

SURVEYS TO MEASURE PROGRAMME COVERAGE AND IMPACT: A REVIEW OF THE METHODOLOGY USED BY THE EXPANDED PROGRAMME ON IMMUNIZATION

Stanley Lemeshow^a & David Robinson^b

Introduction

In order to improve the health status of their populations, most countries are developing their ability to provide primary health care. This relies upon a capacity to manage the national health system which, in turn, is dependent upon information for purposes of planning, supervision and monitoring of health activities. Data are required to define the need for health services, the efficiency of existing services, as well as their impact on morbidity and mortality.

Although much of the necessary information can be obtained from routine sources, some can best be obtained through the use of surveys. The World Health Organization (WHO) and other international agencies have been active in promoting the use of such surveys. As an example, through its Expanded Programme on Immunization (EPI), WHO aims to ensure the availability of immunization for all the children of the world by the year 1990. This effort is considered a vital step towards the attainment of WHO's stated goal of health for all by the year 2000.

The impact of diseases such as neonatal tetanus, poliomyelitis, measles, whooping cough, diphtheria and tuberculosis upon children in the developing world is compounded by the fact that many of these children are severely weakened by malnutrition and repeated episodes of diarrhoea and malaria. Immunization programmes are seen as an important measure for reducing infant and childhood disability and deaths, which act as spurs to sustain high birth rates.

A major need of the EPI has been an appropriate system for gathering information. Reliable data are necessary to document the level of morbidity and mortality from specific target diseases as well as the level of immunization coverage against these diseases. Since this information is not readily available to health managers in many developing countries, EPI has worked to develop a method which could obtain accurate information quickly and cheaply. A method was sought which, in addition, could be implemented in a relatively standardized manner from one country to the next. This would permit training materials and operation manuals to be developed for widespread use, and would facilitate comparison of results between countries. In addition, standardized methodology makes possible the comparison from one time period to another within a particular country, which is needed to measure the impact of the immunization programme's efforts over time.

The primary purpose of the methodology developed by EPI was to assess the level of immunization coverage. Because of the operational success of this methodology, it has been adopted for a range of other purposes including assessments of population morbidity due to specific causes, service coverage and health service needs. It is the purpose of this article to review the EPI methodology, to consider its suitability for other purposes, and to sug-

^a WHO Consultant and Associate Professor, Division of Public Health, School of Health Sciences, University of Massachusetts, Amherst, Massachusetts, United States of America.

^b Medical Officer, Development of Epidemiological and Health Statistical Services, Division of Epidemiological Surveillance and Health Situation and Trend Assessment, World Health Organization, Geneva, Switzerland.

ENQUÊTES SUR LA COUVERTURE ET L'IMPACT DES PROGRAMMES: MÉTHODES QUANTITATIVES UTILISÉES PAR LE PROGRAMME ÉLARGI DE VACCINATION

Stanley Lemeshow^a & David Robinson^b

Introduction

En vue d'améliorer la situation sanitaire de la population, la plupart des pays sont en train de se doter d'une capacité accrue dans le domaine des soins de santé primaires. Il faut pour cela pouvoir gérer le système national de santé, ce qui exige des données aux fins de la planification, de la supervision et de la surveillance des activités sanitaires. Des données sont indispensables pour définir les besoins en matière de services de santé, l'efficacité des services en place et leur influence sur la morbidité et la mortalité.

Si les données nécessaires peuvent en grande partie être trouvées auprès de sources régulières, les enquêtes constituent dans certains cas le meilleur moyen d'information. L'Organisation mondiale de la Santé (OMS), à côté d'autres organismes internationaux, a beaucoup contribué à promouvoir l'utilisation de ces enquêtes. Par exemple, grâce à son Programme élargi de vaccination (PEV), elle s'efforce de rendre la vaccination accessible pour tous les enfants du monde d'ici 1990. Cet effort est considéré comme une étape décisive vers la réalisation de l'objectif formulé par l'OMS, à savoir la santé pour tous en l'an 2000.

Le fardeau que représentent pour les enfants du monde en développement des maladies comme le tétanos du nouveau-né, la poliomyélite, la rougeole, la coqueluche, la diphtérie et la tuberculose, est encore alourdi du fait que le plus souvent, ces enfants sont gravement affaiblis par la malnutrition et des crises de diarrhée et de paludisme à répétition. On voit dans les programmes de vaccination un moyen important de réduire les incapacités et la mortalité chez les nourrissons et les jeunes enfants et, partant, de contribuer au maintien d'une forte natalité.

Un élément important dont avait besoin le PEV consistait dans un système approprié de collecte de l'information. Des données fiables sont nécessaires pour connaître la morbidité et la mortalité exactes associées à telle ou telle maladie cible ainsi que la couverture vaccinale correspondante. Comme, dans de nombreux pays en développement, ce type d'informations n'est pas immédiatement accessible au gestionnaire de la santé, le PEV a travaillé à une méthode qui permettrait d'obtenir des données exactes rapidement et pour un coût modique. De plus, on a recherché une méthode applicable de façon relativement normalisée dans tous les pays. Une telle méthode permet en effet de préparer des manuels opérationnels et des matériels de formation susceptibles d'une utilisation générale et facilite la comparaison des résultats d'un pays à l'autre. En outre, une méthodologie normalisée rend possible les comparaisons entre plusieurs époques dans un même pays, ce qui est indispensable pour suivre l'amélioration progressive obtenue grâce au programme de vaccination.

La méthodologie mise au point par le PEV visait principalement l'évaluation de la couverture vaccinale. Mais devant son succès opérationnel, cette méthodologie a été adoptée à d'autres fins, notamment l'évaluation des taux de morbidité spécifiques dans une population donnée, la couverture assurée par les services existants et les besoins en matière de services de santé. Dans le présent article, on se propose de présenter la méthodologie PEV, d'étudier la possibilité de

^a Consultant OMS et Associate Professor, Division of Public Health, School of Health Sciences, Université du Massachusetts, Amherst, Massachusetts, États-Unis d'Amérique.

^b Médecin au service du Développement des services d'épidémiologie et de statistiques sanitaires, Division de la surveillance épidémiologique et de l'appréciation de la situation sanitaire et de ses tendances, Organisation mondiale de la Santé, Genève, Suisse.

gest limitations, modifications and alternatives to meet the needs of different health programmes.

The EPI survey

Theoretical basis

It is not possible to study every child in a target age range in order to assess the level of immunization coverage. As a result, a sampling scheme was sought which would be effective, relatively easy to carry out in a short period of time and inexpensive. Simple random sampling (SRS) would not be suitable: accurate lists of children could not be readily obtained or constructed, and travel between selected children, who might be widely scattered, could be both costly and time consuming.

An alternative sampling strategy, with considerable operational advantages, is known as "probability proportionate to size" (PPS) cluster sampling. The principle of this method is as follows: it is assumed that the population under consideration can be divided into M groups or "clusters" on the basis of geographic location. The population of each cluster must be known or estimable from census or other government figures.

Let N_i denote the population size of the i^{th} cluster, where i can take on any value between 1 and M . At the first stage of sampling, a subset of m clusters is selected from the complete list of M clusters. This selection is done in such a way that large clusters have a higher probability of being selected than do small clusters. The actual selection is carried out using random numbers, and it should be noted that it is possible for a large cluster to be selected more than once in making up the subset of m clusters.

As with any cluster sampling method, this procedure entails a big saving in time and expense, since detailed frames of children in the target age group need be compiled only for the m selected clusters rather than for the entire population. In fact the compilation of these frames can, if necessary, be delayed until the m clusters to be included in the sample have been identified.

At the second stage of sampling, \bar{n} children are selected from the total population (N_i) of children living in each of the selected m clusters. This selection is prescribed to be carried out *at random*. When carried out in this manner there are a number of distinct advantages with PPS cluster sampling. First of all, resulting estimates are "self-weighting". In other words, the sizes of the clusters do not enter into computations of proportions or associated standard errors.

For example, to estimate the proportion of the population immunized, the following formula is used:

$$\hat{P} = \frac{1}{m} \left\{ \sum_{i=1}^m y_{i*} / \bar{n} \right\} \quad (1)$$

where $y_{i*} = \sum_{j=1}^{\bar{n}} Y_{ij}$ = total number of children adequately immunized from among the \bar{n} children studied in the i^{th} cluster and

$$Y_{ij} = \begin{cases} 1 & \text{if the } j^{\text{th}} \text{ child in the } i^{\text{th}} \text{ cluster is adequately immunized} \\ 0 & \text{if the } j^{\text{th}} \text{ child in the } i^{\text{th}} \text{ cluster is not adequately immunized.} \end{cases}$$

l'appliquer à d'autres fins et, après en avoir indiqué les limites, de proposer des modifications et variantes en vue de répondre aux besoins de programmes de santé différents.

L'enquête PEV

Base théorique

Vu qu'il est impossible d'examiner chaque enfant, dans un groupe d'âges donné, en vue d'évaluer le taux de couverture vaccinale, on a cherché une méthode d'échantillonnage qui serait à la fois efficace, peu onéreuse et d'exécution relativement facile dans un bref laps de temps. Un sondage aléatoire simple (SAS) ne convient pas, car il est difficile de se procurer ou d'établir un recensement des enfants, et les déplacements nécessaires pour aller voir chacun des enfants choisis seraient à la fois longs et onéreux étant donné la dispersion de la population.

Une autre méthode d'échantillonnage, qui a des intérêts considérables sur le plan opérationnel est connue sous le nom d'échantillonnage par grappes avec « probabilité de tirage proportionnelle à la taille » (PPT). Le principe de la méthode est le suivant: on suppose au départ que la population visée peut être répartie en M groupes ou « grappes » en fonction de leur localisation. La population de chaque grappe doit être connue, au moins de façon approchée d'après les résultats d'un recensement ou d'autres statistiques officielles.

Soit N_i l'effectif de la $i^{\text{ème}}$ grappe, i pouvant prendre toutes les valeurs entières comprises entre 1 et M . Lors de la première étape du sondage, on choisit dans la liste complète des M grappes un sous-ensemble de m grappes. Ce choix est effectué de façon que les grappes importantes aient une probabilité d'être tirées plus élevée que les petites grappes. En pratique, on se sert d'une table de nombres au hasard si bien qu'il peut arriver qu'une même grappe d'effectif élevé soit choisie plusieurs fois lors de la constitution du sous-ensemble de m grappes.

Comme c'est le cas chaque fois qu'on applique une méthode de sondage par grappes, on réalise ainsi d'importantes économies de temps et d'argent puisque l'établissement de listes détaillées d'enfants appartenant au groupe d'âges visé n'est nécessaire que pour les m grappes ainsi choisies, et non pour la totalité de la population. En fait, la compilation de ces listes peut, s'il y a lieu, être différée en attendant que les m grappes qui seront incluses dans l'échantillon aient été identifiées.

A la seconde étape du sondage, \bar{n} enfants sont choisis parmi la population totale (N_i) des enfants qui habitent dans chacune des m grappes retenues. La méthode exige que ce choix se fasse *au hasard*. Lorsqu'on procède ainsi, l'échantillonnage par grappes PPT présente un certain nombre d'avantages indéniables. Tout d'abord, les estimations établies de cette façon sont « auto-pondérées ». En d'autres termes, la taille des grappes n'intervient pas dans le calcul des proportions ou des erreurs-types associées.

Par exemple, pour estimer la proportion des sujets vaccinés dans la population, on applique la formule ci-dessous:

$$\hat{P} = \frac{1}{m} \left\{ \sum_{i=1}^m y_{i*} / \bar{n} \right\} \quad (1)$$

où $y_{i*} = \sum_{j=1}^{\bar{n}} Y_{ij}$ désigne le nombre total d'enfants correctement vaccinés parmi les \bar{n} enfants étudiés dans la $i^{\text{ème}}$ grappe et

$$Y_{ij} = \begin{cases} 1 & \text{si le } j^{\text{ème}} \text{ enfant de la } i^{\text{ème}} \text{ grappe est correctement} \\ & \text{vacciné} \\ 0 & \text{si le } j^{\text{ème}} \text{ enfant de la } i^{\text{ème}} \text{ grappe n'est pas correcte-} \\ & \text{ment vacciné.} \end{cases}$$

Expression (1) is equivalent to:

$$\hat{P} = \frac{1}{m} \sum_{i=1}^m \sum_{j=1}^{\bar{n}} Y_{ij/m\bar{n}} \quad (2)$$

which is simply the total number of immunized children divided by the total number of children studied. This estimate is called self-weighting, since there was no need to incorporate the N_i into the formula.

An estimate of the standard error of this estimated proportion (necessary for the construction of confidence intervals) is also relatively straightforward, with computation as follows:

$$SE(\hat{P}) = \left[\frac{1}{m(m-1)} \sum_{i=1}^m (\hat{P}_i - \hat{P})^2 \right]^{1/2} \quad (3)$$

where \hat{P}_i is the proportion immunized in the i^{th} cluster and \hat{P} is the proportion immunized over all sample clusters as given in (2) above. This expression is considerably easier to calculate than typical cluster sampling formulae, since only the variability between the estimated proportions in the sampled clusters is needed. An expression for the intracluster variability is not incorporated into this formula.

It should be noted that, strictly speaking, for these formulae to hold, it is necessary that the m sample clusters be selected with probability proportionate to size and that simple random samples of \bar{n} children be selected from each of the m clusters.

Modifications for EPI

Since 1978, WHO/EPI has been advocating a PPS cluster sampling procedure for surveys of immunization coverages (1, 2). The method adopted is a modification of a survey technique originally used for immunization coverage in the United States of America (3) and later updated for use in the smallpox eradication programme in West Africa (4). By the end of 1982, at least 441 surveys of this type had been carried out worldwide (5).

The EPI survey, as it is currently carried out for determining immunization coverage, involves the detailed review of immunization status of approximately 210 children by trained reviewers. The current convention is to identify and visit 30 "clusters" which may be in cities, towns, or villages, and to study 7 children in each. As a result, the EPI surveys are commonly referred to as "30 x 7" surveys.

The rationale for using 210 individuals is as follows. First of all, it was decided that it was necessary to be able to estimate immunization coverage to within 10 percentage points of the true population proportion. Using a population proportion of 50% coverage as the basis, and desiring to be 95% confident that the resulting estimate would be in the interval 40-60%, it was determined that a simple random sample of size 96 would be required. To select a simple random sample of this size from the population was not operationally feasible and, as a result, a cluster sampling strategy was deemed necessary.

In order to achieve the same precision with cluster sampling as would be possible with simple random sampling, experience suggested that a cluster sample of approximately twice the size ("design effect") of the simple random sample would be needed. Because of the economy afforded by the cluster sampling strategy, this larger sam-

L'expression (1) peut se mettre sous la forme équivalente:

$$\hat{P} = \frac{1}{m} \sum_{i=1}^m \sum_{j=1}^{\bar{n}} Y_{ij/m\bar{n}} \quad (2)$$

dans laquelle le second membre représente simplement le rapport du nombre total d'enfants vaccinés au nombre total d'enfants étudiés. Cette estimation est dite auto-pondérée puisque le paramètre N_i n'intervient pas dans la formule.

Il est également relativement facile d'évaluer l'erreur-type correspondant à cette proportion estimative (qui est indispensable pour la détermination des intervalles de confiance), en appliquant la formule ci-dessous:

$$SE(\hat{P}) = \left[\frac{1}{m(m-1)} \sum_{i=1}^m (\hat{P}_i - \hat{P})^2 \right]^{1/2} \quad (3)$$

dans laquelle \hat{P}_i représente la proportion des enfants vaccinés dans la $i^{\text{ème}}$ grappe et \hat{P} la même proportion d'enfants pour la totalité des grappes, donnée par la formule (2). Le calcul est infiniment plus facile que dans le cas des formules classiques de l'échantillonnage par grappes, car il suffit de connaître la dispersion des proportions estimatives pour les diverses grappes faisant partie de l'échantillon. La formule ne fait pas intervenir la dispersion à l'intérieur de chacune des grappes.

On notera que pour que ces formules soient valables, il faut, en toute rigueur, que la probabilité de tirage de chacune des m grappes constituant l'échantillon soit proportionnelle à sa taille et que les échantillons de \bar{n} enfants constitués à partir de chacune d'elles le soient par sondage aléatoire simple.

Modifications en vue du PEV

Depuis 1978, l'OMS/PEV préconise l'application aux enquêtes sur la couverture vaccinale de la méthode d'échantillonnage par grappes PPT (1, 2). La méthode adoptée constitue une variante d'une technique d'enquête utilisée à l'origine pour l'étude de la couverture vaccinale aux Etats-Unis d'Amérique (3) et qui, par la suite, a été revue en vue de son application dans le programme d'éradication de la variole en Afrique occidentale (4). A la fin de 1982, au moins 441 enquêtes de ce type avaient été réalisées dans le monde (5).

L'enquête PEV, telle qu'elle est actuellement pratiquée pour déterminer la couverture vaccinale, implique que des enquêteurs dûment formés examinent en détail l'état vaccinal d'environ 210 enfants. On est actuellement convenu d'identifier 30 «grappes» formées de personnes habitant par exemple une ville, grande ou moyenne, ou un village et d'étudier sur place 7 enfants dans chacune de ces grappes. C'est pourquoi les enquêtes PEV sont couramment qualifiées d'«enquêtes 30 x 7».

Le choix du nombre 210 s'explique comme suit: tout d'abord, on a estimé nécessaire d'obtenir une estimation de la couverture vaccinale avec une erreur maximale de 10 points (en plus ou en moins) par rapport au pourcentage réel. En prenant pour base une couverture de 50%, on trouve, si l'on veut avoir 95 chances sur 100 que l'estimation correspondante tombe entre 40% et 60%, qu'il faut un échantillon aléatoire simple de 96 personnes. Comme il est impossible, pour des raisons opérationnelles, de prélever un échantillon aléatoire simple de cette taille dans une population, on a jugé indispensable d'appliquer la méthode de l'échantillonnage par grappes.

Pour obtenir la même précision avec cette seconde méthode, il faut, comme le montre l'expérience, utiliser un échantillon d'effectif double («effet de conception»). Vu les économies que permet la méthode de l'échantillonnage par grappes, on peut étudier cet échantillon pourtant plus nombreux à un moindre coût, et plus commodément. Ainsi, la taille de

can be studied both more conveniently and less expensively. As a result, the necessary sample size with cluster sampling was estimated to be 192.

Based on procedures adopted for use in the United States at that time (3) and, taking into account practical as well as logistic factors, it was decided that 30 clusters would be used. This meant that 7 children per cluster must be studied in order to attain the specified sample size. There is no particular statistical advantage in using 30 clusters and it is perfectly reasonable that the operational considerations should dictate the number to be selected. However, it should be noted that if one is satisfied to have an estimate which will be within 10 percentage points of the true P with 95% confidence, a different combination of m and n might result in significant savings in time and cost. On the other hand, for the same time and cost, alternative combinations of m and n could yield increased precision. Decisions as to the exact value of m and n to use in a particular country would have to be tailored to the specific characteristics of that country. These could involve numerous assumptions regarding costs likely to be incurred at each stage of sampling, as well as estimates of intracluster correlation coefficients which, in their own right, might be of questionable accuracy.

The classic PPS cluster sampling scheme as described above has certain features which may make implementation in field conditions difficult. In particular, the random selection at the second stage may not always be possible — particularly in scattered rural areas. The classic methodology has thus been modified, and the 30 x 7 survey currently being used by the EPI may be characterized as a PPS cluster sample without random selection at the second stage.

Field methodology

The first step in carrying out an immunization coverage survey is to identify precisely what the population is and which age groups within this population are of particular interest. Clusters are identified by compiling a list of all cities, sectors of cities, towns and villages included in the target area for which immunization coverage is to be evaluated.

Although detailed lists of individuals or families may well not exist for each of the clusters on the list, all that is really needed at this point is a reasonably accurate population estimate for each cluster so that PPS selection can take place. In practice, this may present some difficulties since it may be several years since the last census update was carried out. The most important assumption that must be made in this case is that any changes that have occurred in the population since the last figures were compiled affect equally all clusters in the country being studied. For example, if cluster A had 3 times as large a population as did cluster B at the time of the last census, that same 3:1 ratio is assumed to hold at the time of the survey. If census data are not available, intelligent estimates must be made of the population sizes in the various clusters. With PPS sampling, the relative sizes of the clusters are more important than the actual sizes.

Once the list of M clusters and associated population sizes has been compiled, m of the clusters can be selected with probability proportionate to size. This is accomplished by computing the cumulative population, cluster by cluster, with the cumulative total equalling the total population size, N . By dividing the total population size by the number of clusters sought (30 in this case), 30 "zones" are identified, each containing $N/30 = K$ individuals. K is referred to as the "sampling interval". By selecting a single random number between 1 and K (call this i), the first cluster to be sampled is identified as the one which includes the i -th individual on the cumulative

l'échantillon nécessaire serait en principe de 192.

En s'inspirant des méthodes retenues à l'époque aux États-Unis (3) et compte tenu de facteurs pratiques et logistiques, on a décidé de fixer à 30 le nombre de grappes. Il fallait donc choisir 7 enfants par grappe pour obtenir un échantillon dont la taille ne soit pas inférieure au chiffre indiqué ci-dessus. Le choix de 30 grappes ne présente pas d'avantages particuliers sur le plan statistique de sorte qu'on peut parfaitement admettre que le choix du nombre précis soit dicté par les considérations opérationnelles. Cependant, on notera que si l'on se satisfait d'obtenir une estimation ne différant pas de plus de 10 points de la proportion réelle P (au niveau de confiance de 95%), une combinaison différente de m et de n pourrait permettre des économies de temps et d'argent. Par ailleurs, pour une durée et un coût donnés, certaines combinaisons des valeurs m et n peuvent permettre une plus grande précision. Le choix des valeurs exactes à retenir pour ces 2 paramètres dans un pays déterminé doit être adapté aux caractéristiques de celui-ci. Cela amène à formuler de nombreuses hypothèses sur les coûts probables à chaque stade de l'échantillonnage et la valeur estimative des coefficients de corrélation intra-grappe, qui peuvent elles-mêmes être d'une exactitude douteuse.

La méthode classique d'échantillonnage par grappes PPT qui vient d'être décrite comporte certains aspects qui en rendent l'application difficile sur le terrain. En particulier, le choix aléatoire n'est pas toujours possible au second stade, en particulier dans les zones rurales à population clairsemée. La méthodologie classique a donc été modifiée de sorte qu'on peut dire que les enquêtes 30 x 7 actuellement pratiquées par le PEV constituent un échantillonnage par grappes PPT sans choix aléatoire au second stade.

Méthodologie pratique

La première étape d'une enquête sur la couverture vaccinale consiste à identifier de façon précise la population en cause et, à l'intérieur de cette population, les groupes d'âges présentant un intérêt particulier. Les grappes sont définies en dressant la liste de tous les quartiers, villes, bourgades et villages inclus dans la zone où l'on cherche à évaluer la couverture.

Il peut fort bien se faire qu'il n'existe pas de liste détaillée des particuliers ou des familles pour chacune des grappes constituant la liste; mais, à ce stade, il suffit en réalité d'une estimation raisonnable de la population de chaque grappe pour qu'on puisse procéder à un tirage du type PPT. En pratique, cela peut soulever des difficultés si le dernier recensement remonte à plusieurs années. En pareil cas, l'hypothèse la plus importante qu'on est contraint de faire est d'admettre que les modifications éventuellement intervenues au sein de la population depuis l'établissement des dernières statistiques atteignent également toutes les grappes du pays étudié. Par exemple, si la grappe A était 3 fois plus nombreuse que la grappe B à l'époque du dernier recensement, on admettra que cette proportion de 3 à 1 est encore valable au moment de l'enquête. En l'absence de résultats d'un recensement, il faut faire une estimation rationnelle de l'effectif des diverses grappes. Dans le cas d'un échantillonnage PPT, c'est la taille relative des grappes, et non les effectifs réels qui a le plus d'importance.

Une fois qu'on a dressé la liste de M grappes et fixé les effectifs correspondants, on peut choisir m de ces grappes avec une probabilité de tirage proportionnelle à la taille. Pour cela, on additionne successivement les effectifs des diverses grappes, le total cumulatif obtenu étant égal à l'effectif total de la population, soit N . En divisant ce total par le nombre de grappes souhaité (30 dans le cas présent), on identifie 30 «zones» dont chacune contient $N/30 = K$ individus. K est désigné sous le nom d'«intervalle d'échantillonnage». En choisissant au hasard un nombre entier compris entre 1 et K (désigné ci-après par i), on définit la première grappe à inclure dans l'échantillon comme étant celle qui contient le $i^{\text{ème}}$ indi-

list. Starting from this position on the cumulative list, 29 further clusters are identified by successively adding the sampling interval, K . Thus, $i, i+K, i+2K, \dots, i+29K$ define the 30 clusters. This type of sampling is called a systematic sample of clusters and the probability that a particular cluster will be included in the sample is directly related to the size of the cluster. In fact, large clusters can be selected more than once in this systematic sampling process.

Once the sample cluster is identified, it is then necessary to determine which individuals to study within the cluster. The method advocated by EPI is as follows: a household is selected "at random" from all households in the cluster. In practice, it may not be possible to make this selection truly at random, since the method used will depend upon the density of the population as well as other factors such as the availability of lists of households. When household lists are available, the households are numbered and 1 random number is selected to represent the first sampled household. If, on the other hand, the cluster is a small village and household lists do not exist, it may be possible to carry out a quick census of households, numbering them, and then selecting a random number to represent the first sampled household. In fact, the process of carrying out a census is to be preferred, since it is then reasonably assured that the list is complete.

Enumerating all households is often impossible in moderately large towns or widely scattered rural populations. In that case, EPI suggests that the interviewer go to some centrally located landmark, (such as a church, school or market) randomly select a direction to walk (e.g., north, south, east or west), count the number of households (call this L) found in that direction from the central point to the town boundary, and finally, select a random number between 1 and L . This number identifies the randomly selected starting household.

In urban areas, the process of identifying a random starting household may be much more difficult. One possibility is a 2-step process where the city is subdivided into geographically contiguous zones, a zone is selected at random, and a starting household is then identified within the zone. If household lists do not exist and if the zones are made small enough, it may be possible to carry out a census in the selected zone before selecting a household. Usually, there is no single strategy which can be applied to all situations, and solutions must often be devised on the spot to deal with living arrangements such as multiple dwelling units and apartment buildings.

It should be noted that there are certain situations where valid lists of target populations are not available and cannot be reliably constructed. For instance, in certain societies there are large numbers of "street dwellers": i.e., individuals who do not live in permanent dwellings and who do not, therefore, figure in local census data. In fact the very concept of household may have unclear meaning with such individuals, and compiling accurate lists to include them is rarely done since such lists are immediately obsolete.

Upon entering the first household, the interviewer must first determine whether there are any individuals living there who are in the target age group. If there are, the required information is collected for each such individual. As presently recommended by EPI, if no one is at home in the selected household, the interviewer moves on to the next household. There is no provision to revisit households.

After the first household is visited, whether there is an individual in the target group or not, the interviewer proceeds to the "next" household. This is defined as that residence whose front door is physically closest to the one just visited. This process of visiting households is

vidu de la liste cumulative. Ce point de départ étant repéré sur la liste cumulative, on définit les 29 grappes suivantes en se décalant chaque fois d'un intervalle égal à l'intervalle d'échantillonnage K . Ainsi, ce sont les individus de rang $i, i+K, i+2K, \dots, i+29K$ qui définissent les 30 grappes. Cette méthode fournit un échantillon systématique de grappes, dont chacune a une probabilité d'être incluse dans l'échantillon qui est directement liée à sa taille. En fait les grappes d'effectif élevé peuvent être tirées plusieurs fois avec cette méthode d'échantillonnage systématique.

Une fois choisies les grappes constituant l'échantillon, il faut déterminer les individus à étudier dans chacune d'elles. La méthode préconisée par le PEV est la suivante: on choisit « au hasard » un ménage parmi tous ceux qui appartiennent à la grappe. En pratique, un choix rigoureusement aléatoire est parfois impossible car la méthode appliquée dépend de la densité démographique ainsi que d'autres facteurs, par exemple l'existence de listes des ménages. Quand on dispose de listes de ce type, on numérote les ménages et l'on choisit le ménage à étudier par tirage d'un nombre au hasard. En revanche, quand la grappe est constituée par un petit village et qu'il n'existe pas de listes des ménages, on peut procéder à un recensement rapide de ceux-ci, puis les numérotter et procéder comme dans le cas précédent par tirage d'un nombre au hasard. En fait, la préférence doit être accordée au recensement car on a ainsi l'assurance à peu près absolue que la liste est complète.

Souvent, il est impossible d'énumérer la totalité des ménages dans une ville de dimension moyenne ou dans une population rurale très disséminée. En pareil cas, le PEV suggère de procéder comme suit: l'enquêteur se rend dans un lieu de situation centrale (par exemple une église, une école ou un marché), il choisit au hasard une direction (par exemple le nord, le sud, l'est ou l'ouest) et s'éloigne dans cette direction en comptant le nombre de ménages (L) qu'il trouve sur son chemin jusqu'à la limite de la ville et, finalement, il choisit au hasard un nombre compris entre 1 et L . Ce nombre sert à repérer le ménage d'où va démarrer l'enquête.

Dans les zones urbaines, le choix de ce ménage au hasard peut être plus difficile. Il existe diverses méthodes: on peut, par exemple, procéder en 2 étapes, en subdivisant d'abord la zone en plusieurs zones géographiquement contiguës et, après avoir choisi l'une d'elles au hasard, en repérant dans cette zone le ménage de départ. Quand il n'existe pas de liste des ménages et que les zones obtenues par découpage sont suffisamment petites, on peut effectuer un recensement dans la zone choisie avant d'y désigner un ménage. A l'évidence, il n'existe pas de méthode passe-partout et il faut souvent découvrir sur place la solution qui permet de tenir compte de la structure de l'habitat par exemple de l'existence de bâtiments à plusieurs cellules individuelles ou d'immeubles d'habitation collective.

A noter qu'il existe des cas où l'on ne dispose pas de liste fiable de la population à étudier et où il n'est pas non plus possible d'en établir valablement. Par exemple, certaines sociétés comptent un grand nombre de « sans-abri » qui dorment dans la rue et qui, par suite, ne figurent pas dans les recensements locaux. En fait, le concept même de ménage est peut-être discutable pour ces personnes, de sorte qu'elles sont rarement comptabilisées dans des listes précises, vu que ces listes sont immédiatement périmées.

Quand il prend contact avec le premier ménage choisi, l'enquêteur doit commencer par établir s'il compte parmi ses membres des sujets du groupe d'âge visé. Dans l'affirmative, l'enquêteur réunit pour chacun d'eux les renseignements voulus. Conformément à la politique actuellement recommandée par le PEV, si l'enquêteur ne rencontre aucun sujet de cet âge dans le ménage choisi, il passe au suivant, sans qu'une seconde visite soit prévue.

Ainsi après s'être rendu dans le premier ménage et qu'il y ait ou non rencontré avec un sujet appartenant au groupe visé, l'enquêteur passe au ménage « suivant ». On entend par là le ménage qui habite le plus près, à en juger par la distance des portes d'entrée des 2 logements. Cette tournée se pour-

repeated until a total of 7 individuals of the appropriate age have been studied in the sampled cluster. As an operating rule, all individuals in that household contributing the seventh individual to the sample are studied, even if that results in 8-10 individuals in the cluster rather than the target number of 7.

Potential problems with the EPI methodology

The EPI survey has proved to be a most useful tool for providing health managers with essential information for planning health programmes. Normally a survey will require about 5 days of work for 4-6 teams, and is routinely included as part of EPI management training. Results from these surveys have provided the incentive for resource allocation which has allowed immunization programmes to expand and increase their impact. Without these surveys, many national programmes would have no means for assessing their progress.

It should be evident from the preceding sections that the procedure followed by EPI has diverged from standard PPS cluster sampling methodology. This divergence occurs at the second stage of sampling (i.e., selection of households and individuals). From a statistical point of view this is a cause for concern since, in order for the formulae presented earlier in this article to hold, it is assumed that individuals studied at the second stage are the result of random selection, and the procedure advocated by EPI does not achieve this.

Theoretically, the households should be selected independently of each other and should be representative of the totality of households in the cluster. The EPI method, by selecting a starting household and then visiting geographically related households, ensures that this will not be the case. The effect of this is impossible to quantify but, intuitively, households which are spatially related may have other factors in common, including access to immunization facilities, water supply, disease exposure, etc. A particular example of this might be the pocketing of unimmunized households in inaccessible slum areas of cities. There is a risk that surveys of adjacent households could either over- or under-estimate the true population coverage depending upon where the starting household happens to fall. In fact, where such pocketing is suspected, and where the objective is to immunize individuals living in such pockets, other survey methods may give the manager more immediately useful data than would be possible even with a perfectly implemented PPS survey. In practice, in circumstances such as these, special arrangements are usually made by EPI managers.

In some situations, the distance between the central point and the edge of the community may be too large for the prescribed EPI procedure to be practical. Although this is not advocated by EPI, the interviewer may find it more realistic to simply select the direction to be taken from a fixed starting point and to pick out and visit a house at random in that direction without first counting the number of households in that direction. Strictly speaking, this method of selecting the starting household is not random, and although it does assure that the interviewer does not exercise personal judgement in the selection process, it still introduces statistical bias.

Leaving selection of successive households to the interviewer presents another opportunity for bias. This may occur when an interviewer must decide which household is closest to the one just visited. If, for example, this choice is between one household in a slum area and another not in a slum area, there is a possibility that the

suit jusqu'à ce qu'on ait examiné au total 7 personnes de l'âge convenable dans la grappe échantillonnée. Une règle pratique consiste à examiner tous les membres du ménage où l'on rencontre le 7^{ème} sujet, même si cela doit aboutir à retenir pour cette grappe 8-10 sujets au lieu de 7 qui constitue en principe le chiffre à atteindre.

Problèmes potentiels soulevés pour la méthodologie du PEV

A l'expérience, les enquêtes du type PEV se sont révélées un outil fort utile en vue de fournir aux gestionnaires de la santé les informations essentielles pour la planification de leurs programmes. En principe, une enquête exige environ 5 jours de travail pour 4-6 équipes, et il en figure systématiquement une dans le cadre de la formation à la gestion du PEV. Les résultats de ces enquêtes ont constitué un argument de poids pour l'obtention de ressources qui ont permis aux programmes de vaccination de se développer et d'avoir un impact accru. Sans ces enquêtes, les responsables des programmes nationaux seraient bien souvent dans l'incapacité d'apprécier le progrès accompli.

Il ressort des sections qui précèdent que la méthode appliquée par le PEV s'écarte de la méthodologie classique de l'échantillonnage par grappes PPT. La différence se situe au niveau du second stade de l'échantillonnage (c'est-à-dire du choix des ménages et des individus). Du point de vue statistique, c'est là un point préoccupant puisque, pour que les formules présentées plus haut soient valables, il faut admettre que les sujets étudiés lors du second stade sont issus d'un échantillonnage aléatoire, ce qui n'est pas le cas avec la méthode du PEV.

En théorie, les ménages devraient être choisis indépendamment les uns des autres et être représentatifs de la totalité des ménages constituant la grappe. Ce n'est pas le cas avec la méthode du PEV puisqu'elle repose sur le choix et la visite d'un ménage de départ suivie de celle de ménages qui ont avec le premier un lien géographique. Il est impossible d'estimer quantitativement l'effet de ce biais mais, intuitivement, il est clair que des ménages voisins dans l'espace risquent de partager d'autres facteurs, par exemple l'accès aux moyens de vaccination, l'approvisionnement en eau, l'exposition aux maladies, etc. Comme exemple particulier, on peut citer la constitution d'îlots de ménages non vaccinés dans les banlieues pauvres et inaccessibles des grandes villes. En limitant l'enquête à des ménages voisins dans l'espace, on risque d'aboutir à une surestimation ou à une sous-estimation de la couverture réelle, selon la localisation aléatoire, du ménage de départ. En fait, lorsqu'on a des raisons de croire à l'existence d'îlots de ce type et qu'on se propose précisément de vacciner leurs habitants, d'autres méthodes d'enquête peuvent fournir aux gestionnaires des données plus directement utiles que celles qu'on peut réunir avec une enquête du type PPT, même parfaitement conduite. En pratique, les gestionnaires du PEV adoptent en général des dispositions particulières en pareilles circonstances.

Dans certaines situations, le point central est trop éloigné de la limite de la collectivité pour qu'on puisse appliquer le procédé prescrit par le PEV. Bien que la méthode ci-après ne soit pas recommandée par le PEV, l'enquêteur peut estimer plus réaliste de fixer simplement la direction à suivre à partir d'un point déterminé et de choisir dans cette direction une maison au hasard, dans laquelle il va procéder à son enquête, sans d'abord compter le nombre total de ménages rencontrés dans cette direction. A strictement parler, cette façon de choisir le ménage de départ n'est pas aléatoire et, bien qu'elle élimine l'intervention du jugement personnel de l'enquêteur dans le choix, elle n'y introduit pas moins un biais statistique.

En laissant le choix des ménages successifs à la discrétion de l'enquêteur, on introduit un autre risque de biais. C'est le cas quand l'enquêteur doit choisir le ménage qu'il estime être le plus proche de celui chez lequel il vient de se rendre. A supposer, par exemple, qu'il ait à choisir entre 2 ménages dont l'un seulement est situé dans une zone de taudis, il

viewer's preferences may result in one or the other being adequately represented. Areas which are not easily accessible may also be underrepresented if left to discretion of the interviewer.

Further potential for bias is inherent in the practice of revisiting households found to be unoccupied at the of the interviewer's visit since certain subgroups will not be adequately represented.

The selection of the starting household is also a cause of concern since the exigencies of field operations may limit the possibility of a truly random selection. This results in households inadvertently being selected on grounds of convenience. In a scattered rural community, for example, the tendency may be to select the starting household in the area of densest population. Such selection is subject to bias since such centres may also be the sites for outreach services and other health care facilities. These circumstances, even if the direction in which the interviewer is to walk is selected at random and the household is selected at random from all houses lying in that direction, considerable bias may still result.

Finally, there is a potential problem in having to use documented evidence of immunization status. This problem cannot be fully overcome but is often taken into account when analysing the data.

Applications of the EPI methodology

An EPI survey was designed for the express purpose of measuring immunization coverage, either in the absence of any data, or when data of doubtful validity exist. In recent years, the sampling methodology developed for such surveys has been applied not only to assessments of immunization coverage but also to assessments of changes in immunization coverage over time. With some modifications, the methodology has also been applied to studies of the incidence of poliomyelitis, neonatal tetanus and diarrhoea, as well as to studies of mortality due to malaria. Recently, the same procedure has been used in many surveys to assess various factors relating to the availability and use of health services.

Based on the above discussion, it should be clear that the particular methodology developed by EPI specifically for coverage surveys might not be appropriate for other surveys having different objectives. For example, for surveys designed to document the expansion of coverage of an immunization programme over time, sample size calculations should incorporate such factors as the estimated coverage rate before the expansion, the anticipated coverage after the expansion as well as requirements for confidence and type II error rates. The rationale underlying the use of 210 does not take these factors into account and if smaller numbers were used, the study would involve an unacceptably high type II error rate (i.e., the probability of failing to detect a difference or change which actually existed).

The lack of modern knowledge concerning neonatal tetanus and poliomyelitis in developing countries arises from information from cluster sample surveys (7). Researchers studying the incidence of these diseases have realized that the procedures advocated for coverage surveys needed to be modified and, in particular, that larger sample sizes were required. The commonly accepted practice is that in each of 30 clusters, 70 livebirths occurred within a stated recall period (usually 4 to 6 months) are sought. The deaths among these livebirths are investigated and the proportion due to neonatal tetanus is estimated. All children aged 5-9 years in the households visited during the survey are examined for the presence of lameness due to poliomyelitis. This usually

risque de se laisser guider par ses préférences personnelles, et l'un des 2 types de ménages ne sera pas correctement représenté. De même, les zones difficilement accessibles risquent d'être sous-représentées si le choix est laissé à l'enquêteur.

Une autre possibilité de biais tient à la pratique adoptée de ne pas retourner dans un logement qui était inoccupé lors du passage de l'enquêteur car, de ce fait, certains sous-groupes ne seront pas correctement représentés.

Le choix du ménage de départ soulève également certains problèmes car les impératifs rencontrés sur le terrain peuvent exclure un choix véritablement aléatoire. Il peut en découler un choix intempestif de certains ménages, pour de simples raisons de commodité. Dans une communauté rurale disséminée, par exemple, on peut avoir tendance à choisir le ménage de départ dans la zone la plus peuplée. Ce choix est une source de biais puisqu'un tel centre peut aussi être privilégié dans la distribution des services périphériques et l'implantation des établissements de soins. Dans ces conditions, même si la direction dans laquelle l'enquêteur va partir est choisie au hasard et si, dans cette direction, le ménage est choisi de façon aléatoire dans tous les logements situés dans cette direction, il peut encore y avoir un biais extrêmement important.

Enfin, une source de difficultés possibles tient à l'emploi d'éléments autres que des certificats pour attester de la situation vaccinale. Ce problème est impossible à éliminer complètement, mais l'on en tient souvent compte lors de l'analyse des données.

Autres applications de la méthodologie du PEV

L'enquête PEV a été conçue dans un but bien défini: mesurer la couverture vaccinale, soit en l'absence totale de données, soit quand les données existantes sont de validité douteuse. Ces dernières années, la méthodologie d'échantillonnage mise au point pour ces enquêtes a été également appliquée à l'évaluation des variations de la couverture vaccinale avec le temps. Moyennant certaines modifications, la méthode a en outre été appliquée à l'étude de l'incidence de la poliomyélite, du tétanos du nouveau-né (6) et de la diarrhée, ainsi qu'à l'étude de la mortalité associée à la rougeole. Récemment, la même méthode a été employée pour des enquêtes sur différents facteurs en rapport avec la disponibilité et l'utilisation des services de santé.

D'après la discussion qui précède, il est clair que la méthodologie particulière mise au point par le PEV en vue, exclusivement, des enquêtes sur la couverture vaccinale ne convient pas forcément pour des enquêtes dont les objectifs sont différents. Par exemple, lorsqu'il s'agit de suivre l'extension, avec le temps, de la couverture d'un programme de vaccination, le calcul de la taille de l'échantillon doit tenir compte d'autres facteurs, par exemple du taux de couverture estimatif avant la mise en œuvre du programme, du taux prévu après son application ainsi que des conditions imposées pour la valeur des erreurs de première ou de seconde espèce. Les raisons qui justifient le choix du nombre 210 ne tiennent pas compte de ces facteurs, de sorte qu'en utilisant cet effectif on aboutirait à une erreur de seconde espèce très élevée (c'est-à-dire à une forte probabilité de ne pas mettre en évidence une différence ou évolution effectivement intervenue).

Ce qu'on sait aujourd'hui du tétanos du nouveau-né et de la poliomyélite dans les pays en développement provient en grande partie de données recueillies à l'aide de sondages par grappes (7). Les chercheurs qui étudient l'incidence de ces maladies ont reconnu que les méthodes préconisées pour les enquêtes de couverture devaient dans ce cas être modifiées et, en particulier, qu'il fallait un effectif plus élevé. La pratique courante consiste à rechercher dans chacune des 30 grappes 70 naissances vivantes ayant eu lieu depuis un délai déterminé (généralement 4 à 6 mois). Le recensement des décès intervenus après la naissance d'un enfant viable permet d'évaluer la mortalité par tétanos du nouveau-né. La recherche d'une paralysie d'origine poliomyélitique est systématiquement recherchée chez les enfants de 5 à 9 ans faisant

results in at least 200 children studied in each of the 30 clusters. The continued use of 30 clusters is based more on tradition and intuition than on statistical theory.

The need to consider carefully the choice of sample size can be illustrated by another example: if a study was planned to estimate the prevalence of a comparatively uncommon disease such as leprosy, the total required sample size would be much greater than that required in either of the previously described surveys in order to make estimates with acceptable precision. Specifically, if the rate of leprosy in a country is 1 per 1000 (i.e., $P = .001$), then a simple random sample of size of 15 350 would be required to be 95% confident that the sample estimate would be within 50% of the true value (i.e., between .0005 and .002) (8). An even larger sample would be required if a more complex sampling scheme were used or if greater precision is required. Clearly, for such uncommon diseases, planning the sample size so that the resulting estimate will be within 10 percentage points of the true P would be of little value.

For surveys of health service utilization and health status, there are often numerous and unrelated parameters being estimated in a wide age range. This difficulty is compounded by the fact that some of the parameters being studied are unlikely to be distributed homogeneously within communities. This is particularly true for accessibility of water supplies, communications and health service facilities. Planning sample size requirements under these conditions is extremely difficult, but it is an issue which must be confronted by agencies or others needing the information which would result from such studies. Blind acceptance of the 30 x 7 sampling scheme may seriously compromise the results of surveys and mislead planners and managers, despite the intensive efforts which may have gone into the execution of such surveys.

Evaluation of the 30 x 7 EPI sampling strategy via computer simulation

Despite the concern expressed over the lack of random selection of household with the EPI method, no studies have been carried out which have objectively evaluated its performance compared to a method which selects households at random. To this end, a computer simulation was designed and implemented at WHO Headquarters by the Epidemiological and Statistical Methodology Unit of the Division of Epidemiological Surveillance and Health Situation and Trend Assessment in cooperation with EPI and the Diarrhoeal Diseases Control Programme (CDD). The details of this study which will be presented in a separate report, are summarized here. Populations were constructed to represent low, moderate, high or very high population density. Households within the populations were distributed in 1 of 4 ways: at random, along a road or stream, centred around a single focus such as a market, school or church, or centred around 2 foci. One out of every 7 households was designated as having a child in the target age range and a certain overall proportion of children was designated "immunized". Four schemes were used to mark the immunized children: at random among all children, restricted to 1 "pocket" in which 100% of children are immunized and outside of which no children are immunized, restricted to 2 pockets of 100% immunization and, finally, restricted to 5 pockets of 100% immunization. The overall immunization levels in the various strata were either 10%, 30%, 50%, 70%, or 90%.

partie des ménages étudiés lors de l'enquête. En général, cela amène à examiner au moins 200 enfants dans chacune des 30 grappes. Ce chiffre de 30 continue d'être utilisé davantage pour des questions d'habitude et d'intuition que pour des raisons statistiques théoriques.

Un autre exemple montre pourquoi il convient de choisir soigneusement l'effectif de l'échantillon: supposons qu'une étude soit mise sur pied en vue d'estimer la prévalence d'une maladie relativement rare, comme la lèpre; dans ce cas, la taille de l'échantillon nécessaire doit être beaucoup plus importante que pour les 2 types d'enquêtes décrites ci-dessus si l'on veut avoir une précision acceptable. Plus précisément, si la lèpre frappe, dans un pays, 1 sujet sur 1000 (autrement dit si $P = 0,001$), il faudrait un échantillon aléatoire simple de 15 350 sujets pour qu'on puisse avoir la certitude, à 95% d'arriver à une estimation comprise entre la moitié et le double de la valeur réelle (c'est-à-dire entre 0,0005 et 0,002) (8). Il faudrait un échantillon encore plus nombreux si l'on appliquait une technique d'échantillonnage plus complexe ou si l'on souhaitait obtenir une plus grande précision. Il est clair que, pour des maladies aussi rares, il serait irréaliste de prévoir un échantillon de taille suffisante pour que l'estimation correspondante ne s'écarte pas de plus de 10 points de la valeur réelle P , en pourcentage.

Dans le cas d'enquêtes sur l'utilisation des services de santé et sur la situation sanitaire, l'estimation porte souvent sur de nombreux paramètres, sans lien les uns avec les autres, et cela dans des groupes d'âge très larges. La difficulté est encore renforcée du fait que certains des paramètres étudiés ont *a priori* une distribution hétérogène à l'intérieur des collectivités. C'est particulièrement vrai pour l'accès aux sources d'approvisionnement en eau, aux moyens de communication et aux établissements de soins. Dans ces conditions, il est fort difficile de définir la taille de l'échantillon nécessaire, mais c'est un problème auquel doivent faire face les organismes ou personnes qui ont besoin des renseignements que fourniraient ces études. L'application irréfléchie de la méthode d'échantillonnage 30 x 7 risque de compromettre gravement les résultats des enquêtes et d'induire en erreur les planificateurs et les gestionnaires, en dépit de tous les efforts qui auront pu être consacrés à l'exécution de ces enquêtes.

Evaluation de la méthode d'échantillonnage 30 x 7 du PEV par simulation sur ordinateur

Si le caractère strictement aléatoire du choix des ménages dans la méthode PEV a pu être jugé préoccupant, aucune étude n'avait été faite jusqu'ici en vue d'évaluer objectivement l'efficacité de cette méthode par comparaison à une méthode où les ménages seraient rigoureusement choisis au hasard. On a donc mis au point une simulation sur ordinateur qui a été effectuée au Siège de l'OMS par le service Méthodologie épidémiologique et statistique de la Division de la Surveillance épidémiologique et de l'Appréciation de la Situation sanitaire et de ses Tendances, en coopération avec le PEV et le Programme de lutte contre les maladies diarrhéiques (LMD). Les résultats de cette étude, qui sont présentés en détail dans un autre rapport, sont résumés ici. Des populations ont été constituées de façon à représenter une densité démographique faible, moyenne, élevée ou très élevée. A l'intérieur de ces populations, les ménages ont été répartis de l'une des 4 façons suivantes: au hasard, le long d'une route ou d'un cours d'eau, autour d'un foyer unique (par exemple un marché, une école ou une église) ou autour de 2 foyers. Un ménage sur 7 a été désigné comme comportant un enfant du groupe d'âge considéré, et l'on a fixé la proportion globale des enfants supposés «vaccinés». Pour repérer les enfants vaccinés, on a procédé de 4 façons: au hasard, parmi tous les enfants, en se limitant à un «îlot» ou la totalité des enfants sont supposés vaccinés tandis qu'aucun ne l'est à l'extérieur, ou à 2 îlots puis, finalement, à 5 îlots où le taux de couverture est toujours supposé égal à 100%. Dans les diverses strates, la couverture vaccinale globale a été fixée à 10, 30, 50, 70 ou

ations with coverage probabilities of 17%, 34%, 46%, 65% or 75%. Once the clusters were established, sampling took place either by selecting a single household as a starting point and then visiting adjacent households until 7 children had been obtained, or by selecting 7 random starting points and visiting adjacent households until 1 child was obtained. The sampling process was repeated 500 times for each of the established clusters.

The results of the simulation can be summarized as follows. Comparison of the EPI estimates to the SRS estimates over the 500 replications within the constructed populations suggests that the EPI estimates have greater bias and are more variable than the corresponding SRS estimates. However, with respect to the stated goal of the EPI to estimate the true immunization coverage to within 10 percentage points of the true population proportion with 95% confidence, the simulation provides strong evidence that this is being accomplished. Over the 500 replications within each of the artificially generated populations, the estimated coverage rate differed from the actual level by more than 10 percentage points in less than 5% of the replications.

At the individual cluster level, the EPI method was typically outdone by the SRS method. This was particularly evident for those clusters having pockets of immunized children where the EPI method demonstrated high bias and extremely high variability which was not the case with the SRS method. This can be explained as follows: if the randomly selected starting household was far from the pocket of immunized children, the estimated proportion with the EPI method would then be much too low, whereas if it was in or close to the pocket, the estimate would be much too large. Since the SRS method selects 7 random starts, pocketing did not result in the same errors of estimation.

It should be reemphasized that the EPI uses these surveys for overall population estimates and has never advocated estimation at the individual cluster level. Hence, even though the individual cluster estimates may, under some conditions, be poor, the averaging process over the 30 clusters produced estimates accurate to within 10 percentage points of the true level of immunization coverage in the vast majority of replications. This certainly provides reassuring testimony concerning the coverage surveys as currently being carried out. It also sets forth a warning to those users who might disaggregate the estimates for the purpose of making statements concerning particular communities, cities or regions. This disaggregation represents an unacceptable procedure within the cluster sampling framework.

Alternatives to the 30 x 7 EPI sampling strategy

Further research into possible alternatives to the currently used 30 x 7 EPI survey is certainly warranted and should be of great value to a wide range of investigators. Several alternatives have been proposed to date. For example the Diarrhoeal Diseases Control Programme at WHO is currently testing survey designs which divide a population into M clusters of roughly equal size. If the clusters are large, they are subdivided into smaller sets, each with the desired number of individuals. In these M clusters, m are selected at random and all individuals residing in those selected clusters are studied. This eliminates the need to select n_i from the N_i in the selected clusters and, as a result, eliminates the need to incorporate a "within cluster" term into the variance estimate. In order for this method to be used, it is necessary for good estimates of population size to exist prior to

90% ce qui, pour l'ensemble des 30 grappes, aboutit à une couverture probable de 17, 34, 46, 65 ou 75%. Une fois les grappes constituées, l'échantillonnage a été effectué soit en choisissant un ménage unique comme point de départ puis en se rendant dans les ménages immédiatement voisins jusqu'à l'obtention de 7 enfants, soit en choisissant au hasard 7 points de départ et en procédant de même jusqu'à l'obtention de 1 enfant. Cette méthode d'échantillonnage a été répétée 500 fois pour chacune des grappes préalablement constituées.

Les résultats de la simulation peuvent être résumés comme suit. En comparant les estimations PEV aux estimations SAS pour l'ensemble des 500 épreuves conduites en double dans les populations théoriques ci-dessus, il semble que les estimations PEV comportent un biais plus important et une plus grande dispersion que les estimations SAS correspondantes. Cependant, si l'on considère l'objectif que s'est fixé le PEV, à savoir estimer la couverture vaccinale réelle à 10 points près, en plus ou moins (au niveau de confiance de 95%), la simulation donne tout lieu de penser que ce but est atteint. Pour l'ensemble des 500 calculs effectués en double dans chacune des populations théoriques, le taux de couverture estimatif ne s'écarte en effet de la valeur réelle, de plus de 10 points que dans moins de 5% des cas.

Au niveau des grappes individuelles, la méthode SAS l'emporte le plus souvent sur la méthode PEV. C'est particulièrement net pour les grappes comportant des îlots d'enfants tous vaccinés, la méthode PEV faisant apparaître un biais important et une dispersion extrêmement forte, contrairement à l'autre méthode. Ce résultat peut s'expliquer comme suit. Quand le ménage de départ, choisi au hasard, se trouve loin de l'îlot, la proportion estimative obtenue avec la méthode PEV est a priori beaucoup trop faible tandis que lorsque ce ménage se situe à l'intérieur ou à proximité de l'îlot, on aboutit à une estimation beaucoup trop forte. Comme la méthode SAS repose sur le choix au hasard de 7 points de départ, la présence d'îlots n'entraîne pas les mêmes erreurs d'estimation.

Il faut de nouveau insister sur le fait que le PEV se sert de ce type d'enquêtes pour obtenir des estimations globales, pour l'ensemble d'une population, et n'a jamais préconisé une estimation au niveau des grappes individuelles. Dans ces conditions, même si les estimations à ce dernier niveau risquent, dans certaines circonstances, d'être médiocres, le calcul d'une moyenne pour l'ensemble des 30 groupes fournit dans la très grande majorité des épreuves effectuées en double, des estimations qui ne diffèrent pas de plus de 10 points de la couverture vaccinale effective. Cette conclusion est rassurante quant à la valeur des enquêtes actuelles sur la couverture vaccinale. Elle constitue en outre une mise en garde pour ceux qui seraient tentés de ventiler les estimations afin d'arriver à des conclusions au sujet de telle ou telle collectivité, ville ou région. Cette ventilation est à proscrire dans le cadre de l'échantillonnage par grappes.

Solutions de rechange à la méthode d'échantillonnage 30 x 7 du PEV

La recherche de solutions de rechange par rapport à la méthode des enquêtes 30 x 7 actuellement appliquée par le PEV est certainement justifiée, et la poursuite des travaux dans ce sens devrait présenter un vif intérêt pour toute une série de chercheurs. Plusieurs solutions ont été proposées à ce jour. Par exemple, le groupe du Programme de lutte contre les maladies diarrhéiques de l'OMS expérimente actuellement un type d'enquête où la population est répartie en M grappes d'effectifs sensiblement uniformes. Quand ces grappes sont importantes, elles sont subdivisées en sous-populations comptant chacune le nombre de sujets souhaité. Parmi ces M grappes, on en choisit m au hasard et l'on étudie dans ces grappes tous les résidents. On évite ainsi d'avoir à choisir n_i sujets parmi les N_i membres des grappes retenues et, de ce fait, il devient inutile d'incorporer dans l'estimation de la variance un terme tenant compte de la dispersion à l'intérieur

ation and selection of clusters. This method is appropriate to the objective of the CDD since the desired sample size is large. It would not be appropriate for coverage or other surveys requiring smaller sample sizes.

Another sampling method which has been proposed which needs much field testing is known as the "Lot Quality Assurance Sampling" (LQAS) technique. In effect, used in an immunization context, this method can be used to say whether or not the immunization coverage is below a predetermined target level. It has the advantage of being able to provide a programme manager with a local estimate of the effectiveness of his/her programme and, as such, could be most effective as a management tool in specific health districts or towns.

With the LQAS method a target level of coverage is specified for the population. A random sample of individuals in the population is then selected and, based on this sample, the population (or "lot") is "rejected" if it is determined that the coverage is below the target, or is "accepted" if the coverage achieves the target. The method parallels simple quality assurance methods used traditionally in industrial applications and is based upon the binomial distribution.

Specifically, let d = number of unimmunized children out of n children sampled and let Q = threshold value or the proportion of children unimmunized in the population. One may wish to test the null hypothesis $H_0: Q \geq .50$ (i.e., the population is not adequately immunized) versus the alternative $H_A: Q < .50$ (i.e., the population is adequately immunized). Note that the null hypothesis is that the proportion of unimmunized people in the population is at least as large as the threshold. If the null hypothesis H_0 is rejected, it can be concluded that the level of coverage is adequate. Rejection of H_0 is equivalent to "accepting the lot" from a quality assurance point of view and, similarly, failing to reject H_0 is equivalent to "rejecting the lot". The type I error in this situation is the probability that the population is deemed adequately immunized on the basis of the survey when, in fact, the proportion immunized is actually less than Q . This very serious error can easily be controlled by choosing a small type I error rate. The type II error is the probability that the population is deemed not adequately immunized when in fact it is. For actual policy decisions, this would not appear to be as serious as an error of the first type.

The strategy with LQAS is to select a sample of n children, determine whether or not each has been immunized and, defining d as number of unimmunized children among the n studied. If $d \leq d^*$, H_0 is rejected and the lot is accepted; if $d > d^*$, H_0 is accepted and the lot is rejected. If the lot is rejected, a new immunization effort might be initiated. Choice of d^* and n depend upon the desired type I and type II error probabilities.

As an example, a district health officer wishes to know whether immunization coverage in his/her district is above 50%. To make this decision, a random sample of children in the target age group is selected and their individual immunization status is determined. (It should be noted here that the process of selecting a random sample will present a difficult barrier and, as a result, the LQAS method may not be practicable for even moderately large populations.) If the number of unimmunized children in the sample is less than or equal to a computed threshold (d^*), the lot is accepted and it is concluded that the coverage exceeds the target level. Otherwise, the lot is rejected. Specifically, if a random sample of $n = 50$ children is taken and, for a type I error of .05, the rule would be to reject H_0

de la grappe. Pour pouvoir procéder ainsi, il faut disposer d'une bonne estimation de l'effectif de la population avant de constituer puis de choisir les grappes. Cette méthode est adaptée à l'objectif du Programme LMD puisque la taille souhaitée pour l'échantillon est importante. Elle ne conviendrait pas pour des enquêtes comme celles sur la couverture vaccinale qui exigent des échantillons de taille moins grande.

Une autre méthode d'échantillonnage qui a été proposée, mais qui devra faire l'objet d'une expérimentation beaucoup plus poussée sur le terrain, est connue sous le nom de technique de « Sondage de contrôle de la qualité des lots » (LQAS). En fait, dans le contexte de la vaccination, cette méthode permet de savoir si la couverture vaccinale est ou non inférieure à un niveau visé, fixé à l'avance. Elle a l'avantage de fournir au responsable d'un programme une estimation locale de l'efficacité de ce programme et, à ce titre, pourrait constituer un excellent outil de gestion dans des villes ou arrondissements sanitaires déterminés.

Dans la méthode LQAS, on fixe pour la population en cause une couverture à atteindre. On choisit alors au hasard un échantillon d'individus dans cette population et, sur la base de cet échantillon, on « rejette » la population (ou « lots ») si la couverture est inférieure au niveau visé ou on l'« accepte » dans le cas contraire. Cette méthode est analogue aux techniques simples d'assurance de la qualité qui sont traditionnellement utilisées dans l'industrie et elle repose sur la loi binomiale.

Plus précisément, soit d le nombre d'enfants non vaccinés parmi les n enfants constituant l'échantillon et Q le seuil ou proportion d'enfants non vaccinés dans la population. On cherche à tester l'hypothèse nulle $H_0: Q \geq 0,50$ (qui correspond à une couverture vaccinale insuffisante) par comparaison à l'hypothèse alternative $H_A: Q < 0,50$ (couverture vaccinale suffisante). On notera que l'hypothèse nulle correspond au cas où la proportion des personnes non vaccinées dans la population est au moins aussi élevée que le seuil. Si l'hypothèse nulle H_0 est rejetée, on en déduit que la couverture vaccinale est suffisante. Le rejet de l'hypothèse nulle équivaut, du point de vue de l'assurance de la qualité, à l'« acceptation du lot » et, inversement, ne pas pouvoir rejeter l'hypothèse nulle équivaut à « rejeter le lot ». En pareille situation, l'erreur de première espèce correspond à la probabilité de juger la couverture vaccinale suffisante au vu d'une enquête alors que, en réalité, la proportion réelle des sujets vaccinés est inférieure à Q . Pour éviter cette erreur, aux conséquences très graves, il suffit de choisir une faible valeur pour le risque de première espèce. L'erreur de deuxième espèce correspond à la probabilité de juger insuffisante une couverture vaccinale qui est en réalité suffisante. En pratique, ce type d'erreur semble moins grave que le premier pour le choix de la politique générale.

Avec la méthode LQAS, on procède comme suit: après avoir choisi un échantillon de n enfants, on détermine pour chacun d'eux s'il a ou non été vacciné et l'on appelle d le nombre d'enfants non vaccinés dans cet échantillon. Si $d \leq d^*$, l'hypothèse nulle est rejetée et le lot est accepté; au contraire si $d > d^*$, l'hypothèse nulle est acceptée et le lot rejeté. Dans ce dernier cas, on peut envisager d'intensifier l'effort de vaccination. Le choix de d^* et de n dépend de la valeur fixée pour les erreurs de première et de seconde espèce.

Par exemple, supposons que le responsable d'un arrondissement sanitaire veuille savoir si la couverture vaccinale y est supérieure à 50%. Pour cela, on choisit au hasard un certain nombre d'enfants dans le groupe d'âge considéré et l'on étudie la situation vaccinale de chacun d'eux. (On notera que, dans ce cas, le choix d'un échantillon aléatoire soulève de réelles difficultés et que, de ce fait, la méthode LQAS risque d'être impraticable même dans une population relativement peu nombreuse.) Si le nombre d'enfants non vaccinés dans l'échantillon est inférieur ou égal à un seuil défini par le calcul (d^*), le lot est accepté et la conclusion est que la couverture dépasse le taux visé. Dans le cas contraire, le lot est rejeté. Plus précisément, supposons qu'on prélève un échantillon aléatoire de $n = 50$ enfants; le critère à appliquer, dans l'hy-

(and accept the lot) if $d \leq d^* = 18$ since, then, $Pr(d \leq d^* | P = 0.50) \leq 0.05$. If it were possible to select a sample of $n = 100$, the criterion would be to reject H_0 if $d \leq 41$. Clearly, larger sample sizes will result in smaller type II errors against a range of alternatives.

Confidence interval estimates can be constructed within each of the areas for which the LQAS sample was selected and, using straightforward stratified sampling theory, estimates can be produced for the totality of the areas sampled. It should be noted that this is not a cluster sampling strategy and that LQAS is not likely to be a substitute for a sampling scheme which attempts to estimate a parameter for a large population.

Conclusions

In many situations the use of sample surveys is essential to provide some of the data necessary for the management of health programmes. The adoption of survey methodology by national programmes should be encouraged and supported. At the same time, there is great danger in taking for granted the idea that because a particular survey design works for one stated objective, that this same design can be successfully employed in other contexts. It is essential that prospective users tailor the available survey designs to suit their particular needs and, at very least, recognize the limitations inherent in the use of any survey methodology adopted — particularly when practical considerations force divergence from the ideal path. Programme managers cannot expect reliable results from surveys if the methodology is not adhered to and if the underlying assumptions of the sampling strategy are not pertinent to the survey in question.

Acknowledgements

The authors would like to acknowledge the assistance of the Director and staff of the Expanded Programme on Immunization, the Diarrhoeal Diseases Control Programme, and the Division of Health Situation and Trend Assessment at WHO Headquarters for their review and suggestions in the preparation of this document.

REFERENCES — RÉFÉRENCES

HENDERSON, R. H. & SUNDARESAN, T. Cluster sampling to assess immunization coverage: a review of experience with a simplified sampling method. *Bulletin of the World Health Organization*, **60**: 253-260 (1982).

WORLD HEALTH ORGANIZATION. *Training for mid-level managers. Evaluate vaccination coverage*. Geneva, WHO Expanded Programme on Immunization in cooperation with US Department of Health and Human Services, Public Health Service, Center for Disease Control, 1979.

ORGANISATION MONDIALE DE LA SANTÉ. *Formation des cadres moyens. Evaluer la couverture vaccinale*. Genève, OMS Programme élargi de vaccination avec US Department of Health and Human Services, Public Health Service, Center for Disease Control, 1979.

SERFLING, R. E. & SHERMAN, I. L. *Attribute sampling methods*. Washington, DC., US Department of Health and Human Services, Public Health Service, 1975. (Publication No. 1230).

HENDERSON, R. H. ET AL. Assessment of vaccination cover-

pothèse d'une erreur de première espèce fixée à 0,05, est de rejeter l'hypothèse nulle H_0 (et donc d'accepter le lot) si $d \leq d^* = 18$ car, alors, $Pr(d \leq d^* | P = 0,50) \leq 0,05$. Si l'on pouvait choisir un échantillon de taille $n = 100$, le critère serait de rejeter l'hypothèse nulle quand $d \leq 41$. Il est clair que l'emploi d'un échantillon de plus grande taille entraîne la diminution des erreurs de seconde espèce par rapport à toute une série d'hypothèses alternatives.

Pour chacune des régions dans laquelle on a prélevé un échantillon LQAS, on peut déterminer les intervalles de confiance estimatifs et, en s'appuyant directement sur la théorie de l'échantillonnage stratifié on obtient une estimation pour l'ensemble des zones étudiées par sondage. On notera qu'il ne s'agit pas là d'une méthode d'échantillonnage par grappes et que la méthode LQAS ne peut certainement pas remplacer une méthode d'échantillonnage visant à estimer un paramètre relatif à une population nombreuse.

Conclusions

Dans de nombreux cas, les enquêtes par sondage constituent un moyen essentiel pour obtenir certaines des données qui sont indispensables à la gestion des programmes de santé. L'adoption d'une méthodologie des enquêtes par les programmes nationaux est à encourager et à soutenir. D'un autre côté, il serait fort hasardeux d'admettre sous prétexte qu'un type particulier d'enquête donne de bons résultats pour un objectif déterminé, que la même méthodologie peut être employée avec succès dans un contexte différent. Il est essentiel que les utilisateurs potentiels adaptent la conception des enquêtes de type connu à leurs besoins particuliers et que, pour le moins, ils soient conscients des limites inhérentes à l'emploi de n'importe quelle méthodologie — en particulier quand des considérations pratiques forcent à s'écarter de la procédure idéale. Les responsables de programme ne sauraient attendre de résultats fiables des enquêtes si la méthodologie n'est pas respectée et si les hypothèses sur lesquelles reposent la technique d'échantillonnage ne peuvent être remplies dans le cadre de l'enquête considérée.

Remerciements

Les auteurs tiennent à remercier de leur aide au Siège de l'OMS, le Directeur et le personnel du Programme élargi de vaccination, du Programme de lutte contre les maladies diarrhéiques et de la Division de la surveillance épidémiologique et de l'appréciation de la situation sanitaire et de ses tendances qui ont bien voulu revoir l'avant-projet du présent article et formuler des suggestions pour sa mise au point.

age, vaccination scar rates, and smallpox scarring in five areas of West Africa. *Bulletin of the World Health Organization*, **48**: 183-194 (1973).

5. Expanded Programme on Immunization. Global status report. *Weekly epidemiological record*, **58** (23): 173-180 (1983).

Programme élargi de vaccination. Rapport sur la situation mondiale. *Relevé épidémiologique hebdomadaire*, **58** (23): 173-180 (1983).

6. WORLD HEALTH ORGANIZATION. *Eastern Mediterranean Region/South-East Asian Region Meeting on the prevention of neonatal tetanus*. Lahore, 1982. (EMRO technical publication No. 7, SEARO technical publication No. 3).

7. DIRECTORATE GENERAL OF HEALTH SERVICES, INDIA. *Guidelines. Combined survey to estimate the incidence of neonatal tetanus and poliomyelitis*. New Delhi, Ministry of Health, 1981.

8. LEVY, P. S. & LEMESHOW, S. A. *Sampling for health professionals*. Belmont, Cal., Lifeline Learning Publications, 1980.

204