Tropical Medicine and International Health

VOLUME 3 NO 9 PP 747–750 SEPTEMBER 1998

Pooling sera to reduce the cost of HIV surveillance: a feasibility study in a rural Kenyan district

T. Verstraeten[1], B. Farah[2], L. Duchateau[3] and R. Matu[4]

1 *Kenya National AIDS Control Programme, Nairobi, Kenya*
2 *DMLT, University of Nairobi, Kenya*
3 *European Organization for Research and Treatment of Cancer, Brussels, Belgium*
4 *University of Nairobi, Kenya*

Summary Seroprevalence studies are crucial in HIV control programs but too expensive at district level. We evaluated the applicability of pooling sera and how it can reduce cost and affect accuracy at district level. 740 samples collected from antenatal clinic attendants for a sentinel survey in a rural Kenyan district were screened individually and in pools of 10. The seroprevalence when measured individually was 7.30%, while the calculated seroprevalence from pooled testing was 7.49%. Pooling was practicable and reduced costs by 62% for a marginal loss of accuracy. It is a useful tool in increasing the affordability of surveillance at district level. A pool size of 8 would have resulted in optimal cost reduction at minimal loss of accuracy.

keywords HIV, seroprevalence, pooling, cost reduction, accuracy, applicability

correspondence Dr T. Verstraeten, Hellegatstraat 21, 9070 Destelbergen, Belgium

Introduction

Knowing the seroprevalence and incidence of a disease can be of crucial importance for planning and evaluating a disease control programme. Surveillance of HIV seroprevalence within a population is an essential part of HIV control programs, but can be too expensive and time-consuming at district level. Pooling sera is one method which has been shown to reduce the cost of seroprevalence surveys (Cahoon-Young *et al*. 1989; Kline *et al*. 1989; Behets *et al*. 1990; Babu *et al*. 1993). Pooling has been used since the end of World War II when large numbers of returning American soldiers were screened for syphilis. Recently, its application has also been discussed in relation to HIV seroprevalence studies. Most authors agree that pooling is feasible for various HIV test kits and can save costs (Babu *et al*. 1993; Cahoon-Young *et al*. 1989; Kline *et al*. 1989; Behets *et al*. 1990; Perriens *et al*. 1993). Pool sizes used in these trials seem to have been chosen empirically and the possible use of pooling at district level was not tested. In Kajiado, a rural district south-east of Nairobi, we have been running a HIV sentinel surveillance system on specimens collected from ANC attendants for unanimous unlinked HIV screening since 1995. To reduce the cost and increase sustainability, we decided in 1996 to evaluate testing of pooled specimens in parallel with individual testing. We measured the accuracy of pooled testing and evaluated its applicability and cost-saving in our district setting. Finally we calculated how different pool sizes would have affected our findings and thus identified the 'ideal' pool size.

Materials and methods

Samples were collected from 740 pregnant women at three sites over a period of three months. Samples from 2 sites (hospitals) were frozen at − 20 °C and transported once a month to the central laboratory. Samples from the third site (a health centre) were stored at 4 °C for a maximum of 7 days before transport and freezing at the central laboratory. Seventy-four pools were constituted by mixing 150 μl from 10 consecutive samples. Both individual and pooled sera were first tested by EIA (Innotest HIV-1/HIV-2, Innogenetics, Belgium). Positives were confirmed by a rapid assay test (Capillus HIV-1/HIV-2, Cambridge Biotech, Ireland). We used these two test kits as they are issued by the National AIDS Control programme to all government hospitals.

Samples with differing results (EIA-positive and rapid assay-negative) were tested by a second EIA (Viranostika, Organon, Holland) and classified positive or negative according to the result of the second EIA. A sample was

therefore judged to be positive if both the first EIA and the Capillus were positive or if both EIAs tested positive. Quality was controlled by the WHO reference laboratory at the University of Nairobi through examination of every fifth negative and all positive samples.

Data analysis

The prevalence found in the pools ($P_{pool} = S/N$, whereby S = number of positive pools, N = total number of pools) was converted to an estimation (P_{est}) of the true prevalence (π) using the following equation (Kline *et al*. 1989):

$$P_{est} = 1 - (1 - (P_{pool})^{1/A} \quad (1)$$

where A = pool size.

An approximate standard error of this estimate of π (Hauck 1991) is given by

$$\text{var}(P_{est}) = \frac{(1 - P_{est})^2\left[(1 - P_{est})^{-A} - 1\right]}{A^2N} \quad (2)$$

while the approximate 95% confidence interval can be obtained using the normal approximation given by

$$P_{est} \pm 1.96 \times \sqrt{(\text{Var}(P_{est})} \quad (3)$$

The total cost C of a seroprevalence survey is estimated for A = 1 by

$$C = N \times s + N \times e_1 + N \times (P_{est} + P_{(+,-,-)}) \times c + N \times (P_{(+,-,-)} + P_{(+,-,+)}) \times e_2 \quad (4)$$

and for A > 1 by

$$C = NA \times s + N \times e_1 + N \times (P_{pool} + P_{(+,-,-)}) \times c + N \times (P_{(+,-,-)} + P_{(+,-,+)}) \times e_2 + N \times g \quad (5)$$

whereby

s = cost for collecting one sample,

g = cost for composing one pool,

e_1 = cost per first EIA,

c = cost per Capillus,

e_2 = cost per second EIA,

$P_{(+,-,-)}$ = proportion of the pools that test positive at the first EIA but negative at the Capillus and the second EIA,

$P_{(+,-,+)}$ = proportion of the pools that test positive at the first and second EIA but negative at the Capillus.

Results

Of the 740 individual results, 64 were positive on the first EIA and 11 of these were negative on rapid assay. Two of those were confirmed positive on the second EIA, whereas the other 9 were false positive at the first EIA (FP). The seroprevalence for individual testing was therefore 7.3%. Forty-one of the pools tested gave a positive reading on the first EIA, one of which gave a negative rapid assay and a second negative EIA result (FP). The seroprevalence was thus estimated by the pooled test as 7.49%.

The distribution of positive samples among the pools was as follows:

Number of positive samples	1	2	3	0
Number of pools	29	9	2	34

All pools containing positive samples tested positive, one pool containing only negative samples tested positive on the first EIA but negative on confirmation. Two of the false positive samples went into pools containing positive samples, while the remaining seven went into pools containing only negative samples. Only one of these pools tested falsely positive on the first EIA.With no loss of sensitivity and a minor decrease of specificity (from 98.7 to 97.1%), the positive predictive value (PPV) for pooled samples (97.6%) far exceeded that of individual samples (85.7%).

The amount of time spent on pooling was 0.2 h per pool, while 0.1 h was needed per EIA and Capillus test. For a commercial price of USD 0.1 per needle, USD 0.1 per syringe, USD 0.01 per pipette tip, USD 0.5 per vial and a wage of USD 1/h for a lab technician, the estimated cost for collecting one sample (cost of 1 syringe + 1 needle + 1 vial + 0.1 h wage lab technician) was USD 0.8. The cost of pooling was equally USD 0.8 per reconstructed pool (cost of 10 pipette tips + 1 vial + 0.2 h wage lab technician). For a commercial price of USD 2 per EIA and USD 4 per Capillus, the estimated cost per analysed sample (cost of 1-test + 0.1 h wage lab technician) was USD 2.1 for an EIA and USD 4.1 for a Capillus. The total cost of the study based on pooled samples amounted to USD 905, whereas the study based on individual samples cost USD 2381. The cost was thus reduced by 62%. The 95% confidence intervals for the prevalence were 0.073 $\pm$0.01874 for the study based on individual samples and 0.0748 $\pm$0.02286 for the study based on pooled samples. Thus there is a substantial cost reduction for a moderate drop in accuracy.

Practical problems were expected mostly during the pooling process, where a sample could easily be included twice or excluded from a pool. By introducing a wooden pooling board designed for fixing 10 consecutive vials and a empty vial to contain one pool, together with specific forms graphically illustrating the samples to be included in each pool, we tried to reduce the risk for such errors. As no discordance was found between positive samples and positive pools, we assumed no mistakes were made in the pooling process and found it practicable.

Discussion

We considered three variables in analysing our data: accuracy, cost and applicability. Accuracy, expressed as the length of the 95% CI of the prevalence estimate, is slightly less for a pool size of 10 compared to individual testing, but the cost reduction is substantial. We chose a pool size commonly used in other experiments (Cahoon-Young *et al*. 1989; Behets *et al*. 1990; Sherlock *et al*. 1995). We calculated how different pool sizes would have affected our findings and tried to identify the ideal pool size for our setting by considering three scenarios: First, the cost can be fixed and the pool size chosen which minimizes the length of the confidence interval. Alternatively, the accuracy can be fixed and the pool size chosen as to minimize the cost. Finally, for a fixed sample size, as was the case for our survey, variations of both price and accuracy have to be evaluated to choose an appropriate combination. We calculated the cost and accuracy for pool sizes ranging from 1 to 15. Results are summarized in Table 1 and illustrated in Figure 1. In general, both cost and accuracy decrease with increasing pool size. The gain in cost reduction becomes negligible at pool sizes above 8, whereas the accuracy still decreases substantially. Therefore a pool size of 8 can be considered as 'ideal' for our expected prevalence (7.3%), wage (USD 1/h) and test prices (USD 2 and 4).

As for feasibility, the larger the pool size is, the greater the chance for technical errors to occur in constructing the pools. We found a pool size of 10 an easy number to work with for laboratory procedures. We assume that a pool size below 10 would equally not have resulted in pooling mistakes. Pool sizes above 10, however, as suggested by Behets *et al*. (1990) or as used by Kline *et al*. (1989), might lead to more frequent mistakes. More study is needed to determine the probability of pooling and testing errors and their relative importance.

Table 1 Estimated total cost and accuracy for a seroprevalence of 7.3%

Pool size	Cost	Accuracy
1	2449	0.0375
2	1919	0.0382
3	1540	0.0390
4	1347	0.0397
5	1229	0.0405
6	1148	0.0414
7	1089	0.0422
8	1043	0.0431
9	1006	0.0440
10	976	0.0450
11	951	0.0459
12	929	0.0470
13	910	0.0480
14	893	0.0491
15	878	0.0502

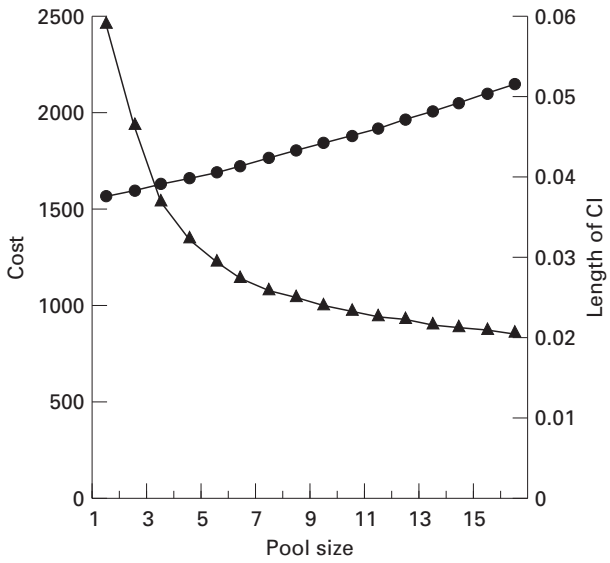

Figure 1 Estimated total cost and accuracy for a seroprevalence of 7.3%.

The increase in PPV is due to the artificial increase of prevalence with the specificity nearly unchanged. We noticed that OD values of the pools were all either strongly positive or negative and not near the cutoff value. Note that we used undiluted positive and negative controls for calculating the cut-off value, since diluting the positive control (1:10) resulted in an OD value rejected by the test protocol.

The various potential benefits of pooled testing have repeatedly been described by authors from both medical and statistical backgrounds. Still there is a reluctance among many epidemiologists to implement this technique. This reluctance probably finds its origin in the assumption that technical errors are likely to have an exponential negative impact on accuracy. Pooling, however, was shown by Tu *et al*. (1994) to be more accurate than individual testing by reducing the number of tests (and thus the possible errors) and increasing the prevalence (and thus the PPV). Another misconception is the belief that pooling is only useful in low-prevalence settings. In our calculations we also looked at the potential benefits for different prevalences and found a cost reduction of 50% for prevalences up to 15% without significant loss of accuracy.

Note

We have developed a model that summarizes our calculations for ideal pool sizes for varying scenarios (fixed price,

accuracy or sample size), prevalences, and costs. This model can be downloaded as a zipped Excel file from the Website of the Limburg University Centre in Belgium on: http://www.luc.ac.be/research/groups/statistics/members/luc.HTLM.

Acknowledgements

We thank the Belgian Administration for Development Co-operation in Nairobi for financial support, Prof. Ndinya Achola and the Microbiology Department of University of Nairobi for technical support, Prof. Paul Janssen, Mr Nico Nagelkerke and Dr Daniel Sondag for their advice and the Limburg University Centre for providing the website.

References

Babu G, Saraswathi NK, Vaidyanathan H & Jacob JT (1993) Reduction of the cost of testing for antibody to HIV virus, without losing sensitivity, by pooling sera. *Indian Journal of Medical Research* **97**, 1–3.

Behets F, Bertozzi S, Kasali M, Kashamuka M, Atikala L, Brown C *et al*. (1990) Successful use of pooled sera to determine HIV-1 seroprevalence in Zaire with development of cost-efficiency models. *Aids* **4**, 737–741.

Cahoon-Young B, Chandler A, Livermore T, Gaudino J & Benjamin R (1989) Sensitivity and specificity of pooled versus individual sera in a human immunodeficiency virus antibody prevalence study. *Journal of Clinical Microbiology* **27**, 1893–1895.

Hauck WW (1991) Confidence intervals for seroprevalence determined from pooled sera. *Annals of Epidemiology* **1**, 277–281.

Kline RL, Brothers TA, Brookmeyer R, Zeger S & Quinn TC (1989) Evaluation of human immunodeficiency virus seroprevalence in population surveys using pooled sera. *Journal of Clinical Microbiology* **27**, 1449–1452.

Perriens JH, Magazani K, Kapila N, Konde M, Selemani U, Piot P *et al*. (1993) Use of a rapid test and an ELISA for HIV antibody screening of pooled serum samples in Lubumbashi, Zaire. *Journal of Virological Methods* **41**, 213–222.

Sherlock CH, Strathdee SA, Le T, Sutherland D, O'Shaughnessy MV & Schechter MT (1995) Use of pooling and outpatient laboratory specimens in an anonymous seroprevalence survey of HIV infection in British Columbia, Canada. *Aids* **9**, 945–950.

Tu XM, Litvak E & Pagano M (1994) Studies of Aids and HIV surveillance. Screening tests: can we get more by doing less? *Statisitics in Medicine* **13**, 1905–1909.