)</td>
<td>+</td>
<td>+++</td>
<td>+++</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>Equipment, parts and supply shortages</td>
<td>+++</td>
<td>+++</td>
<td>+++</td>
<td>+++</td>
</tr>
<tr>
<td>Solid waste handling</td>
<td>Damage to civil engineering structures</td>
<td>+++</td>
<td>++</td>
<td>++</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>Transportation failures</td>
<td>+++</td>
<td>+++</td>
<td>+++</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>Equipment shortages</td>
<td>+++</td>
<td>+++</td>
<td>+++</td>
<td>+++</td>
</tr>
<tr>
<td></td>
<td>Personnel shortages</td>
<td>+++</td>
<td>+++</td>
<td>+++</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>Water, soil and air pollution</td>
<td>+++</td>
<td>+++</td>
<td>+++</td>
<td>+</td>
</tr>
<tr>
<td>Food handling</td>
<td>Damage to food preparation facilities</td>
<td>+++</td>
<td>+++</td>
<td>++</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>Transportation failures</td>
<td>+++</td>
<td>+++</td>
<td>+++</td>
<td>++</td>
</tr>
<tr>
<td></td>
<td>Power outages</td>
<td>+++</td>
<td>+++</td>
<td>+++</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>Flooding of facilities</td>
<td>+</td>
<td>+++</td>
<td>+++</td>
<td>+++</td>
</tr>
<tr>
<td></td>
<td>Contamination/degradation of relief supplies</td>
<td>++</td>
<td>+++</td>
<td>+++</td>
<td>+</td>
</tr>
<tr>
<td>Vector control</td>
<td>Proliferation of vector breeding sites</td>
<td>+++</td>
<td>+++</td>
<td>+++</td>
<td>+++</td>
</tr>
<tr>
<td></td>
<td>Increase in human-vector contacts</td>
<td>+++</td>
<td>+++</td>
<td>+++</td>
<td>+++</td>
</tr>
<tr>
<td></td>
<td>Disruption of vector-borne disease control programmes</td>
<td>+++</td>
<td>+++</td>
<td>+++</td>
<td>+++</td>
</tr>
<tr>
<td>Home sanitation</td>
<td>Destruction or damage to structures</td>
<td>+++</td>
<td>+++</td>
<td>+++</td>
<td>+++</td>
</tr>
<tr>
<td></td>
<td>Contamination of water and food</td>
<td>++</td>
<td>+++</td>
<td>+++</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>Disruption of power, heating, fuel, water supply or waste-disposal services</td>
<td>+++</td>
<td>+++</td>
<td>+++</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>Overcrowding</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
</tbody>
</table>

+++ severe possible effect; ++ less severe possible effect; + least or no possible effect.
Source: Reference (21).

**FIG. 5**

SURVIVAL CURVE OF VICTIMS IN THE 1980 EARTHQUAKE (CAMPANIA, ITALY)

COURSE DE SURVIE POUR LES VICTIMES DU TREMBLEMENT DE TERRE DE 1980 (CAMPANIE, ITALY)

 Accumulative number of trapped people extricated - Nombre cumulatif de personnes extraites des décombres.
Accumulative number of people extricated alive - Nombre cumulatif de personnes extraites vivantes.

\[
\text{Time (Lib}_{50}) = 8 \text{ h.} \quad \text{- Temps (Lib}_{50}) = 8 \text{ h.}
\]

N = 548

Rapp. trimest. statist. sanit. mond., 44 (1991)
majority of cases, rapid assessment of nutritional status is required for drought or other severe food-shortage situations. The aims of assessment in these cases are to: (i) confirm that a nutritional emergency exists and estimate the number of those severely affected; (ii) assess how severe the food crisis is likely to become in the short and medium term; (iii) identify the groups most affected and the risk factors that could potentially worsen the status; and (iv) assess the need for a more detailed evaluation.

In nutritional crises or disasters resulting in food shortages, anthropometric indicators, principally weight-for-height, height-for-age, weight-for-age and mid-upper-arm circumference, have been used and tested for their potential. The field inadequacies of age-based measurements have been discussed by Bairagi et al. (13), Chen et al. (14) and others. Bairagi estimates the extent of bias in age misstatement to be serious enough to invalidate many results.

Trowbridge & Staehling (15) have examined the sensitivity and specificity of these indicators vis-à-vis their cut-off points. There is general agreement that arm circumference (at 125 mm) provides a quick and sensitive assessment of undernourished children aged 12-60 months. However, the robustness of this indicator with regard to the cut-off point is low. Small shifts in the threshold generate major changes in false positives and false negatives, making the tool questionable in field conditions.

**Rapid health assessment among displaced persons**

Displaced persons include both refugees and internally displaced persons. The aims of a rapid assessment in displacement situations are to: (i) assess the magnitude of the displacement; (ii) assess the major health and nutrition needs of the displaced population; (iii) initiate a health and nutritional surveillance system; and (iv) assess local resource capacity and immediate needs.

Displaced populations, especially those fleeing from drought, have the most wide-ranging and pervasive health implications. Drought-related displacement usually involves groups severely undernourished and physically exhausted. The environmental and social conditions in which they find themselves following displacement are particularly deficient and chaotic. Health-needs assessment in these situations is extremely difficult and the ranking of priorities doubtful.

As far as rapid assessment is concerned the cause of displacement is not entirely irrelevant. The risk of certain diseases (for example malaria) is higher if the community is moving from a nonendemic to an endemic region. Glass (6) reports on the elevated incidence of malaria morbidity and mortality among refugees arriving at the Sa Keo camp at the border between Thailand and Cambodia. The arrival of these refugees from a malaria-free zone into an endemic area resulted in a health situation that was highly vulnerable to outbreaks of the disease. If the displacement is motivated by drought, it is likely that the population will be seriously malnourished and the demographic profile distorted in favour of families without males or a disproportionate number of very old and very young. This is because, as the food situation deteriorates over time, the adults leave home to search for work or food and the families then move on their own, when the village decides to move. Civil war or returnees, such as the ones seen on the frontline states of Southern Africa (e.g. Botswana, Mozambique, Zambia and Zimbabwe), are frequently in good health and nutritional condition when they arrive, but if health services are not quickly organized, their status can deteriorate rapidly. These background conditions can determine the focus of an assessment mission and the types of methods to use.

Since refugee health problems have the dubious distinction of encompassing all aspects of a normal health structure with some additional specificities, assessment in these situations is essentially a telescoped version of the planning of a health programme in five days. The worst-case scenarios are those already observed in the 1983-1984 droughts of Ethiopia, Sudan and other Sahel countries. The emergency consisted of the sudden arrival of 20,000-40,000 persons in an advanced state of destitution, with high rates of disease and malnutrition and no apparent means of survival. Rapidity of the health-needs assessment in these conditions takes on a different significance and the methods or techniques are selected according to different criteria.

**Rapid assessment in acute natural disasters**

Rapid assessment in this type of disaster tends to focus on mortality and expected morbidity estimations. The types and quantities of injuries and subsequently the need for food and disease control, especially in floods or cyclones, typify the assessment procedures. Earthquakes do not generally produce a situation in which the health of the surviving population is severely affected, although the death rate is high and fast, and survival curves level off within 24-48 hours of the impact. Most deaths occur in the first 6-10 hours and those surviving are generally unaffected. Injuries and trauma among survivors are relatively limited and are also concentrated within the first couple of days. The survivors among the affected population are healthy and require shelter, food and water.

Floods and cyclones have greater health implication. Besides mortality and morbidity from immediate impact, waterborne diseases, respiratory tract infections and in the longer term, decreases in nutritional status have been observed (16,17).

Overall, the main public health concern following disaster has been the fear of disease as a secondary consequence to the acute natural disaster. This has not been observed to occur very frequently. There are, however, some epidemiological determinants that influence the risk of an outbreak after an earthquake, flood or cyclone. These are:

- the endemic levels of disease in the community;
- ecological change;
- population displacement;
- population density;
- interruption in health services;
- disruption of sanitary facilities.

---

*Guha-Sapir, D. Nutritional and non-nutritional impact of supplemental feeding programmes. A study in India. Department of Epidemiology and Preventive Medicine, University of Louvain, 1990. (Doctoral thesis).*

Rapid assessment in epidemics and disease outbreaks

The principal aims of a rapid assessment in epidemics are to: (i) confirm that an epidemic exists or is threatening; (ii) estimate its geographical distribution; (iii) estimate its health impact and; (iv) identify local capacity to control transmission and reduce mortality. The usual approach in epidemics has been to carry out a classical epidemiological investigation which normally takes 2-3 weeks to complete. In fact, if an epidemic is already in place, as it usually is before any emergency action is taken, the following information may be sufficient for an assessment. Firstly, the establishment of a working case-definition is essential for any further field investigation and case identification. Secondly, the geographical distribution of the epidemic should be defined not only from hospitals, but also from new graves and a review of death certificates. Improved recognition and better reporting of disease due to rumours of an epidemic can artificially increase the number of cases when there is no real increase. Thirdly, the mode of transmission of the disease which will determine the control mechanisms should be established. An investigation of the type and source of infection, severity of illness and prognosis will also be required for an initial assessment. The data on the sample of cases should ideally be mapped by community groups and geographical regions to filter out high-risk bias or groups. Assessment methods in epidemic situations will use survey (hospital- or population-based depending on the expected prevalence of the disease) and laboratory methods. Time-series data can be important in this context for identifying the location on an epicurve.

Limitations and weaknesses in present methods and approaches

Currently, the methods used for rapid assessment are variations of those used for normal assessment or epidemiological investigations. The variations are generally ad hoc, depending entirely on the assessor's individual capacities, the time frame and the donor mood. There is little use of standardized methods. Information on assessment techniques used to arrive at the final conclusions and the biases in the methods have rarely been provided. However, experience has shown that among the techniques used, there are some sources of error and bias that are revealed repeatedly in emergency health assessment.

An initial assessment often focuses on what are thought to be the most affected areas, rather than on an overview of the entire disaster area. There are two main problems with this approach. Firstly, the areas identified as most affected may not, in fact, be the worst-off. Frequently, the selection of sites is the worst affected results in a sample biased by the source of information. Secondly, by selecting the hardest hit areas only, it may be difficult to assess the full impact of the disaster. The source of information on morbidity may also severely influence morbidity estimates. Information from health providers may not be accurate or representative. On the one hand, injuries may be underreported because of poor record keeping or because health facilities may be inaccessible to many of the injured. On the other hand, injuries may be overreported because the same injuries are registered or counted several times.

The problems encountered in obtaining information on injuries can be illustrated by the following experience of needs assessment after an earthquake in Tangshan (China) in 1976. It was suggested that the low levels of head and chest injuries recorded in the hospital registers were mainly due to the fact that persons with this type of injury died on impact or soon thereafter. Thus, the morbidity profile of the earthquake, based only on those surviving several days in the hospital, revealed artificially low rates of trauma to the upper torso. When compared to other more complete injury profiles reported by Beinin (17), it is clear that head and chest injuries are the predominant categories of immediate health impact in earthquakes.

Other errors in the immediate impact phase arise from partial information on mortality. An important consideration in using mortality data is that it is not as useful for immediate relief as the evolving injury pattern. However, for assessing future need priorities, it is useful to determine leading causes of death and associated risk factors in specific types of disasters. In rapid-onset disasters, it is particularly difficult to estimate the number of bodies that have not been recovered. For this reason, reported mortality is often limited to the number of bodies recovered, thus underestimating the true mortality. The differentiation between mortality estimates based on body counts and those which include persons missing is critical for any population-based estimate. In addition, while it is easier to attribute the event as the cause of death in acute, rapid-onset disasters such as earthquakes and cyclones, this is more problematic in slow-onset disasters such as famines or floods. The problem occurring as an indirect effect of the disaster, such as deaths due to diseases aggravated by malnutrition, may be attributed to the drought by some definition and not by others. This definition of attributed cause of mortality or morbidity has serious implications for how need is assessed and response planned. In slow-onset disasters, the problem of excluding that portion of morbidity and mortality that would have occurred in the normal course of events is also significant.

Bias can be a serious weakness in rapid assessment made during emergencies. Often, sample sizes are too small to be representative of the population or survey, and a purposive sample rather than a random sample must be used. Sample representativeness can be increased in several ways. Besides avoiding markets, centres of towns or main roads, the range of the health impact can be estimated by selecting the most-affected and least-affected villages (accordance to a local authority). This can be further refined by asking different individuals, such as a religious leader, a local political figure, government official, non-governmental organization, missionary, for information on worst affected and least affected. If there is no clear consistency in the replies, several worst-affected areas should be included in the survey. Within the village, random starting points should be used and contiguous houses avoided. Urban centres should be treated differently from rural centres, since living patterns are different. Some marginal groups should be specially sampled, such as nomadic settlements, slums, or forest and mountain people.

There are other sources of bias leading to low representativeness. The timing of surveys, for example, can produce a misleading sample. Surveys done at certain times of the day will oversample women or...
In recent years, a few epidemiological studies have been published that have contributed significantly to the development of the rapid assessment techniques proposed by WHO, and have furthered the cause of robust techniques for use in the field. Survey techniques have been proposed based on experience, particularly in sudden, massive population-displacement situations. Epidemiological and other indicators that are easy and quick to measure have also been examined for their potential in reflecting health needs (18-20). However, much of this knowledge today is based on occasional field experience, and relatively little systematic and scientifically-controlled research has been undertaken to validate the methods. The typical conditions of time and resource constraints under which rapid assessments take place in emergencies make sound, field-tested methods critical.

In emergencies, particularly in sudden, massive population-displacement situations, the consequences of errors are very serious. Inadequate assessment in a rapidly-evolving situation such as a flood or famine implies significant numbers of lives lost and serious long-term health consequences.Since disasters in developing countries generally affect large groups of population, the individualized approach to emergencies as practised in Western countries is inappropriate.

An important methodological issue in surveys for rapid assessment in the epidemiological context is the need for a reliable denominator. This is evidently easier said than obtained, especially in rapidly-evolving situations. However, the significance of a valid denominator should not be underestimated. The population affected is essential to any rapid assessment initiative. Creative methods such as counting the number of houses without roofs by an aerial survey in an earthquake and multiplying it by the average family size of the country would already provide a working figure for the total number of homeless. Standard coefficients (for example, 0.15 for proportion of children <5) would provide further estimates for infants and small children for feeding and immunization planning.

In camp conditions, a method for estimating deaths prior to having a proper registration system established has been used by Toole & Waldman (18). It consists of posting a guard around the clock at the burial place to note all bodies brought in by relatives. Alternatively, in a camp where no burial facilities were available and the dead were removed by an external contractor, the epidemiologists responsible for assessment paid the contractor a flat fee to report the numbers and some basic details on the bodies he transported. Mortality assessment can be very problematic in many emergency situations, since food rations or other relief goods are supplied on a per capita basis. This naturally discourages families from reporting deaths for fear of losing part of their share.

Finally the choice of sampling methods for the assessment (household- or institution-based) will depend on the expected prevalence of the disease or phenomenon in question. At a rate of 5 cases per 1000 at the peak of an epidemic, as expected in epidemic meningococcal meningitis, a household survey is meaningless. For yellow fever, on the other hand, a household survey is indicated, but since the aim is to cut off transmission and improve protection rather than prevent mortality, much effort for an accurate assessment of the number of cases may be of little practical value for mounting an immediate emergency response. The severity of a disease and therefore its duration can also lead to misleading conclusions in rapid assessment missions. Any example of such misjudgement is provided by the rapid assessment mission sent to Chad in 1975, during the height of the famine. The report of the team of experts concluded that no serious malnutrition existed and there was no cause for alarm. It was reported later, after examination of data over a longer period than was considered by the rapid assessment team, that the children most severely malnourished were dying very quickly at the peak of the famine. Therefore only the survivors (moderately or mildly malnourished) were available to be surveyed by the team. This phenomenon of a low-point prevalence of a disease as duration of illness decreases is a pitfall in rapid assessments that could seriously invalidate the results. Sampling methods and survey content should be modified according to the health problem in question and the phase of the emergency (for example, when the rapid assessment is taking place in the epicurve of a meningitis epidemic).

Limited surveys, if properly done, may provide a rough idea of the extent of damage, prevalences and incidences of malnutrition and diseases. The larger and better-designed the survey, the more reliable the results. But in emergencies there is a trade-off to be made between accuracy and timeliness, in addition to accounting for resource and logistical constraints. The most practical view in these circumstances is that being roughly right is generally more useful than being precisely wrong. The delay in reporting the assessment conclusions generally means that relief response will have been initiated without any consideration of the actual needs. This does not by any means imply that spurious methodologies or amateurism can replace rigorous thinking. The use of indicators that do not reflect the phenomena in question or surveying samples that mislead the assessor in the conclusions can create more damage than not taking any action. The challenge is to modify regular methods to fit the constraints of the situation.

Acknowledgement
The author wishes to thank Ms Martha Anker for her constructive and useful comments on previous drafts.

Wid hith statist. quart., 44 (1991)
The increase in the number of natural disasters and their impact on population is of growing concern to countries at risk and agencies involved in health and humanitarian action. The numbers of persons killed or disabled as a result of earthquakes, cyclones, floods and famines have reached record levels in the last decade. Population density, rampant urbanization and climatic changes have brought about risk patterns that are exposing larger and larger sections of populations in developing countries to life-threatening natural disasters. Despite substantial spending on emergency relief, the approaches to relief remain largely ad hoc and amateurish, resulting generally in inappropriate and/or delayed action. In recent years, mass emergencies of the kind experienced in Bangladesh or the Sahelian countries have highlighted the importance of rapid assessment of health needs for better allocation of resources and relief management. As a result, the development of techniques for rapid assessment of health needs has been identified as a priority for effective emergency action.

This article sketches the health context of disasters in terms of mortality and morbidity patterns; it describes initial assessment techniques currently used and their methodological biases and constraints; it also discusses assessment needs which vary between different types of disasters and the time frame within which assessments are undertaken. Earthquakes, cyclones, famines, epidemics or refugees all have specific risk profiles and emergency conditions which differ for each situation. Vulnerability to mortality changes according to age and occupation, for earthquakes and famines. These risk factors then have significant implications for the design of rapid assessment protocols and checklists.

Experiences from the field in rapid survey techniques and estimation of death rates are discussed, with emphasis on the need for a reliable denominator even for the roughest assessment. Finally, the importance of adapting normal epidemiological and statistical methodologies to crisis situations is underlined in order to rationalize the recurrent and substantial expenditures made in response to natural disasters today.

Résumé

Evaluation rapide des besoins sanitaires en situation d'urgence de grande ampleur: méthodes et concepts actuels

La fréquence accrue des catastrophes naturelles et la portée de leurs effets sur les populations préoccupent de plus en plus les pays exposés et les organismes à vocation sanitaire et humanitaire. Tremblements de terre, cyclones, inondations et famines ont fait un nombre record de tués et de handicapés ces 10 dernières années. Les schémas de risque dont s'assortissent une forte densité de population, une urbanisation galopante et les changements climatiques mettent en danger la vie de sections de plus en plus importantes de la population des pays en développement lors d'une catastrophe naturelle. Malgré les sommes considérables dépensées pour les secours d'urgence, la façon d'apprécier les secours reste largement ponctuelle et empirique. Les mesures adoptées se sont souvent révélées inadéquates ou ont été prises tardivement, voire les deux à la fois. Les situations d'urgence de grande ampleur survenues récemment, comme celles du Bangladesh et des pays du Sahel, ont fait ressortir la nécessité d'une évaluation rapide des besoins sanitaires pour améliorer l'allocation des ressources et la gestion des secours. C'est ainsi que l'élaboration de techniques d'évaluation rapide des besoins sanitaires a été reconnue comme une priorité pour l'efficacité des secours.

Cet article donne un aperçu du contexte sanitaire des catastrophes sous l'angle des schémas de morbidité et de mortalité. Il décrit les techniques d'évaluation initiale en usage ainsi que les biais et obstacles méthodologiques. Il examine les besoins en évaluation, qui varient selon le type de catastrophe et le moment de l'évaluation. Les profils de risque diffèrent selon qu'il s'agit d'un tremblement de terre, d'un cyclone, d'une famine, d'une épidémie ou de populations réfugiées, et les conditions d'urgence varient dans chaque situation. En cas de tremblement de terre ou de famine, le risque de mortalité est fonction de l'âge et de la profession. Ces facteurs de risque modifieront profondément la conception des protocoles d'évaluation rapide et des listes de contrôle.

Divers exemples concrets de techniques d'enquête rapide et d'évaluation des taux de mortalité sont examinés et l'auteur insiste sur la nécessité d'un dénominateur fiable, même pour les évaluations les plus approximatives. Elle souligne enfin qu'il importe d'adapter les méthodes épidémiologiques et statistiques normales aux situations de crise pour rationaliser les dépenses fréquentes et considérables qu'entraînent aujourd'hui les catastrophes naturelles.

Références


Maternal Mortality
A Global Factbook

Compiled by C. AbouZahr and E. Royston
1991, 608 pages (English only)
ISBN 92 4 159001 7
Sw.fr. 50.-/US $45.00
In developing countries: Sw.fr. 35.-
Order no. 1930024

This book sets out the facts and figures needed to understand why so many women continue to die as a result of pregnancy and childbirth despite the fact that the technical means to prevent such deaths have long been available. Drawing upon a vast data base of some 3,000 reports and studies, the factbook shows, in the form of country profiles, where women are dying, what they are dying of and what other aspects of their lives contribute to their deaths. Noting that maternal death is most often the tragic end to a life-long chain of events and disadvantages, the book tracks down the underlying factors, often rooted in sex discrimination present since infancy, as well as the more immediate factors, such as lack of access to lifesaving care, that reveal the true complexity of the forces at work. Information such as that contained in this factbook provides the key for effective action, making the best use of limited resources despite the often difficult circumstances.

The main body of the factbook, which runs to some 600 pages, consists of country profiles which, for the first time ever, bring together and analyse the results of all available surveys and studies on maternal mortality, women's reproductive health and allied subjects, as well as indicators of the coverage of maternity care, family planning and other background factors. Profiles are given for each of 117 developing countries in Africa, Latin America, Asia and Oceania. Data on developed countries are also tabulated for comparison. In compiling the profiles the authors have drawn upon the unique WHO women's health data base which, in addition to the more readily available government reports and articles from scientific journals, contains information from a large variety of disparate sources, including unpublished articles, doctoral theses and consultant briefings.

To make it easier to compare countries, each profile follows a common format, starting with a section containing demographic and socioeconomic indicators that shed light on women's lives in each country: their chances of going to school, eating well, and receiving health care, the age at which they are likely to marry, their chances of planning their families, and the number of children they are likely to bear. These data provide a backdrop for the detailed statistics on coverage of care and maternal mortality which follow, and which detail the numbers of deaths, the mortality rates and ratios, the causes and circumstances surrounding each case, the groups of women most at risk of dying, and the kinds of preventive and curative actions that might have averted death. Each country profile also includes an annotated bibliography, a list of further reading, and lists of information sources.

The interpretation of this vast amount of information is facilitated through the inclusion of four background chapters. The first provides an overview of the dimensions and causes of maternal mortality and morbidity in the world today as well as of the extent of the coverage of care. The different ways of measuring maternal mortality are described in the second chapter, which discusses the strengths and weaknesses of each method. The third explains how the results of surveys should be interpreted and analyses the information that can, or cannot, be obtained from hospital studies, community surveys or registration data. The book also features a comprehensive listing of general resource materials for readers who wish to expand their knowledge on this complex issue.

Throughout the book, the picture that emerges is one of deaths that could have been prevented if women in need had access to good prenatal monitoring and adequate maternity care. In showing where and why women die, the book also gives cause for hope. With the facts now available, policy makers, health professionals and others striving to improve women's health will be better equipped to attack the underlying and immediate problems on a broadened range of fronts.
Manual of Epidemiology for District Health Management

edited by J.P. Vaughan and R.H. Morrow
1989, vii + 198 pages (available in English and French)
ISBN 92 4 154404 X
Sw.fr. 35.–/US $31.50
Order no. 1150335

This book provides a simple, practical, step-by-step guide to the use of epidemiology as a tool for improving the management of health services. Addressed to general health workers, the book uses clear definitions, analogies, examples, checklists, sample forms and calculations, and abundant illustrations to demystify the methods of epidemiology and show how they can work in concrete situations. Particular emphasis is placed on the simple knowledge and skills needed to collect and then use epidemiological data to monitor health problems commonly found in developing countries.

The book has 14 chapters. Readers are first introduced to the main tasks involved in the management of district health services and the types of information that can contribute to more effective management. The second chapter shows how a four-phase epidemiological approach, involving descriptive, analytical, intervention, and evaluation epidemiology, can supply virtually all the information needed to pinpoint health problems, design targeted interventions, and define reliable indicators for monitoring progress. Basic definitions of incidence versus prevalence, of numbers versus rates, and of episodes versus attendances are also set out in an effort to simplify the concepts of epidemiology and prevent common errors in the design or interpretation of studies. Other chapters offer guidance in the collection of demographic data, the conduct of routine health surveillance, the use of epidemiology to control an epidemic, and the design of special surveys to collect additional information.

The second half of the book concentrates on the analysis, presentation, and use of results. Topics covered include the use of record forms and coding, methods of data processing and analysis, and the presentation of health information in tables, figures, graphs, diagrams, charts, and maps. The final chapter, which constitutes the core of the manual, shows how the knowledge and skills previously described can be used to formulate plans for the management and monitoring of district health services.

"... The methods of epidemiology are applied clearly to the analysis of the health of the district and to plans for its health care ... Admirable practical checklists are given and basic epidemiological techniques are precisely described ... deserves to be a standard text in the district..."
— The Lancet

"... logical in sequence and eminently readable ... should be on every district general manager's and management trainee's bookshelf and referred to frequently in order to make informed judgements on health needs..."
— Nursing Times

"... a good overview of epidemiological methods that are relevant to district health systems..."
— American Journal of Epidemiology

"... a very practically based introduction to epidemiology in the context of planning and management ... certain to be of immense value to local health workers wanting an introduction to the area ... the ideas and advice presented are universally valid...
— Health Services Management Journal
Cet ouvrage répond au besoin d'un guide simple, pratique, progressif sur l'utilisation de l'épidémiologie en tant qu'outil destiné à améliorer la gestion des services de santé. À l'intention des agents de santé généraux, ce livre utilise des définitions claires, des analogies, des exemples, des listes de contrôle, des formules et des calculs d'échantillons ainsi que d'abondantes illustrations pour démystifier les méthodes épidémiologiques et montrer comment elles peuvent s'appliquer dans des situations concrètes. L'accent porte sur les connaissances et les compétences simples requises pour recueillir puis utiliser les données épidémiologiques permettant de surveiller les problèmes de santé constants dans les pays en développement.

Le livre se compose de 14 chapitres. Sont d'abord exposées les principales tâches liées à la gestion des services de santé de district et les informations propres à accroître l'efficacité de la gestion. Le deuxième chapitre montre comment une démarche épidémiologique en quatre phases - épidémiologie descriptive, épidémiologie analytique, intervention et évaluation - peut fournir la quasi-totalité des informations requises pour recenser les problèmes de santé, concevoir des interventions ciblées et définir des indicateurs fiables pour suivre les progrès accomplis. Des définitions de base - incidence et prévalence, effectifs et taux, épisodes et consultations - sont aussi données dans le but de simplifier les concepts de l'épidémiologie et de prévenir les erreurs courantes dans la conception et l'interprétation des études. Les autres chapitres donnent des conseils sur la façon de recueillir des données démographiques, de mener des activités de surveillance sanitaire systématiques, d'utiliser l'épidémiologie pour lutter contre une épidémie et de concevoir les enquêtes spécialement destinées à recueillir des informations supplémentaires. Les détails vont d'une formule permettant d'évaluer le taux brut de natalité, en passant par des exemples de diagnostics utilisés pour la définition des cas, à des avis sur l'utilisation des échantillons en grappes et la détermination de la taille des échantillons. L'attention des lecteurs est aussi appelée sur les problèmes liés à l'utilisation des questionnaires et sur le devoir absolu de confidentialité dans le déroulement des enquêtes.

La deuxième partie du manuel est consacrée à l'analyse, à la présentation et à l'utilisation des résultats. Au nombre des sujets couverts figurent l'utilisation des formulaires d'enquête et de codage, les méthodes de traitement et d'analyse des données et la présentation des informations sanitaires en tableaux et figures, graphiques, histogrammes de fréquences, diagrammes en barres, diagrammes en secteurs, diagrammes de dispersion et cartes. Des directives pour la rédaction des rapports sanitaires, y compris un cadre modèle, sont aussi données. Le dernier chapitre, élément central du manuel, montre comment utiliser les connaissances et compétences précédemment décrites pour planifier la gestion et la surveillance des services de santé de district.

L'ouvrage s'achève sur une série de six annexes contenant des détails supplémentaires sur les méthodes exposées dans le manuel.

«... Les méthodes épidémiologiques sont appliquées de manière appropriée à l'analyse de la santé au niveau du district et aux plans de soins de santé... Des listes de contrôle extrêmement commodes y sont proposées et l'on y décrit avec précision les techniques épidémiologiques de base... Il mérite de devenir l'ouvrage de référence du district... »

The Lancet

---

**BON DE COMMANDE**

- Veuillez m'envoyer ______ exemplaire de *Manuel d'épidémiologie pour la gestion de la santé au niveau du district* au prix de Fr.s. 35.--- l'exemplaire (N° de commande 2150335)

Nom

Adresse

---

- chèque joint
- Veuillez débiter ma carte de crédit comme suit
  - Visa
  - American Express
  - Eurocard/Mastercard/Access

N° de carte

Date d'expiration

Signature
EPIDEMIOLOGICAL AND STATISTICAL METHODS FOR RAPID HEALTH ASSESSMENT

The use of rapid assessment techniques has increased markedly in recent years. Many of these techniques are new, others are innovative adaptations of more traditional methods. Rapid assessment methods aim to provide information quickly, at low cost, in a simple format, without detriment to accuracy. Emphasis is placed on obtaining information that can be used at the local level.

This issue of the Quarterly reviews a number of statistical and epidemiological methods of rapid assessment from the point of view of their strengths and weaknesses, their applicability to field settings, their validity and reliability, and their potential for application to new situations.

Both quantitative and qualitative methods are covered. Specific articles deal with: a simplified general method for cluster-sample surveys in developing countries; an evaluation of the Expanded Programme on Immunization survey methodology for estimating relative risk; lot quality assurance sampling - both the methodology behind it, and its application to field settings; the potential of case-control methodology for rapid epidemiological assessment; the value of questionnaires aimed at key informants and distributed through an existing administrative system, for rapid cost-effective health assessment; the application of geographical information systems in public and environmental health; and rapid assessment of health needs in mass emergencies.

MÉTHODES ÉPIDÉMIOLIGIQUES ET STATISTIQUES POUR UNE ÉVALUATION RAPIDE DE LA SITUATION SANITAIRE

L'utilisation de techniques d'évaluation rapide s'est sensiblement accrue ces dernières années. Nombre d'entre elles sont nouvelles, d'autres sont des adaptations novatrices de méthodes éprouvées. Les méthodes d'évaluation rapide visent à fournir des informations sans délai, dans des conditions économiques et sous une forme simple, sans que l'exactitude ait à en souffrir. L'accent est mis sur la collecte d'informations pouvant être utilisées au niveau local.

Le présent numéro du Trimestriel examine un certain nombre de méthodes statistiques et épidémiologiques d'évaluation rapide pour déterminer leurs avantages et inconvénients, leur applicabilité sur le terrain, leur validité et leur fiabilité et enfin, leurs possibilités d'adaptation à des situations nouvelles.

Les articles, qui passent en revue des méthodes tant quantitatives que qualitatives, portent sur les questions suivantes: méthodes générales simplifiées pour les enquêtes sanitaires utilisant le sondage par grappes dans les pays en développement; évaluation de la méthodologie d'enquête du Programme élargi de vaccination pour l'estimation du risque relatif; l'échantillonnage par lots pour l'assurance de la qualité - sa méthodologie et ses applications sur le terrain; le potentiel de la méthode cas-témoin pour une évaluation épidémiologique rapide; l'utilisation des questionnaires destinés aux répondants clés et distribués dans le cadre d'un système administratif existant en vue d'une appréciation rapide et rentable de la situation sanitaire; l'application de systèmes d'information géographique dans les domaines de la santé publique et de l'hygiène de l'environnement; et l'évaluation rapide des besoins sanitaires dans les situations d'urgence de grande ampleur.