Survey and surveillance development in settings with low human immunodeficiency virus prevalence

D.J. Nanan,[1] M.M. Kadir[1] and F.M.M. White[1]

تطوير المسوحات والترصُّد في مناطق الانتشار المنخفض لفيروس العَوَز المناعي البشري

ديبرا ج.نانان ومحمد مسعود قادر وفرنكلين م.م. هوايت

خلاصة: في معظم البلدان، توجد نزعة في المراحل المبكرة من أوبئة فيروس العَوَز المناعي البشري، إلى المبادرة بإجراء مسوحات متفرقة حول الجماعات التي يُعتقد أنها معرَّضة لخطر مرتفع، قبل أن تُنشأ أنظمة رسمية للترصد. ولكن من سوء الحظ أن العينات الصغيرة الحجم في المناطق ذات الانتشار المنخفض، تؤدي إلى تقديرات لا يعتمد عليها لمعدلات الانتشار واتجاهاتها، مع ميل لا يمكن اجتنابه نحو النتائج الإيجابية. وفي هذه الدراسة نعرض حسابات لأحجام العينات، لعينات نموذجية استُعملت في مسوحات فعلية. وكان نموذجنا المرجعي هو باكستان. إن بالإمكان استنباط بيانات أكثر فائدة عن السلوكيات المحفوفة بالخطر وعن احتمالات الانتشار، من دراسة الأمراض المنقولة جنسياً، الأكثر شيوعاً، وما يصاحبها من سلوكيات خطيرة، بما في ذلك إجراء تقديرات للمعارف والاتجاهات والمعتقدات والممارسات.

ABSTRACT In most countries. during the early phases of a human immunodeficiency virus epidemic, independently initiated surveys of perceived high-risk groups tend to precede the development of formal surveillance systems. Unfortunately, in low-prevalence settings, small sample sizes produce unreliable estimates of prevalence and trends, with an inevitable tendency towards positive results. In our study, we present sample size calculations and typical samples used in actual surveys, with Pakistan as our example. More useful data on risk behaviour and potential for spread can be derived from the study of commoner sexually transmitted diseases and associated risk behaviours, including assessments of knowledge, attitudes, beliefs and practices.

Mise en place d'enquêtes et de la surveillance dans les environnements à faible prévalence du virus de l'immunodéficience humaine

RESUME Dans la plupart des pays, pendant les phases préliminaires de l'épidémie de virus de l'immunodéficience humaine, des enquêtes mises en œuvre indépendamment sur les groupes perçus à haut risque ont tendance à précéder la mise en place des systèmes de surveillance formelle. Malheureusement, dans les environnements à faible prévalence, des échantillons de petite taille produisent des estimations de la prévalence et des tendances peu fiables. avec une orientation inévitable vers des résultats positifs. Dans notre étude, nous présentons des calculs de la taille des échantillons et des échantillons typiques utilisés dans des enquêtes réelles, prenant le Pakistan comme exemple. Des données plus utiles sur le comportement à risque et le potentiel de propagation peuvent être tirées de l'étude des maladies sexuellement transmissibles les plus courantes et des comportements à risque associés, y compris les évaluations des connaissances, attitudes, croyances et pratiques.

[1]*Department of Community Health Sciences, The Aga Khan University, Karachi, Pakistan.*
Received: 10/02/00; accepted: 02/05/00

Introduction

The pandemic of human immunodeficiency virus (HIV) infection was first recognized in North America in the early 1980s, although the disease is now known to have been present in human hosts at least two decades earlier. Since then, both HIV and acquired immunodeficiency syndrome (AIDS) have been reported in virtually every country. In 1997, global AIDS deaths totalled 2.3 million, with 11.7 million cumulative deaths since the pandemic began. More than 40 million people will be living with HIV/AIDS in the year 2000 [*1*].

Eastern Mediterranean Region (EMR) countries are still considered low incidence settings. It is imperative in such settings that appropriate prevention and control measures are established based on sound surveillance data while the infection is still confined to high-risk groups, in the event that it may spread more generally throughout the population. Our focus was on methodological issues regarding surveys of specific population groups at different levels of perceived risk in low-incidence countries, using Pakistan as the example.

HIV/AIDS surveillance in Pakistan

The first reported death from AIDS in Pakistan was an African sailor in 1986 [*2*]. The following year, the first HIV-infected person was detected in Pakistan, and in 1988 the first indigenous transmissions of HIV infections were reported in two independent studies [*3,4*]. In 1992, the World Health Organization (WHO) estimated a 0.2% prevalence of HIV in a study of prisoners in Pakistan; in 1993, 3 cases of HIV infection were identified among a survey of 183 prisoners [*5*]. Taking into account data from existing serosurveys, recent estimates show there are approximately 80 000–100 000 cases nationally, or a prevalence rate of 64–83 per 100 000 population [*6,7*]. The estimates were extrapolated from a number of relatively small surveys in mostly high-risk settings. However, in 1998, the National AIDS Control Programme (NACP) reported only 1364 cumulative cases of HIV infection [*8*]. The majority of these cases were men, of whom 170 had developed AIDS. Sexual contact followed by the use of blood and blood products are cited as the most common modes of transmission. However, mode of transmission has not been recorded in 41.0% of total cases.

Assuming that the extrapolations noted are a reasonable basis for assessing the adequacy of reporting, the discrepancy between reported versus projected cases and the relative dearth of information on mode of transmission suggest the following important deficiencies in HIV surveillance in Pakistan, which may also apply to other low-incidence countries.

- The small number of reported cases implies that most infections (98.0%) are thus far unreported, and are likely to be undetected and hence unprotected in terms of transmission.
- Major deficiencies in data exist in the case reports that are received.

The potential for the incidence of HIV infection increasing in countries such as Pakistan, which is currently regarded as low incidence, must be considered seriously as there is evidence of underlying sexual dynamics that increase the risk of transmission. For example, although commercial sex is illegal in Pakistan, red-light areas can be found in major cities and are frequented mostly by traders, truck drivers, seafarers and travellers. Potential for spread also exists through unsatisfactory and unhygienic

practices in health care and community settings through excessive and nonsterile use of syringes. Intravenous drug abuse also exists in the country, although few data are available. Other conditions conducive to transmission include: an unprotected blood supply, low educational levels, high fertility and low use of barrier contraception [*9*]. In these respects, Pakistan may be similar to other Asian countries such as Thailand and India that moved from a low-incidence phase with infection confined to high-risk groups, to a higher incidence when the infection spread to the wider population [*10*].

Thus far, a small number of independent studies have been conducted on selected groups in Pakistan: migrant workers [*7*]; male international travellers [*11.12*]; blood donors [*12,13*]; prison inmates [*14,15,18*]; long-distance truck drivers [*16*]; drug abusers [*17.18*]; female sex workers [*18*]; transvestite sex workers [*19*] and university students [*20*]. These studies provide a useful starting point for estimating the occurrence of HIV infection in specified risk groups but are more opportunistic in nature, contributing mainly to ad hoc surveillance that is typical of HIV surveillance in most countries during the early phase of an epidemic. A more scientific approach would be the development of a national surveillance framework, which includes the monitoring of selected groups using standard protocols.

Methodological limitations of ad hoc surveillance

The use of opportunistic surveys as a basis for surveillance is seriously flawed. Comparisons of such studies in terms of person, place and time are difficult as the designs are not standardized. The following examples are some of the biases inherent in such studies

Sampling bias

- Convenience sampling: by definition non-random and non-representative;
- Membership bias: membership in a group may imply differences which affect reported frequency of behaviour (e.g. some may be over-reported, others under-reported);
- Sample size bias:
  - sample size may be inadequate to detect a rare infection
  - sample size may be inadequate to detect a trend.

Measurement bias

- Diagnostic suspicion: knowledge of a person's prior exposure may influence the intensity or outcome of a diagnostic process;
- Exposure suspicion: knowledge of a person's disease status may affect the intensity and outcome of the search for plausible or acceptable exposure;
- Recall bias: questions about specific exposures may be asked many times of cases but only once of controls.

Other

- Positive results bias: studies which show evidence of a problem (e.g. HIV-positive individuals) are more likely to be submitted and accepted for publication than those which do not, thereby potentially distorting the apparent frequency of the problem.

A critical consideration in any study design is sample size. Assuming that a sample is random, if it is desired to increase the precision of an estimate, then the sample size must be increased. However, to obtain reasonable confidence intervals (CI) when estimating the prevalence of relatively rare conditions, sample size requirements can

become extreme. For example, a sample size of 250 is needed to provide 95% CI of 1.0%–5.0%. If the observed prevalence derived from the sample is 2.0%, then a larger sample size of 1000 is required to narrow the 95.0% CI to 1.0%–3.0% (Table 1) [*21*].

In practice, much smaller samples tend to be used. In Pakistan, the majority of reported HIV surveys have used smaller sample sizes (Table 2), and it follows that their estimates of HIV prevalence must have wide CI (i.e. the resulting estimates are less precise). For example, one study of 183 prisoners revealed 3 infected with HIV, an estimated prevalence of 1.6%, and a CI of 0%–10.0%. Conversely, a study among intravenous drug users (sample size = 77) revealed no HIV infection, yet the CI were 0%–6.0%. Table 3 illustrates sample size requirements needed to detect a statistically significant change ($P < 0.05$) in prevalence between two points in time [*20*]. Based on Table 2, neither of the two studies illustrated, if repeated after a specified interval, would be adequate in size to reveal a statistically significant difference in prevalence with desirable CI.

Therefore, far larger sample sizes than we have seen to date for HIV studies in Pakistan are required to detect differences over time. Table 4 is based on two reported studies of HIV from Pakistan. It shows the inadequate power that one of these studies (on 302 seafarers) would have if it were repeated in an attempt to determine an increase in prevalence of 20.0% from 2.0% to 2.4% over a given time period. The second reported prevalence is from the NACP and is based on screening carried out over several years, hence a much larger sample (1 350 000). Although this appears close to the sample size sufficient to measure a

Table 1 **95% confidence intervals for observed prevalence (0% and 2%) by sample size**

Prevalence (%)	Sample size 50	100	250	500	1000
0	0–7	0–4	0–2	0–1	0
2	0–11	0–7	1–5	1–4	1–3

Table 2 **Typical sample sizes and HIV findings from selected studies, Pakistan**

Study group	No. tested	No. HIV-positive
Truck drivers	77	0
Commercial sex workers	81	0
Intravenous drug users	77	0
Seafarers	302	2
Prisoners	183	3
STD patients	215	0
National data	1 350 000	869

STD = sexually transmitted disease

Table 3 **Sample sizes required to detect a significant increase between two proportions, with power of 0.80, by baseline prevalence and proportional increase**

Baseline prevalence (%)	Proportional increase compared with baseline prevalence 10%	20%	30%	40%	50%	60%	70%	80%	90%
1	145 000	34 000	14 000	7 290	4 280	3 000	2 070	1 459	1 060
5	28 000	6 550	2 800	1 500	903	585	400	282	204

Table 4 **Sample size required to detect a 20% prevalence change at 95% confidence intervals and varying power**

Power (%)	Sample size required to detect a 20% prevalence change from: 2.0% to 2.4%[a]	0.06% to 0.07%[b]
80	21 605	1 039 594
70	17 095	821 725
50	10 826	518 872
20	3 848	181 935
10	1 702	78 520
7	1 073	48 389
3	324	13 290

[a]*Based on a study of 302 seafarers*
[b]*Based on screening by the National AIDS Control Programmes of 1 350 000 people over several years*

change, the figure is cumulative and renders this method less than optimal.

These observations help us recognize the basic problem of studying relatively rare conditions such as HIV in low-prevalence countries. The alternative is to shift the emphasis from HIV to measuring behavioural risk factors and more common STDs. More common STDs would allow more modest sample sizes to be used, treating these in turn as "proxies" for the risk of HIV transmission, should it be introduced into a particular setting.

Characteristics of an ideal surveillance model

Disease surveillance is defined by the International Epidemiological Association (IEA) as "the continuing scrutiny of all aspects of occurrence and spread of a disease that are effective to pertinent control" [*22*]. Ideally, surveillance includes systematic collection and evaluation of: morbidity and mortality reports, special reports of field investigations, identification of infectious agents by laboratories, information on immunity levels in populations and other relevant epidemiological data. Existing surveillance systems can be assessed, or a proposed system conceptualized, by applying the following eight attributes: a) sensitivity, b) specificity, c) predictive value positive, d) representativeness, e) timeliness, f) simplicity, g) flexibility and h) acceptability [*23,24*]. Attributes may vary in importance depending on the situation and may involve trade-offs. Emphasis should be on addressing the most relevant attributes, while accepting that others may not be fully achieved. The measure of whether a surveillance system has achieved an optimal balance of attributes lies in its usefulness to disease prevention and control.

Surveillance based on morbidity and mortality case reporting can be described as active or passive. Active surveillance involves periodic stimulation (telephone calls, written notices, periodic visits) of reporting units. Passive surveillance relies on reporting to take place without direct intervention, and in most countries is incomplete and unreliable. When a condition is rare, the situation is compounded by a lack of diagnostic awareness and recognition, aside from the unreliability of reporting if and when the condition is revealed.

Sentinel surveillance

In situations where the prevalence of the condition under study is low, or where surveillance is not well established, "sentinel surveillance" provides a viable alternative. Sentinel surveillance involves actively monitoring specifically selected population groups, defined on the basis of sociobehavioural, demographic, geographic or other

relevant variables. Such groups are not in themselves representative of the entire "population", but should be sufficiently stable and reliable to allow for periodic assessment and trend analyses. The sentinel approach is not the same as ad hoc surveillance based on opportunistic studies of conveniently selected groups. Sentinel sites should be identified within an overall framework of selected settings. Sites and their groups should be amenable to repeated study using standardized methodology, including attention to issues such as adequacy of sample size to detect trends. Sites can be periodically validated to determine the performance characteristics of the system. The selected sentinel sites may be used as a proxy for populations with similar characteristics; a reasonable number and variation of sites would therefore give a more accurate picture of the situation at the national or provincial level.

This type of surveillance is appropriate for the study of HIV infection in low-incidence countries. Repeat studies using standard protocols can be conducted in selected, operationally stable sentinel sites. Provided that characteristics of sentinel sites do not change over time, trends observed at sentinel sites may be extrapolated within the overall context [*25*]. The key is to develop standardized methodologies to minimize the effect of bias and to optimize the validity of interpretation. Given the sensitive nature of human sexuality, adequate information about related infections cannot be obtained by a single study design. A combination of qualitative and quantitative methods and innovative approaches is required to develop a broad but well focused knowledge in community settings, in order to adequately define the dynamics of STDs in general, and HIV transmission in particular. It is beyond the scope of this paper to list the options in any detail, but examples include case–control studies to identify risk factors, and focus groups and in-depth interviews to identify underlying behavioural issues, including attitudes towards health services and care-seeking behaviour.

Sexually transmitted diseases and associated behaviours as a proxy for HIV risk

In countries where HIV prevalence is low, direct surveillance is impractical because of sample size and other cost considerations, especially given the high likelihood of clustering in particular settings. However, since HIV is primarily transmitted by sexual contact, useful data on potential for spread can be inferred from studying more common STDs and the associated knowledge, attitudes, beliefs and practices (KABP) which increase an individual's risk of acquiring an STD. In addition, there is strong evidence that transmission of HIV is facilitated in the presence of other STDs, such as genital ulcers [*26*]. Thus, the integration of HIV and other STDs in an overall sentinel surveillance framework becomes advantageous when viewed in the context of prevention and control of diseases that share common behavioural risk factors.

The following list illustrates four rationales for an integrated sentinel surveillance system as a basis for prevention and control of STDs:

- to provide a situation analysis on variables relevant to transmission, including behavioural factors;
- to identify trends in disease and risk factors over defined time periods;
- to provide essential data as a basis for the design of appropriate interventions;
- to serve as a baseline against which the impact of interventions can be assessed.

Such distinct rationales produce different requirements in terms of study design,

sample size and resources needed to execute surveys that will combine to provide a basis for HIV/STD surveillance.

Conclusion

Based on the foregoing considerations, we advocate a formal sentinel surveillance system to capture three levels of disease risk (low, medium and high) in different population subgroups. The overall design consists of a framework within which various types of studies and their products interrelate to produce an assessment of behavioural risk factors and the occurrence of STDs relevant to defined geographic areas. Examples of subgroups that could form part of a sentinel surveillance network include: antenatal clients (low risk), migrant workers (medium risk) and STD clinic attendees (high risk). Sample size should be sufficient to allow calculation of estimates of prevalence and underlying risk in relation to person, place and time. Further, the system will ideally measure trends in underlying behavioural risk factors, or shifts in the prevalence of more common STDs through repeat surveys at sentinel sites. The sentinel surveillance system can be evaluated periodically on the basis of the eight attributes listed earlier. Using data from such a system, demographic and risk-factor projection models can be applied in order to predict the future disease burden in both health and economic terms, which has implications for public policies such as health and social sector resource allocations.

Finally, STDs other than HIV are important in their own right so that the development of integrated sentinel surveillance will also address this public health need.

Acknowledgement

An earlier version of this paper was presented at the 6th Congress of the Asia Pacific Association of Societies of Pathologists, 25–28 November, 1999, Karachi, Pakistan.

References

1. *The World Health Report 1998. Life in the 21st century. A vision for all.* Geneva, World Health Organization, 1998.
2. Mujeeb SA, Kayani N, Khursheed. HIV/AIDS in Pakistan. *CARC calling*, 1994, 7(2):31.
3. Khanani MR et al. HIV associated disorders in Pakistan. *AIDS research and human retroviruses*, 1988, 4(2):149–54.
4. Abdul Mujeeb S, Hashmi MRA. A study of HIV antibody in sera of blood donors and people at risk. *Journal of the Pakistan Medical Association*, 1988, 38(8):221–2.
5. Mujeeb SA, Hafeez A. Prevalence and pattern of HIV infection in Karachi. *Journal of the Pakistan Medical Association*, 1996, 46:255–8.
6. Hyder AA, Khan OA. HIV/AIDS in Pakistan: the context and magnitude of an emerging threat. *Journal of epidemiology and community health*, 1998, 52(9):579–85.
7. Khawaja ZA et al. HIV/AIDS and its risk factors in Pakistan. *AIDS*, 1997, 11(7):843–8.
8. Hakim A, Cleland J, Batti MH. *Pakistan Fertility and Family Planning Survey*

1996–97. Preliminary report. Islamabad, National Institute of Population Studies, and London, Centre for Population Studies, London School of Hygiene and Tropical Medicine, 1998.

9. *UNAIDS country profile, Pakistan.* Geneva, UNAIDS, 1999.

10. Cleland J, Ferry B. *Sexual behaviour and AIDS in the developing world.* Geneva, World Health Organization, 1995.

11. Malik A, Tariq WUZ. Dilemma of spread of AIDS. *Pakistan journal of pathology*, 1993, 4:79–80.

12. Kayani N et al. A view of HIV-1 infection in Karachi. *Journal of the Pakistan Medical Association*, 1994, 44(1):8–11.

13. Kakepoto GN et al. Epidemiology of blood-borne viruses: a study of healthy blood donors in southern Pakistan. *Southeast Asian journal of tropical medicine and public health*, 1996, 27(4):703–6.

14. Baqi S et al. *HIV-1 antibody seroprevalence and associated risk factors among prison inmates in Sindh, Pakistan.* Paper presented at the 1st National Symposium, The Aga Khan University, Karachi, Pakistan, 4–14 September, 1994.

15. Khan AJ et al. *Prison inmates in Sindh their knowledge and attitudes towards HIV-related risk behaviour.* Paper presented at the 2nd National Symposium, The Aga Khan University, Karachi, Pakistan, September, 1995.

16. Khan AJ et al. *Prevalence of sexually transmitted diseases among long-distance truck drivers in Pakistan.* Paper presented at the 2nd National Symposium, The Aga Khan University, Karachi, Pakistan, September, 1995.

17. Baqi S. *HIV seroprevalence and risk factors in drug abusers in Karachi.* Paper presented at the 2nd National Symposium, The Aga Khan University Karachi, Pakistan, September, 1995.

18. Baqi SA et al. HIV antibody seroprevalence and associated risk factors in sex workers, drug users and prisoners in Sindh, Pakistan. *Journal of acquired immune deficiency syndromes and human retrovirology*, 1998, 18(1):73–9.

19. Khan OA, Hyder AA. HIV/AIDS among men who have sex with men in Pakistan. *Sexual health exchange*, 1998, 2.

20. Raza MI et al. Knowledge, attitude and behaviour towards AIDS among educated youth in Lahore, Pakistan. *Journal of the Pakistan Medical Association*, 1998, 48(6):179–82.

21. *Guidelines for sexually transmitted infections surveillance.* Geneva, World Health Organization, 1999.

22. Last JM, ed. *A dictionary of epidemiology*, 3rd ed. New York, Oxford University Press, 1995.

23. Thacker SB, Parrish RG, Trowbridge FL. A method to evaluate systems of epidemiologic surveillance. *World health statistics quarterly*, 1988, 41:11–8.

24. Klaucke DN et al. Guidelines for evaluating surveillance systems. *Morbidity and mortality weekly report*, 1988, 37(suppl. S5):1–20.

25. Betts CD, Zacarias F. Surveillance for HIV, AIDS and STDs. In: Lamptey P et al., eds. *The handbook for AIDS prevention in the Caribbean.* Research Triangle Park, North Carolina, Family Health International, 1992.

26. *Reproductive tract infections. A set of fact sheets.* Thailand, The Population Council Incorporated, 1999.