Tropical Medicine and International Health doi:10.1111/j.1365-3156.2006.01799.x

VOLUME 12 NO 4 PP 475–484 APRIL 2007

How to measure the prevalence of tuberculosis in a population

Marieke J. van der Werf[1,2] and Martien W. Borgdorff[1,2]

1 *KNCV Tuberculosis Foundation, The Hague, The Netherlands*
2 *Department of Infectious Diseases, Tropical Medicine & AIDS, University of Amsterdam, The Netherlands*

Summary The World Health Assembly has defined targets for measuring performance of tuberculosis (TB) control programmes, which have been included in the framework of the Millennium Development Goals. To be able to measure progress towards these targets TB control programmes need epidemiological information. At the moment, surveillance data of countries with a high burden of TB are insufficient to assess performance of the TB control programme as a result of incompleteness or low quality. Several high-burden countries have performed TB prevalence surveys to obtain epidemiological information. As a standardized method for TB prevalence surveys has not been defined, we discuss the different options for measuring prevalence and their advantages and disadvantages. Most surveys use comparable strategies. Alternatives strategies at lower cost need to be evaluated.

keywords Tuberculosis, Burden of disease, prevalence survey, Millennium Development Goals, methodology

Preamble

Prof Martien W Borgdorff was awarded the Eijkman Medal on October 12, 2005 for his special contribution at both national and international level to the epidemiology and control of tuberculosis. He is a member of the WHO Strategic and Technical Advisory group and chairman of the Tuberculosis Surveillance and Research Unit of WHO.

The Eijkman Medaille Foundation (E.M.F.) was established in 1923 in honour of Christian Eijkman during his 25th anniversary as professor of Hygiene at the University of Utrecht, The Netherlands. Eijkman was renowned for his research in the former Dutch East Indies (now Indonesia), which lead to the elucidation of the cause of beri beri, for which he received the Nobel Prize for Physiology and Medicine in 1929.

Prof Hugo J van der Kaay
President, Eijkman Medal Foundation, The Netherlands

Introduction

In 1991, two targets for tuberculosis (TB) control were defined by the World Health Assembly (WHA): to detect 70% of all new sputum smear-positive cases arising each year and to successfully treat 85% of these cases. In 2000, the United Nations formulated the Millennium Development Goals (MDG) with target 8 being 'Have halted by 2015 and begun to reverse the incidence of malaria and other major diseases'. The indicators for measuring progress to this target for TB are prevalence and death rates associated with TB (indicator 23) and proportion of TB cases detected and cured under DOTS (indicator 24)(http://unstats.un.org/unsd/mi/mi_goals.asp, accessed on 22 March 2006). To be able to assess progress towards the WHA and the MDG targets TB control programmes need epidemiological information about the burden of TB (Attaran 2005).

In most countries with a high burden of TB and poor resources, routine surveillance and vital registration are weak and incomplete or even nonexistent. For instance, World Health Organization (WHO) estimated that in 2004 only 53% of global TB cases were notified (WHO 2006). Therefore, additional epidemiological information is required. The most valuable information is TB incidence. Unfortunately measuring TB incidence is difficult. It requires two TB prevalence surveys conducted with a short interval between the two surveys and rigorous monitoring of the study population to detect new cases that die or migrate out in between the two prevalence surveys. Owing to these requirements, no country measures TB incidence regularly. Alternatively,

tuberculin surveys are performed which measure the prevalence of tuberculous infection and provide information about the annual risk of tuberculosis infection (ARTI). The ARTI can be used to estimate the incidence of TB by assuming a fixed relationship between the two (Styblo 1985). This relationship has been assessed before DOTS was introduced and HIV infection was prevalent. Therefore, it may no longer apply as DOTS is introduced in most TB high-burden countries and HIV infection has become epidemic. In countries where multiple registers with TB patient data are available, under-reporting and the corrected incidence of TB can be estimated by the capture–recapture method (Baussano *et al.* 2006; Guernier *et al.* 2006).

Studies that measure the prevalence of TB have been performed by several TB high-burden countries (HBC) and provide extremely useful information for monitoring progress towards the targets. Because prevalence surveys seem to be the best alternative at present for measuring progress we will discuss strategies that can be applied for measuring prevalence. We start with summarizing issues of study design in recent TB prevalence surveys, i.e. performed after 1990, in general populations. Thereafter, we discuss the advantages and disadvantages of the applied methods. Finally, we provide recommendations for performing TB prevalence surveys.

Study design

Case definition

The design of a TB prevalence survey depends on the case definition. The main objective of most recently performed studies was to determine the prevalence of smear-positive TB. Definitions used for a smear-positive case differed substantially: at least one positive smear (one acid-fast bacilli (AFB) found during examination of 100 fields) (China Tuberculosis Control Collaboration 2004); one positive smear (>3 AFB) (Gopi *et al.* 2003); at least one slide with ≥4 AFB/100 fields (Hamid Salim *et al.* 2004); at least two sputum specimens positive for AFB or one sputum positive for AFB and radiological abnormalities consistent with TB (Zaman *et al.* 2006); and at least two slides positive (Soemantri & Senewe 2005). In two surveys, a case definition was not provided (Hong *et al.* 1998; Tupasi *et al.* 1999). One survey used the case definition of WHO (WHO 2003; NTP Cambodia 2005). For other surveys, it was not specified whether they used a special definition, the case definition of the TB programme or the WHO definition.

The outcome of one survey was bacillary-positive TB cases, i.e. sputum positive by smear (≥1 AFB on Ziehl-Neelsen (ZN) and/or ≥4 AFB on Fluorescence microscopy (FM)) and/or yielding *Mycobacterium tuberculosis* by culture (Datta *et al.* 2000).

Sample size

The main difference between TB and other infectious diseases such as HIV and malaria is that the prevalence of TB is much lower. Therefore, it is a challenge in TB prevalence surveys to accurately identify TB cases in the study population. To illustrate the size of the task, in 2003, the estimated prevalence of all cases of TB in the 22 countries with most new cases arising each year (HBC) ranged from 77/100 000 to 888/100 000 population (WHO 2006). Most surveys concentrate on identifying sputum smear-positive cases and approximately 45% of the prevalent cases is expected to be smear-positive (Dye *et al.* 1999). Thus, in the 22 HBC we can expect 35 to 399/100 000 population to be smear-positive cases. To be able to measure a prevalence of a disease with such a relatively low frequency with adequate precision a large sample size is required. In recent national prevalence surveys, study populations of 21 960 to 365 097 individuals were included (Tupasi *et al.* 1999; China Tuberculosis Control Collaboration 2004) (Table 1). Required sample sizes were calculated using an expected prevalence of sputum smear-positive TB which was obtained from previous surveys (Hong *et al.* 1998), WHO estimates (NTP Cambodia 2005) or an unspecified source. In general, a precision of 20–25% and a confidence level of 95% was used to calculate the sample size. Moreover, the sample size was increased anticipating a compliance or coverage of the study population between 75% and 90%. Furthermore, all surveys used multi-stage cluster sampling (see next section), which provides a prevalence estimate less accurate than if simple random sampling is used because individuals within a cluster will be more alike than individuals from different clusters. Thus, the sample size was multiplied by a factor 1.25 to 2 to take into account this design effect.

Sampling

As it is almost impossible to use simple random selection of individuals for inclusion of the required large sample sizes all surveys used cluster sampling. Two methods were used to select districts or villages: selection proportional to population size (Tupasi *et al.* 1999; China Tuberculosis Control Collaboration 2004; NTP Cambodia 2005) and random sampling (Gopi *et al.* 2003). One study used a combination of the two methods (Datta *et al.* 2000). Two surveys were conducted in the context of other surveys, the

Table 1 Summary of recently (after 1990) conducted tuberculosis (TB) prevalence survey in general populations

Country	Number participating in survey	Symptoms or chest X-ray abnormalities (%)	Number providing sputum samples (%)*	Reference
Korea	64 713	5462 (8.4)	5368 (98.3)	(Hong *et al.* 1998; MoH Korea 1995)
Philippines	12 850	1619 (12.6)	1390 (85.9)	(Tupasi *et al.* 1999)
India	64 077	8229 (12.8)	8032 (97.6)	(Datta *et al.* 2000)
India	83 390	10 446 (12.5)	10 028 (96.0)	(Gopi *et al.* 2003)
Bangladesh	223 936	7001 (2.6)		(Hamid Salim *et al.* 2004)
China	365 097	4667 (1.3)	4558 (97.7)	(China Tuberculosis Control Collaboration 2004)
Cambodia	30 032	Cough 1614 (7.3) Chest X-ray 2406 (10.9)	Cough 1465 (90.8) Chest X-ray 2272 (94.4)	(NTP Cambodia 2005)
Indonesia	50 154	4158 (8.3)	n.p.	(Soemantri & Senewe 2005)
Bangladesh	59 359	4235 (7.1)	3834 (90.5)	(Zaman *et al.* 2006)

*n.p., not provided in publication.

survey in Matlab in Bangladesh included all individuals participating in a Health and Demographic Surveillance System (Zaman *et al.* 2006) and the survey in Indonesia used a subsample of the households included in the National Household Health Survey (Soemantri & Senewe 2005). Most studies stratified districts or villages into urban or rural before sampling. If initially districts were sampled simple random sampling was used thereafter to select villages or subdistricts. In the selected villages or sub districts, clusters comprised a fixed number of individuals: 600 in the Philippines (Tupasi *et al.* 1999) and 720 in Cambodia (NTP Cambodia 2005); or the total population of a village, town street or subdistrict (Hong *et al.* 1998; Datta *et al.* 2000; Gopi *et al.* 2003; Hamid Salim *et al.* 2004; China Tuberculosis Control Collaboration 2004).

In the selected clusters, households were visited by the field teams. In surveys where a fixed number of individuals were included, households or groups of households were selected by random sampling. This requires that all households or groups of households in the selected village or subdistrict be identified before selection can be performed.

Study population

Individuals eligible for inclusion in the survey were those who are members of the selected households (Tupasi *et al.* 1999; Gopi *et al.* 2003; Hamid Salim *et al.* 2004; Zaman *et al.* 2006), or permanent residents of the selected villages or town streets (Datta *et al.* 2000). In two surveys, the local authority office was requested to prepare a name list of all individuals living in the household a few weeks before the survey (Hong *et al.* 1998; NTP Cambodia 2005). All individuals on this list were considered eligible for inclusion in the survey.

As TB prevalence surveys focus on identification of bacteriological-positive TB patients children were often not included in the study population. This is because bacteriological-positive TB is rarely diagnosed in children. Furthermore, it is difficult to collect sputum especially of very young children. The age above which individuals were included in the survey differed from ≥5 years (Hong *et al.* 1998); ≥10 years (Tupasi *et al.* 1999; NTP Cambodia 2005); ≥12 years (Hamid Salim *et al.* 2004); to ≥15 years (Datta *et al.* 2000; Gopi *et al.* 2003; China Tuberculosis Control Collaboration 2004; Zaman *et al.* 2006).

Case finding

To limit the workload of sputum examination by the laboratory, all surveys first screened the study population to identify individuals with a high risk of bacteriological-positive pulmonary TB (suspects) and to exclude those with an extremely low risk of bacteriological-positive pulmonary TB. Sputum samples were only collected from those with a relatively high risk, i.e. suspects. Three methods for screening were used: identification of suspects based on the presence of

(1) symptoms (Hamid Salim *et al.* 2004; Soemantri & Senewe 2005; Zaman *et al.* 2006);
(2) abnormal chest X-ray (Hong *et al.* 1998; Tupasi *et al.* 1999); or
(3) symptoms or abnormal chest X-ray (Datta *et al.* 2000; Gopi *et al.* 2003; China Tuberculosis Control Collaboration 2004; NTP Cambodia 2005).

Symptoms or combinations of symptoms that were used to identify suspects were cough for 2 weeks or more, chest pain for 1 month or more, fever for 1 month or more, or haemoptysis within the last 6 months (Datta *et al.* 2000);

chest symptoms (Gopi *et al.* 2003); cough for at least 3 weeks (Hamid Salim *et al.* 2004; Zaman *et al.* 2006); persistent cough (China Tuberculosis Control Collaboration 2004); coughing up sputum or blood over the past 1 month (Soemantri & Senewe 2005); and cough lasting for 3 weeks or more and/or blood-contained sputum (NTP Cambodia 2005). The percentage of individuals eligible for sputum examination ranged from 1.3% to 13% (Table 1).

Suspects were requested to provide one (Hamid Salim *et al.* 2004), two (Datta *et al.* 2000; Gopi *et al.* 2003; NTP Cambodia 2005) or three sputum samples (Hong *et al.* 1998; Tupasi *et al.* 1999; China Tuberculosis Control Collaboration 2004; Soemantri & Senewe 2005; Zaman *et al.* 2006). Sometimes, only spot specimens (sample provided on the spot) were collected (Hong *et al.* 1998). Other surveys focused on collection of morning samples (sample produced immediately after waking up) (Hamid Salim *et al.* 2004) or spot and morning samples (Tupasi *et al.* 1999; Datta *et al.* 2000; China Tuberculosis Control Collaboration 2004; NTP Cambodia 2005; Soemantri & Senewe 2005; Zaman *et al.* 2006). One study used sputum induction by nebulized supersaturated saline aerosol for suspects unable to expectorate spontaneously (Tupasi *et al.* 1999). The percentage of suspects providing sputum ranged from 86% to 98% (Table 1).

Data collection

The procedures for data collection depended on the method used for screening of suspects. If only symptoms were used to identify suspects, field workers performed house-to-house visits during which household members were interviewed using a structured questionnaire (Hamid Salim *et al.* 2004; Soemantri & Senewe 2005; Zaman *et al.* 2006). In cases where chest X-ray was used for identification of suspects, registration and sometimes interview of subjects in the cluster was done through house-to-house visits after which those eligible for chest X-ray examination were invited to have a chest X-ray taken (Hong *et al.* 1998; Tupasi *et al.* 1999; Datta *et al.* 2000; Gopi *et al.* 2003; NTP Cambodia 2005).

All surveys collected information about sex and age. Additionally, information about marital status, financial income, TB symptoms, TB history, BCG vaccination status, contact with TB patients and awareness and healthcare-seeking behaviour for TB was sometimes collected.

Those identified as suspects were requested to submit sputum sample(s) by the survey team (Tupasi *et al.* 1999; Datta *et al.* 2000; Gopi *et al.* 2003; Hamid Salim *et al.* 2004), were visited by a laboratory technician who collected the samples under supervision (Soemantri & Senewe 2005; NTP Cambodia 2005) or were referred to a clinic for sputum examination (Zaman *et al.* 2006).

Sputum handling and examination

Sputum samples were examined on the spot and at a provincial or central laboratory (Hong *et al.* 1998; Datta *et al.* 2000), or transported to a local laboratory (Hamid Salim *et al.* 2004) or a research laboratory (Tupasi *et al.* 1999; NTP Cambodia 2005). Alternatively, the samples were sent to a local healthcare facility where slides were prepared from the sample, and the slides were forwarded to the provincial health laboratory for examination (Soemantri & Senewe 2005). In one survey, the patient instead of the sample was sent to the local healthcare facility (Zaman *et al.* 2006). Information about methods of transportation is limited. Two surveys mentioned to ship samples on ice to the laboratory in 24–72 hours (Tupasi *et al.* 1999; NTP Cambodia 2005).

All surveys used microscopy to examine the sputum samples. Both fluorescence microscopy (FM) and examination by light microscopy after Ziehl Neelsen staining (ZN) were used. Sometimes, both methods were applied in one survey (Tupasi *et al.* 1999; Datta *et al.* 2000). All but the two surveys in Bangladesh also used culture techniques to examine samples. The surveys not using culture techniques had 10% of the smears reread at the reference laboratory for quality assurance (Hamid Salim *et al.* 2004; Zaman *et al.* 2006). In some surveys, it was specified that a positive smear had to be confirmed by a second microscopist (Tupasi *et al.* 1999; Datta *et al.* 2000; China Tuberculosis Control Collaboration 2004) or by checking two additional sputum samples (Hamid Salim *et al.* 2004).

Advantages and disadvantages of applied methods

Sample size

Obviously, a larger sample size will provide a more accurate estimate of the prevalence of TB. However, most surveys are restricted by the amount of money, human resources and time available. Furthermore, if the number of individuals is sufficient to provide an accurate estimate, it may be more useful to focus on quality of data collection instead of increasing the number of included individuals.

Sampling

The most efficient sampling strategy is probability proportional to size sampling to select districts, thereafter

selection of villages by simple random sampling and inclusion of a fixed number of individuals from each village. Selection of districts and villages by simple random sampling and inclusion of a number of individuals from each district proportional to the population size of the district will also provide a self-weighting sample but is less efficient. In practice, self-weighting is not achieved fully. There will be variation in the number of individuals included per village even if it is planned to include a fixed number. Therefore, adjustments are required in the analysis to correct for unequal probabilities of selection and to achieve absence of bias in the final prevalence estimate, i.e. the individual sampling probability (probability of a person to be included in the sample) needs to be calculated and the individuals should be weighted by this.

The advantage of stratification is that separate estimates with a pre-defined precision can be obtained for different strata (e.g. regions or urban *vs.* rural). Moreover, stratification may increase overall precision, if prevalence varies strongly between strata. However, stratification over two or more sampling stages will complicate the survey design.

Study population

Defining the study population as those who are members of the selected household or permanent residents of the selected villages or town streets may work well in countries with limited internal migration. In other areas, individuals may be considered members of a household even though they have not been present in the household for quite some time. Thus, these persons are not present to participate in the survey and will be recorded as absent. They will also not have a chance to be included in the survey in the location where they are residing at the moment of the survey as they are not members of the household there but visitors. If this occurs frequently, it may bias the survey. Alternatively, only individuals who resided in the household the night before the survey can be included. This will result in fewer absent individuals and limit bias. However, there is a concern that, especially if chest X-ray is used for identification of suspects, individuals from neighbouring areas may come to the village included in the survey to participate in the survey and have a free chest X-ray taken.

Surveys that included individuals <15 years of age identified few smear-positive cases in this group. Prevalence of smear-positive TB was only 12% of the total prevalence in those aged 5–19 in Korea (Hong *et al.* 1998), 18% in those aged 10–14 years in the Cambodia survey (NTP Cambodia 2005), 28% in those aged 10–19 years in Philippines (Tupasi *et al.* 1999) and zero cases were identified in the 12–14-year age group in Bangladesh (Hamid Salim *et al.* 2004). Global notification data of 2004 show that the notification rate of new smear-positive cases in the 0–14-year age group is 5.7% of the total notification rate (WHO 2006). Thus, exclusion of individuals <15 years of age will not substantially affect the prevalence estimate whereas it decreases the workload considerably.

Case finding

There are two main advantages of screening populations for suspects. First, the number of individuals who are requested to provide sputum is limited (1.3–13% of all individuals, Table 1) which enables the data collectors to pay special attention to these individuals. The technique for production of a high quality sputum sample can be clearly explained; they can motivate the suspects to provide sputum; and if a sample is not received suspects can be traced. The percentage of suspects providing sputum is therefore high in most surveys that use screening (86–98%, Table 1). The second advantage of screening is that the number of samples that has to be examined by the laboratory is limited facilitating high quality of the laboratory work.

Screening of the study population for suspects is most accurately performed by using both symptoms and abnormality on chest X-ray. With only chest X-ray to identify suspects, two studies reported an underestimation of the prevalence of smear-positive TB of 18% and 29%, respectively (Gothi *et al.* 1976; Datta *et al.* 2000). If only symptoms were used to identify suspects the prevalence would be underestimated by 37% (Gothi *et al.* 1976). A study that used chest X-ray and requested a sputum sample from all individuals in the study population showed that using chest X-ray to identify suspects missed 43% of the culture-positive cases (Gatner & Burkhardt 1980). In a study in South Africa, only 6 (50%) of 12 smear-positives had a suspicious chest X-ray (Fourie & Austoker 1981). In another study of 59 culture-positives, only 38 (64%) had a chest X-ray suspicious of TB (Fourie *et al.* 1980). Also in clinical practice, pulmonary TB is diagnosed in patients with a normal chest X-ray (Marciniuk *et al.* 1999) especially in HIV-infected individuals (Greenberg *et al.* 1994). In conclusion, all methods currently applied for screening of populations to identify suspects will result in an underestimation of the true prevalence, the underestimation will be less if screening is performed using chest X-ray and a combination of TB symptoms. Furthermore, there is no fixed 'correction factor' that can be used to estimate the true prevalence from the prevalence obtained by using a screening method.

Applying chest X-ray for screening involves high cost of mobile chest X-ray equipment, logistical difficulties

(accessibility of survey sites by trucks, power supply) and human capacity demands, in particular the need to have experienced radiologists who can read chest X-rays in the field teams.

Alternatively, a screening step is not applied and sputum samples are collected from all eligible individuals. FM microscopy can be used to examine the specimens as this allows for faster examination of slides. This strategy was recently tested in the national TB prevalence survey in Eritrea. The new strategy was feasible. However, it needs to be compared with the strategies that apply screening. The main risk of a strategy without screening may be the high number of people who need to provide samples which may result in less effort of the field teams to motivate individuals to produce a good quality sample so that the collected specimens are of a lower quality.

The number of sputum samples collected per individual will affect the prevalence estimate. Most TB control programmes have guidelines, which state that three sputum samples (spot-morning-spot) should be examined from each individual before a suspect can be regarded as smear-negative. To collect three sputum samples from individuals in a TB prevalence survey will give a substantial workload for the field teams and for the laboratory that performs the examinations. Furthermore, most smear-positives will be diagnosed by examination of the first slide (77–94%), examination of a second slide will add some more cases (12–15%) whereas examination of a third slide will only provide few additional cases (0.8–8%) (Harries *et al.* 1996,2000; Ipuge *et al.* 1996; Walker *et al.* 2000; Wu & Wang 2000; Van Deun *et al.* 2002; Gopi *et al.* 2004).

In the framework of TB prevalence surveys, it may be easier to collect spot specimens than morning specimens. However, it has been reported repeatedly that morning samples result in a higher positivity rate than spot samples (Andrews & Radhakrishna 1959; Urbanczik 1985). Thus, limiting sputum collection in the survey to spot samples may result in an underestimation of the prevalence of smear-positive TB.

Data collection

Visiting households for data collection seems to be a method ensuring high participation rates. Issues that need to be taken into account when using this procedure is the need for revisiting of households where no one is present during the initial visit and the fact that the interviewee may be easily distracted during the interview when it takes place in the household.

A TB prevalence survey is a good opportunity to collect information from a randomly selected group of individuals. Thus, adding more questions to the questionnaire may increase the amount of information obtained from the survey. However, the balance between quality and quantity of the information should be well guarded.

Collecting high-quality sputum samples is not easy in a clinical setting and will be even more difficult in a survey setting. Thus, good instruction is a necessity. Having a laboratory technician supervising sputum production may ensure high-quality specimens even though part of the collected samples may be saliva instead of sputum. Microscopic examination of saliva is considered to be less sensitive than examination of sputum for diagnosis of TB. It has been reported that in approximately 50% of the saliva samples of patients with a positive sputum sample bacilli could be demonstrated by the ZN method although the number of bacilli was considerably lower in saliva compared with sputum (Neild & Dunkley 1909; Yeager *et al.* 1967).

Sputum handling and examination

Results of microscopy examination are not affected by long delays between collection of sputum and examination (Paramasivan *et al.* 1983), even if sputum is kept at high temperatures for up to 1 month (Harries *et al.* 1998). This enables transportation of samples to a central laboratory for examination, which may facilitate high quality examination. Performing examination on the spot or close to the village allows for providing quick information to the patients and communities about the test results and if necessary to collect additional information or samples from the cases.

In a clinical setting, sensitivity of FM for diagnosis of culture-positive TB in TB suspects is higher than of ZN microscopy whereas specificity is comparable (Githui *et al.* 1993; Kivihya-Ndugga *et al.* 2003). FM examination uses 400 times magnification and ZN 1000 times. Therefore, the FM technique enables faster examination of slides, which is especially convenient in settings where a high number of sputum samples need to be examined such as TB prevalence surveys. TB prevalence surveys can still diagnose smear-positive TB using the same methods as the TB programme by restaining smears that are positive with FM using ZN staining and reexamination by light microscopy.

Recommendations

Case definition

The information obtained from the survey will be compared with the notification rates of the TB programme. Therefore, we recommend using the same definition for a TB case as is used by the TB programme.

Sample size

In the preparation phase of a survey it should be considered whether the resources available will enable inclusion of a sufficiently large sample size to obtain a useful estimate of the prevalence. Furthermore, the sample size should not be larger than necessary to provide a prevalence estimate with an accuracy needed for the purpose of the survey.

Sampling

Probability proportional to population size sampling is in general more efficient compared with simple random sampling (Nagelkerke *et al.* 2000). Simple random sampling will select more small districts. In the analysis, small districts weight less than large ones and if costs of including districts are equal then simple random sampling is less effective.

If there is large variation in prevalence between districts, this will contribute substantially to the uncertainty in the national prevalence. This uncertainty can be reduced by using stratification. This works best if the units within a stratum are rather homogeneous, i.e. have approximately the same value and if the mean strata values are quite different. Stratification may be also applied if separate estimates for the strata (e.g. regions or urban *vs.* rural) are required. The number of strata should be kept small and it is best to limit stratification to only one stage of sampling, in order to keep the design simple. For example, if there is likely to be a difference in TB situation by region as well as by rural–urban areas, the simplest possible stratification should be used (in this instance urban–rural). If stratification is used, data analysis should consider the stratification design.

Selection of households using lists of all households or groups of households is a valid method. If lists of households are not readily available or cannot be easily obtained another method for selection of households that can be applied is based on the Expanded Program on Immunization method (EPI) (Lemeshow & Robinson 1985). Using this method implies that the household where inclusion of individuals should start is determined by spinning a pen or bottle on the ground in the middle of the cluster and choosing the direction indicated by the nip of the pen or bottle. All eligible persons in the first house in this direction are included in the survey. The next household to be included can be the household whose front door is closest to the one just visited or a household in a specified direction of the first household. The EPI method is easy to apply in field situations. However, the probability of inclusion is not equal for all households. Those in the centre of the cluster have a higher probability of inclusion.

Study population

Defining which individuals should be eligible for inclusion in the survey should depend on the characteristics of the population concerning migration and the local concept of who is a member of a household. If membership of a household is a clearly defined concept and migration is limited, being a member of a household can be used to define eligibility. Else, other definitions such as 'having slept in the household the night before the survey team visit' may be used.

As smear positivity is infrequent in those <15 years of age and collection of sputum samples is difficult, exclusion of this group will reduce workload and have only a small effect on the prevalence estimate for the total population.

To increase feasibility of data collection the institutional populations may be excluded from the sampling frame. Furthermore, those living in inaccessible or restricted areas are often excluded from the sampling frame as it is considered inappropriate to expose the field teams to the dangerous conditions present in areas where there is a conflict or other incidents ongoing. This may bias the results of the survey and exclusion of areas should therefore be clearly stated in the report of the survey.

Case finding

The methods currently used for screening are not perfect. Use of symptoms to identify suspects will miss a considerable number of cases and use of chest X-ray is logistically difficult and expensive. Thus, there is a need for a new screening test. The main requirement of a screening test for TB prevalence surveys is high sensitivity as the great majority of the TB patients in the population should be detected in the screening step and a sufficiently high specificity as otherwise the workload for the laboratory will be too high. For instance, with tuberculin skin testing 78% of the population over 10 years of age was required to undergo sputum testing, resulting in a high workload for the laboratory making this screening method not useful in India (Chakraborty *et al.* 1995). Until a new screening tool has been identified and tested, the most appropriate screening method is a combination of symptoms and chest X-ray abnormalities.

Examination of a second specimen identifies a considerable additional number of smear-positive cases. Therefore, it is recommended to collect at least two specimens. Following the same argument it is recommended that at least one of the samples is a morning sample.

Data collection

To ensure high quality of the collected data well-trained and motivated field workers are essential. Furthermore, a quality control system using repeat interviews, double examination of chest X-rays, re-examination of slides and the like should be put in place.

Sputum handling and examination

FM is recommended for clinical settings were more than 100 slides per day are examined (Pio & Chaulet 1998). TB prevalence surveys normally cope with large numbers of slides thus it makes sense to use FM in TB prevalence surveys also because sensitivity is higher than that of ZN microscopy.

Conclusion

We showed that most recent surveys use comparable strategies but some issues still need to be evaluated. Especially strategies used for identification of cases need to be compared and alternatives designed and tested before a standardized study design can be proposed.

References

Andrews RH & Radhakrishna S (1959) A comparison of two methods of sputum collection in the diagnosis of pulmonary tuberculosis. *Tubercule* **40**, 155–162.

Attaran A (2005) An immeasurable crisis? A criticism of the millennium development goals and why they cannot be measured. *PLoS Medicine* **2**, e318.

Baussano I, Bugiani M, Gregori D *et al.* (2006) Undetected burden of tuberculosis in a low-prevalence area. *International Journal of Tuberculosis and Lung Disorders* **10**, 415–421.

Chakraborty AK, Suryanarayana HV, Murthy VV, Murthy MS & Shashidhara AN (1995) Prevalence of tuberculosis in a rural area by an alternative survey method without prior radiographic screening of the population. *Tuberculosis and Lung Disorders* **76**, 20–24.

China Tuberculosis Control Collaboration (2004) The effect of tuberculosis control in China. *Lancet* **364**, 417–422.

Datta M, Gopi PG, Appegowda BN, Bhima Rao KR & Gopalan BN (2000) Tuberculosis in north Arcot district of Tamil Nadu – a sample survey. *Indian Journal of Tuberculosis* **47**, 147–154.

Dye C, Scheele S, Dolin P, Pathania V & Raviglione MC (1999) Consensus statement. Global burden of tuberculosis: estimated incidence, prevalence, and mortality by country. WHO Global Surveillance and Monitoring Project. *Journal of American Medical Association* **282**, 677–686.

Fourie PB & Austoker LH (1981) Tuberculosis prevalence survey in the Daveyton (Benoni) urban black community. *South African Medical Journal* **60**, 64–67.

Fourie PB, Gatner EM, Glatthaar E & Kleeberg HH (1980) Follow-up tuberculosis prevalence survey of Transkei. *Tubercle* **61**, 71–79.

Gatner EM & Burkhardt KR (1980) Correlation of the results of X-ray and sputum culture in tuberculosis prevalence surveys. *Tubercle* **61**, 27–31.

Githui W, Kitui F, Juma ES, Obwana DO, Mwai J & Kwamanga D (1993) A comparative study on the reliability of the fluorescence microscopy and Ziehl-Neelsen method in the diagnosis of pulmonary tuberculosis. *East African Medical Journal* **70**, 263–6.

Gothi GD, Narayan R, Nair SS, Chakraborty AK & Srikantaramu N (1976) Estimation of prevalence of bacillary tuberculosis on the basis of chest X-ray and/or symptomatic screening. *Indian Journal of Medical Research* **64**, 1150–1159.

Gopi PG, Subramani R, Selvakumar N, Santha T, Eusuff SI & Narayanan PR (2004) Smear examination of two specimens for diagnosis of pulmonary tuberculosis in Tiruvallur District, south India. *International Journal of Tuberculosis and Lung Disorders* **8**, 824–828.

Gopi PG, Subramani R, Radhakrishna S *et al.* (2003) A baseline survey of the prevalence of tuberculosis in a community in south India at the commencement of a DOTS programme. *International Journal of Tuberculosis and Lung Disorders* **7**, 1154–1162.

Greenberg SD, Frager D, Suster B, Walker S, Stavropoulos C & Rothpearl A (1994) Active pulmonary tuberculosis in patients with AIDS: spectrum of radiographic findings (including a normal appearance). *Radiology* **193**, 115–119.

Guernier V, Guegan JF & Deparis X (2006) An evaluation of the actual incidence of tuberculosis in French Guiana using a capture-recapture model. *Microbes Infection* **8**, 721–727.

Hamid Salim MA, Declercq E, Van Deun A & Saki KA (2004) Gender differences in tuberculosis: a prevalence survey done in Bangladesh. *International Journal of Tuberculosis and Lung Disorders* **8**, 952–957.

Harries AD, Kamenya A, Subramanyam VR, Salaniponi FM & Nyangulu DS (1996) Sputum smears for diagnosis of smear-positive pulmonary tuberculosis. *Lancet* **347**, 834–835.

Harries AD, Nyirenda TE, Banerjee A, Mundy C & Salaniponi FM (1998) District sputum smear microscopy services in Malawi. *International Journal of Tuberculosis and Lung Disorders* **2**, 914–918.

Harries AD, Mphasa NB, Mundy C, Banerjee A, Kwanjana JH & Salaniponi FM (2000) Screening tuberculosis suspects using two sputum smears. *International Journal of Tuberculosis and Lung Disorders* **4**, 36–40.

Hong YP, Kim SJ, Lew WJ, Lee EK & Han YC (1998) The seventh nationwide tuberculosis prevalence survey in Korea 1995. *International Journal of Tuberculosis and Lung Disorders* **2**, 27–36.

Ipuge YA, Rieder HL & Enarson DA (1996) The yield of acid-fast bacilli from serial smears in routine microscopy laboratories in rural Tanzania. *Transactions of the Royal Society of Tropical Medical Hygiene* **90**, 258–261.

Kivihya-Ndugga LE, Van Cleeff MR, Githui WA *et al.* (2003) A comprehensive comparison of Ziehl-Neelsen and fluorescence microscopy for the diagnosis of tuberculosis in a resource-poor

urban setting. *International Journal of Tuberculosis and Lung Disorders* **7**, 1163–1171.

Lemeshow S & Robinson D (1985) Surveys to measure programme coverage and impact: a review of the methodology used by the expanded programme on immunization. *World Health Statistics Quarterly* **38**, 65–75.

Marciniuk DD, McNab BD, Martin WT & Hoeppner VH (1999) Detection of pulmonary tuberculosis in patients with a normal chest radiograph. *Chest* **115**, 445–52.

MoH Korea, Ministry of Health and Welfare, and Korean National Tuberculosis Association (1995) Report on the 7th Tuberculosis Prevalence Survey in Korea.

Nagelkerke NJ, Borgdorff MW, Kalisvaart NA & Broekmans JF (2000) The design of multi-stage tuberculin surveys: some suggestions for sampling. *International Journal of Tuberculosis and Lung Disorders* **4**, 314–320.

National Tuberculosis Control Program CAMBODIA (2005) *National TB prevalence survey 2002 Cambodia*.

Neild N & Dunkley EV (1909) The role of the saliva in the transmission of tubercle. *Lancet* **176**, 1096–1105.

Paramasivan CN, Narayana AS, Prabhakar R, Rajagopal MS, Somasundaram PR & Tripathy SP (1983) Effect of storage of sputum specimens at room temperature on smear and culture results. *Tubercle* **64**, 119–124.

Pio A & Chaulet P (1998) *Tuberculosis Handbook (WHO/TB/98.253)*. World Health Organization, Geneva, Switzerland.

Soemantri S & Senewe FP (2005) *Tuberculosis Prevalence Survey 2004*. Report jointly published by the National Institute of Health Research and Development (Indonesia), Directorate General of Communicable Disease Control and Environmental Health (Indonesia), and the World Health Organization.

Styblo K (1985) The relationship between the risk of tuberculous infection and the risk of developing infectious tuberculosis. *Bulletin of International Union of Tuberculosis* **60**, 117–119.

Tupasi TE, Radhakrishna S, Rivera AB *et al.* (1999) The 1997 Nationwide Tuberculosis Prevalence Survey in the Philippines. *International Journal of Tuberculosis and Lung Disorders* **3**, 471–477.

Urbanczik R (1985) Present position of microscopy and of culture in diagnostic mycobacteriology. *Zbl Bakt Hyg A* **260**, 81–87.

Van Deun A, Salim AH, Cooreman E *et al.* (2002) Optimal tuberculosis case detection by direct sputum smear microscopy: how much better is more? *International Journal of Tuberculosis and Lung Disorders* **6**, 222–230.

Walker D, McNerney R, Mwembo MK, Foster S, Tihon V & Godfrey-Faussett P (2000) An incremental cost-effectiveness analysis of the first, second and third sputum examination in the diagnosis of pulmonary tuberculosis. *International Journal of Tuberculosis and Lung Disorders* **4**, 246–251.

WHO (2003) *Treatment of tuberculosis: guidelines for national programmes*. Geneva, World Health Organisation (WHO/CDS/TB/2003.313).

WHO (2006) Global tuberculosis control: surveillance, planning, financing. WHO report 2006. Geneva, World Health Organization (WHO/HTM/TB/2006.362).

Wu ZL & Wang AQ (2000) Diagnostic yield of repeated smear microscopy examinations among patients suspected of pulmonary TB in Shandong province of China. *International Journal of Tuberculosis and Lung Disorders* **4**, 1086–1087.

Yeager H Jr, Lacy J, Smith LR & Lemaistre CA (1967) Quantitative studies of mycobacterial populations in sputum and saliva. *American Review of Respiratory Disorders* **95**, 998–1004.

Zaman K, Yunus M, Arifeen SE *et al.* (2006) Prevalence of sputum smear-positive tuberculosis in a rural area in Bangladesh. *Epidemiological Infection* **134**, 1052–1059.

Corresponding Author **Marieke J. van der Werf,** KNCV Tuberculosis Foundation, PO Box 146, 2501 CC The Hague, The Netherlands. Tel.: +31 70 416 7222; Fax: +31 70 358 4004, E-mail: vanderwerfm@kncvtbc.nl

Comment mesurer la prévalence de la tuberculose dans une population

L'Assemblée mondiale de santé a défini des objectifs pour l'évaluation de la performance des programmes de contrôle de la tuberculose, qui ont été inclus dans le cadre des objectifs du développement du millénium. Afin de pouvoir mesurer le progrès vers ces objectifs, les programmes de contrôle de la TB ont besoin d'informations épidémiologiques. Á l'heure actuelle les données de surveillance des pays avec une charge élevée de TB sont insuffisantes pour évaluer la performance des programmes de contrôle de la TB soit parce qu'elles sont incomplètes ou de mauvaise qualité. Plusieurs pays à charge élevée de TB ont conduit des surveillances de prévalence de TB afin d'obtenir l'information épidémiologique nécessaire. Comme une méthode standard pour les surveillances de la prévalence n'a pas été définie, nous discutons ici des différentes options pour la mesure de la prévalence avec leurs avantages et inconvénients. La plupart des surveillances utilisent des stratégies comparables. Des stratégies alternatives à coût plus réduit devraient être évaluées.

mots clés tuberculose, charge de la maladie, surveillance de prévalence, objectifs de développement du millénium, méthodologie

Punto de vista Cómo medir la prevalencia de tuberculosis en una población

La Asamblea Mundial de la Salud ha definido metas para medir el desempeño de los programas de control que han sido incluidos dentro de la estructura de los Objetivos de Desarrollo del Milenio. Para ser capaces de medir el progreso hacia estas metas, los programas de control de la TB necesitan de información epidemiológica. En este momento, los datos de vigilancia epidemiológica de países con una carga alta de TB, son insuficientes para evaluar el desempeño de un programa de control de la TB por tener ser incompletos o de mala calidad. Algunos países con una carga de enfermedad alta han realizado estudios de prevalencia de TB, con el fin de obtener información epidemiológica. Puesto que no se ha definido un método estandarizado para estudios de prevalencia de TB, discutimos diferentes opciones para medir la prevalencia, sus ventajas y desventajas. La mayoría de los estudios utilizan estrategias comparables. Es necesario evaluar estrategias alternativas de menor coste.

palabras clave Tuberculosis, carga de enfermedad, estudio de prevalencia, Objetivos de Desarrollo del Milenio, metodología