Epidemiol. Infect. (1999), **122**, 267–272. Printed in the United Kingdom

Non-participation bias in unlinked anonymous HIV-prevalence surveys in England and Wales

T. DUONG[1]*, A. E. ADES[1], P. ROGERS[2] AND A. NICOLL[2]

[1] *Department of Epidemiology and Public Health, Institute of Child Health, London WC1 1EH*
[2] *Public Health Laboratory Service AIDS & STD Centre at the Communicable Disease Surveillance Centre, 61 Colindale Avenue, London NW9 5EQ*

(*Accepted 20 November 1998*)

SUMMARY

The objective was to assess the potential bias in unlinked anonymous HIV-seroprevalence surveys from objections to specimens being included. Objection rates in seroprevalence surveys were examined. Statistically large clusters of objections were considered to be the result of health care worker behaviour, and were disregarded. Underlying objection rates were estimated from remaining data and compared to seroprevalence. Overall objection rates approached or exceeded seroprevalence in many participating centres. However, underlying objection rates declined with time while prevalences were generally unchanging. Also, underlying rates correlated poorly with observed seroprevalences. Findings were therefore consistent with processes producing the clusters of objections and underlying objection rates independently of serostatus of individuals. Although national seroprevalence estimates produced by the surveys are reasonably free from objection bias, regional seroprevalence estimates outside London remain vulnerable to bias as a result of some centres returning data whose quality cannot be guaranteed.

INTRODUCTION

Unlinked anonymous surveys have been used to monitor HIV-1 seroprevalence among specific populations in a number of countries [1–5]. In the UK, the Department of Health has supported similar programmes since 1990 [5]. Participants in the surveys in the UK have the opportunity of refusing testing. Although the number of refusers are minimal compared to those tested [5], they are sufficiently high in relation to the number of seropositives to have the potential to bias seroprevalence estimates. In addition, it was shown that among STD clinic attenders in the Netherlands where participants are individually asked to participate (voluntary unlinked anonymous testing), HIV seroprevalence is likely to be higher among those who refuse unlinked testing compared to those accepting [6].

We report here on an analysis aimed at assessing the extent and implications of non-participation bias in the major unlinked anonymous surveys conducted in England. These are based on residual specimen left over after completion of routine clinical tests for Sexually Transmitted Disease clinic attenders (STD) and for women attending for antenatal care and on infant samples collected for routine metabolic testing.

METHOD

Unlinked anonymous testing is currently underway in 15 STD clinics (7 in London, 8 elsewhere) and in 26 antenatal units (15 in London, 11 in Yorkshire). The neonatal survey covers 70% of newborns in England

* Author for correspondence.

Table 1. *Example of clusterings of objections to testing among heterosexual women attending an STD clinic*

Year	1992	1993				1994				1995				1996			
Quarter	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4
Objections	3	7*	11*	0	1	9*	21*	57*	41*	3	1	1	3	5	2	0	0
Tested	156	159	193	254	135	312	286	369	401	458	469	455	442	487	335	394	364
Seropositives	0	0	1	1	2	2	0	4	0	3	3	2	0	3	2	4	4

* Observation identified to be an outlier and hence was removed from estimate of underlying objection rate.

and Wales [7]. Enrolments of centres were staggered. Before testing, all personal identifiers are irreversibly removed from the specimens. Survey collaborators are requested to display and make available multilingual posters and leaflets for patients which detail the aims and methodology of the programme, and which explain that if an individual objects to testing, their wishes will be respected. Healthcare workers are not supposed to bring up the issue of anonymous testing with individuals but they are supposed to be informed of the survey in order to answer any questions raised.

A preliminary examination of the data showed that in some participating centres there is a tendency for objections to cluster together in time. An example of such heterogeneity in objection rates can be seen in the data from a single centre shown in Table 1. One interpretation of this pattern is that there is a continuous, rather low, underlying objection rate (UOR) on which sporadic large clusters of objections are overlaid. Seroprevalence, by contrast, remains generally smooth over time, as would be expected from a process that reflects cumulative incidence over a long period [8]. Anecdotal evidence suggests that these clusters of objections may be due to the way in which particular health workers present the testing to patients rather than the risk behaviours or perceptions of the patients themselves [9]. The approach adopted here is therefore based on the premise that the clusters do not represent seropositive subjects but that those in the UOR may do. As a first step, clusterings of objections are systematically identified and excluded to allow UOR to be estimated.

Seroprevalence results up to the end 1996 were summarized by centres and into either quarterly intervals or, in the case of neonatal surveys where there were few objections, half-year intervals. Separate analyses were conducted for each of the male homosexual, male heterosexual and female heterosexual STD surveys and the antenatal and neonatal surveys. Assuming objection to testing is binomially distributed, the probability of individuals objecting is fitted using binomial logistic regression with explanatory variables centre and time from when centres entered the survey. The inclusion of a linear trend in time is conservative as it allows for the possibility that centres have high objection rates when the surveys are initially implemented [9].

Clusterings of objections are suggested where the number of objections observed within a time period at a centre is unduly higher than the number expected from the logistic model. Three different statistical criteria, as outlined below, are applied in sequence to the data and the first centre-time observation to satisfy any of them is considered to be an outlier. Statistical tests for outliers are one-tailed and set at 5% significance level.

(a) Likelihood residuals which are computed from the logistic model provide a basis for measuring how consistent an observation is in relation to the rest of the data. A test-statistic, τ, is assigned the value of the maximum positive likelihood residual. Assuming that the residuals are normally distributed, the observation yielding this residual is an outlier if τ is greater than the standard normal deviate corresponding to the area $(1-0{\cdot}05/n)$, where n is the number of observations [9]. For example, if $n = 1000$, τ would have to be greater than 3·89.

(b) By comparing the observed results with what would have been expected by simulation we can avoid making any assumption of the distribution of the residuals. The data are simulated 99 times based on the fitted binomial distributions. These are then used to construct simulated envelopes in which 90% of observations are expected to fall [10, 11]. The observation with the largest likelihood residual is an outlier if it falls beyond the upper 95% boundary of the envelope.

(c) It was found that centres in which objections were high in several time periods without having any observations meeting either of the above two criteria for being outlying tend to be characterized by a large

mean deviance under the logistic model. By using the simulated data, we can obtain a distribution of the mean deviance for each centre which can be compared to the observed value. If the maximum observed mean deviance exceeds its 95% centile, the observation with the largest positive deviance residual from this centre is then an outlier.

The centre-time observation identified as outlying is then removed and the whole process repeated sequentially until no more outliers remain.

Underlying objection rates and seroprevalence are estimated for each centre based on the remaining data. To assess the degree of potential bias for regional surveys, stratified by London and elsewhere, we compare the average UOR weighted by size of centres to the average seroprevalence. A 'worst case' scenario is considered in which all underlying objections represent infected individuals. A ratio of UOR to seroprevalence of one would therefore indicate that the true number of seropositives could potentially be underestimated by a factor of two. In order to assess how individual centres perform in terms of producing reliable results, their estimated UORs are compared to the average regional prevalence, stratified by London and elsewhere. This allows for a situation where, in centres with extremely low estimated seroprevalence, even if all objectors were seropositive, the true prevalence could still remain low. Because there is a need to consider current performance of the surveys, there is a focus on results for the recent period of 1994–6 in order to take into account changes in objection rates with duration of surveys. The correlation between UOR and seroprevalence was also examined.

RESULTS

Observed objection rate

Table 2 summarizes observed objection rates in relation to seroprevalence. Objection rates in the STD surveys are surprisingly homogenous, between 0·44 and 0·73%, in spite of very large differences in seroprevalence (0·10% female heterosexual attenders outside London to 15·2% male homosexuals within London). As a result, the male homosexual survey is least vulnerable to bias, because of the low ratio of objections to seropositives, and the female heterosexual survey outside London the most vulnerable, with objections exceeding seropositives by a factor of 5·9.

Similarly, while objection rates were constant at 0·10% in the antenatal survey within and outside London, potential vulnerability to bias was extreme outside London, where there were only 0·01% seropositives. The neonatal surveys had the lowest observed objection rates (0·1–0·4%), but again surveys outside London were vulnerable to extreme bias due to the very low seroprevalence.

Complete participation was achieved in 11 (50%) of the units in the antenatal survey and 77 (64%) of the centres participating in the neonatal survey. However, all 15 clinics in the STD surveys encountered non-participation in all risk groups apart from one where none of the homosexual men recorded any objection to testing.

Estimated underlying objection rates

All surveys included centres with a clustering of objections such as that in Table 1. Such clusters, while only occurring between 0·3% (neonatal survey in London) to 11·4% (female heterosexual STD survey outside London) of the total time periods in the surveys, accounted for a proportionately large number of objections (Table 2). The effect of removing the clusters on the UOR was only slight in the male homosexuals attending STD clinics. In the other surveys estimated underlying objection rates fell to between 30 and 60% of the observed rates and in the neonatal survey in London, it was as little as 8% of the observed rate (Table 2). In contrast, re-estimated seroprevalence based on the remaining data (which can be obtained from the UOR and its ratio to seroprevalence in Table 2) hardly changed.

In all surveys, participation rates improved significantly over time (Fig. 1 illustrates this for the antenatal survey), generating slight reductions in estimated UOR for the period 1994 and 1996 (Table 3). The reduction in UOR was substantial for the antenatal and neonatal surveys in London, and no objections in neonatal surveys were recorded in London in this period.

However, in spite of the general improvement, estimated UOR between 1994 and 1996 remained high relative to the seroprevalence in an appreciable number of centres, virtually all outside London (Table 3). Estimated UOR was at least 5 times higher than regional seroprevalence in 13 centres. In 5 of the 15 male heterosexual STD clinics, 7/15 female heterosexual STD clinics, 2/26 antenatal surveys and

Table 2. *Summary of observed and estimated underlying objection rate in unlinked anonymous HIV-surveys between 1990 and 1996*

Surveys	No. of centres	Sample size	% seropositive	% objection rate (ratio of obj. rate to seroprevalence)		No. (%) time periods with clusters of objections
				Observed	Underlying	
Male homosexual STD						
London	7	21878	15·24	0·61 (0·04)	0·50 (0·03)	3 (2·4)
Elsewhere	8	7439	4·05	0·55 (0·14)	0·50 (0·12)	1 (0·5)
Male heterosexual STD						
London	7	50238	1·05	0·45 (0·43)	0·28 (0·27)	9 (7·3)
Elsewhere	8	90269	0·16	0·44 (2·69)	0·27 (1·74)	12 (6·6)
Female heterosexual STD						
London	7	72828	0·69	0·73 (1·06)	0·42 (0·58)	14 (11·0)
Elsewhere	8	80206	0·10	0·57 (5·90)	0·23 (2·37)	20 (11·4)
Antenatal						
London	15	361368	0·27	0·10 (0·37)	0·03 (0·11)	14 (3·4)
Elsewhere	11	249623	0·01	0·10 (8·37)	0·06 (5·34)	4 (1·4)
Neonatal						
London	29	642862	0·16	0·008 (0·05)	0·0006 (< 0·01)	1 (0·3)
Elsewhere	92	1394140	0·01	0·04 (3·10)	0·02 (2·04)	23 (2·8)

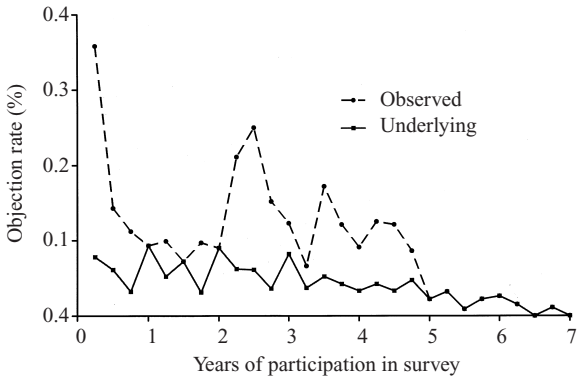

Fig. 1. Observed and estimated rate of objection in unlinked anonymous testing in the antenatal survey.

26/92 neonatal surveys, UOR exceeded regional seroprevalence. It can be seen that exclusion of these centres would have the effect of minimizing the potential bias in the surveys, except possibly in the female heterosexual STD survey in London, and in the male heterosexual survey outside London, where if all objectors were infected, the true seroprevalence could be approx. 1·5 times higher than observed.

Correlations between seroprevalence and UOR were 0·01 for male homosexuals, −0·03 for male heterosexual, 0·46 ($p = 0·08$) for female heterosexual STD clinic attenders, −0·24 in the antenatal survey, and −0·13 in the neonatal survey.

DISCUSSION

Although the numbers of objectors in unlinked anonymous surveys have been minimal in comparison to the numbers tested [5], they are quite substantial when compared to the number of seropositives, particularly outside London. The pattern of objections, however, has several general features that suggest that the processes generating objections are unrelated to serostatus of individuals. Firstly, objection rates are characterized by many sudden spikes, unlike seroprevalence which is quite smooth over time. Secondly, although seroprevalence varies very widely between centres inside and outside London, observed objection rates were very similar, in each of the five surveys. Thirdly, while seroprevalence has remained generally constant or has risen over time [5, 7], objection rates have fallen.

The more formal analysis relied on identification and removal of observation periods with statistically high objection rates, on the grounds that they were due to healthcare worker behaviour and unrelated to seroprevalence. This was verified by the fact that seroprevalence within these periods did not differ from the rest of the data.

The extent of bias in these surveys will depend on whether seroprevalence is higher among objectors than non-objectors. It is reassuring that the underlying

Table 3. *Estimated underlying objection rate (UOR) in unlinked anonymous HIV-surveys compared to seroprevalence between 1994 and 1996*

			Excluding centres with ratio of UOR to regional prevalence			
			> = 5		> = 1	
Surveys	No. of centres	UOR (UOR: seroprevalence)	UOR (UOR: seroprevalence)	No. of centres excluded	UOR (UOR: seroprevalence)	No. of centres excluded
Male homosexual STD						
London	7	0·36 (0·03)	0·36 (0·03)	0	0·36 (0·03)	0
Elsewhere	8	0·33 (0·10)	0·33 (0·10)	0	0·33 (0·10)	0
Male heterosexual STD						
London	7	0·25 (0·24)	0·25 (0·24)	0	0·25 (0·24)	0
Elsewhere	8	0·23 (1·70)	0·16 (1·14)	1	0·06 (0·47)	5
Female heterosexual STD						
London	7	0·40 (0·54)	0·40 (0·54)	0	0·31 (0·46)	1
Elsewhere	8	0·20 (2·52)	0·08 (0·81)	2	0·03 (0·17)	6
Antenatal						
London	15	0·01 (0·04)	0·01 (0·04)	0	0·01 (0·04)	0
Elsewhere	11	0·04 (3·14)	0 (0)	2	0 (0)	2
Neonatal						
London	29	0 (0)	0 (0)	0	0 (0)	0
Elsewhere	92	0·02 (1·69)	0·01 (0·74)	8	0·002 (0·12)	26

objection rate was not shown to be correlated with seroprevalence, which would have suggested seropositive individuals were consistently more likely to refuse testing than those uninfected. In both the antenatal and neonatal surveys, the objection rate was in fact lowest in London where seroprevalence was highest. Furthermore, the disparity in seroprevalence between STD surveys was not mirrored by similar patterns in objection rates. However, the possibility of differential behaviour in seropositive individuals within and outside London cannot be ruled out.

Even if all objectors were seropositive, it was found that seroprevalence estimates in all of the London surveys would not be substantially higher. This is a particularly useful finding as HIV prevalence in London is considerably higher than elsewhere. Other surveys (all outside London) remained quite severely vulnerable to bias even after removal of time periods with extreme numbers of objections. This could, however be minimized by excluding centres where underlying objection rates were high compared to regional seroprevalence. If accurate monitoring of regional seroprevalence outside London is to be ensured, action should be taken to monitor and improve participation in these centres, otherwise, the value of their surveys needs to be assessed.

Underlying objection rate was highest in the male homosexual STD survey and lowest in the neonatal survey. It is possible that refusal of testing is more likely in settings where individuals are also being confronted with the issue of voluntary named testing. Since patient data and specimens are collected at the same time in the STD surveys, it is also possible that healthcare workers in these settings are more likely to discuss anonymous testing in a similar way to named testing. Improving the understanding of anonymous testing programmes by healthcare workers could lead to improved participation rates.

Although these conclusions are generally reassuring and support the general view that the unlinked anonymous HIV-surveys are generating reliable seroprevalence data, it has to be of some concern that a proportion of centres have been returning data whose quality cannot be guaranteed. The potential for using routine serological samples to study prevalence of other infection in the UK is now being realized. For example, some samples from the antenatal, STD and neonatal surveys are now being tested for hepatitis C antibody and anonymous studies of human T-cell leukaemia/lymphoma virus have been carried out locally [12]. The approach presented here provide a means for monitoring potential bias at local and regional levels in unlinked anonymous surveys where individuals have the option to refuse testing.

The views in this paper are those of the authors alone.

REFERENCES

1. HIV and STD Division. Sexual health and health care: HIV, AIDS and sexually transmitted infections – global epidemiology, impact and prevention. London: Overseas Development Administration, 1996.
2. Dondero TJ, Gill ON. Large-scale HIV serologic surveys: what has been learned? AIDS 1991; **5** (suppl 2): S1–7.
3. European Centre for the Epidemiological Monitoring of AIDS. European HIV prevalence database. Updated version, 1992.
4. US Bureau of the Census. Recent HIV seroprevalence levels by country. US Health Studies Branch, International Programs Center, Population Division, 1996; 21.
5. Unlinked Anonymous HIV Surveys Steering Group. Unlinked anonymous HIV prevalence monitoring programme: England and Wales, data to the end 1996. London: Department of Health, Public Health Laboratory Services, Institute of Child Health (London), 1997.
6. Postema EJWP, Ridder de MAJ. Comparison of patients refusing with patients accepting unlinked anonymous HIV testing in an outpatient STD department in The Netherlands. Int J STD AIDS 1997; **8**: 368–72.
7. Nicoll A, McGarrigle C, Brady AR, et al. Epidemiology and detection of HIV-1 among pregnant women in the United Kingdom: results from national surveillance 1988–96. BMJ 1998; **316**: 253–8.
8. Ades AE. Serial HIV seroprevalence surveys: interpretation, design, and role in HIV-AIDS prediction. J Acquir Immune Defic Syndr 1995; **9**: 490–9.
9. Unlinked Anonymous HIV Surveys Steering Group. Unlinked anonymous HIV seroprevalence monitoring programme in England and Wales, data to the end of 1993. Department of Health, Public Health Laboratory Service, Institute of Child Health (London), 1995.
10. Collet D. Modelling binary data. London: Chapman and Hall, 1994: 121–46.
11. Atkinson A. Two graphical displays for outlying and influential observations in regression. Biometrika 1981; **68**: 13–20.
12. Hale A, Leung T, Sivasubramaniam S, et al. Prevalence of antibodies to HTLV in antenatal clinic attenders in south east London. J Med Virol 1997; **52**: 326–9.