32 Andersson SO, Adami HO, Bergstrom R, Wide L. Serum pituitary and sex steroid hormone levels in the aetiology of prostatic cancer—a population-based case-control study. *Br J Cancer* 1993;68:97-102.
33 Ross RK, Bernstein L, Lobo RA, Shimizu H, Stanczyk FZ, Pike MC, *et al.* 5-alpha-reductase activity and risk of prostate cancer among Japanese and US white and black males. *Lancet* 1992;339:887-9.
34 Klopper A, Jandial V, Wilson G. Plasma steroid assay in the assessment of foetoplacental function. *J Steroid Biochem* 1975;6:651-6.
35 McFadyen IR, Worth HG, Wright DJ, Notta SS. High oestrogen excretion in pregnancy. *Br J Obstet Gynaecol* 1982;89:994-9.
36 Gerhard I, Vollmar B, Runnebaum B, Kliilnga K, Haller U, Kubli F. Weight percentile at birth. II. Prediction by endocrinological and sonographic measurements. *Eur J Obstet Gynecol Reprod Biol* 1987;26:313-28.
37 Petridou E, Panagiotopoulou K, Katsouyanni K, Spanos E, Trichopoulos D. Tobacco smoking, pregnancy estrogens and birth weight. *Epidemiology* 1990;1:247-50.
38 Long PA, Abell DA, Beischer NA. Fetal growth and placental function assessed by urinary estriol excretion before the onset of pre-eclampsia. *Am J Obstet Gynecol* 1979;135:344-7.
39 Rosing U, Carlstrom K. Serum levels of unconjugated and total estrogens and dehydroepiandrosterone, progesterone and urinary estriol excretion in pre-eclampsia. *Gynecol Obstet Invest* 1984;18:199-205.
40 Bernstein L, Depue RH, Ross RK, Judd HL, Pike MC, Henderson BE. Higher maternal levels of free estradiol in first compared to second pregnancy: early gestational differences. *J Natl Cancer Inst* 1986;76:1035-9.
41 Panagiotopoulou K, Katsouyanni K, Petridou E, Garas Y, Tzonou A, Trichopoulos D. Maternal age, parity and pregnancy estrogens. *Cancer Causes Control* 1990;1:119-24.
42 Nomura AMY, Kolonel LN. Prostate cancer: a current perspective. *Am J Epidemiol* 1991;13:200-27.
43 Loeb LA. Mutator phenotype may be required for multistage carcinogenesis. *Cancer Res* 1991;51:3075-9.
44 Wang X, Guyer B, Paige DM. Differences in gestational age-specific birthweight among Chinese, Japanese and White Americans. *Int J Epidemiol* 1994;23:119-28.
45 Sakr WA, Haas GP, Cassin BF, Pontes FE, Crissman JD. The frequency of carcinoma and intraepithelial neoplasia of the prostate in young male patients. *J Urol* 1993;150:379-85.
46 Cheville JC, Bostwick DG. Postatrophic hyperplasia of the prostate. A histologic mimic of prostatic adenocarcinoma. *Am J Surg Pathol* 1995;19:1068-76.

(Accepted 1 June 1996)

Fluctuations of HIV load in semen of HIV positive patients with newly acquired sexually transmitted diseases

M C Atkins, E M Carlin, V C Emery, P D Griffiths, F Boag

Most HIV infections worldwide are acquired sexually, yet most coital episodes involving an infected partner do not result in acquisition of HIV. Acquisition of HIV must depend on both the volume of secretions transferred from the infected partner (donor) and the concentration of HIV present in the secretions. However, exposure alone to such a virus inoculum is clearly insufficient to ensure transmission. The coexistence of other sexually transmitted diseases in either the recipient or donor could potentially increase the risk of transmission by causing genital ulcers or by releasing inflammatory cytokines which increase HIV replication.[1]

We measured the HIV load in the semen and blood of HIV infected patients presenting with acute sexually transmitted diseases to determine if viral load in semen decreases when patients receive specific treatment for sexually transmitted diseases.

Subjects, methods, and results

Between October 1994 and June 1995, four asymptomatic homosexual men positive for HIV-1 with CD4 T cell counts between 330 and 600 cells/μl participated in the study. The men had been HIV positive for 3-10 years and none was receiving anti-retroviral treatment. The patients presented with a purulent urethral discharge which developed after unprotected orogenital sex. In each case, microscopy of a Gram stained urethral smear revealed 10 or more neutrophils per high power field (×1000), and in three cases (patients 1, 2, and 4) Gram negative intracellular diplococci were identified; urethral infection with *Neisseria gonorrhoea* was diagnosed at presentation and subsequently confirmed by culture. Non-gonococcal urethritis was diagnosed in patient 3. Patients with gonococcal urethritis were treated with single oral doses of ciprofloxacin 500 mg (patients 1 and 4) or ampicillin 3 g with probenecid 1 g (patient 2). The patient with non-gonococcal urethritis received doxycycline 100 mg bd for seven days. Subsequent urethral samples from patients 1, 3, and 4 were non-purulent and negative on microscopy and culture. Patient 2 had a post-gonococcal urethritis with a persistent discharge which resolved when treated with a course of doxycycline. No patient had any signs or symptoms of prostatitis. Blood and semen samples were collected on the day of presentation (before treatment) and at subsequent visits.

DNA was extracted from whole blood or semen using a commercially available DNA extraction method (Digene Laboratories, Germany). Quantitative polymerase chain reaction for proviral DNA was performed using nested primers to the *gag* region of HIV-1 (outer primers 5' GAGGAGCCACCCCACAATATT and 5' TAGGTGGATTATTTGTCATCCA; inner primers 5' TGCTAAACACAGTTGGGGGGA and 5' CCTGAAGGGTACTAGTAGTT) and coamplification of an internal control sequence.

Figure 1 shows HIV proviral loads during the period of treatment. HIV-1 proviral load in the semen of these subjects declined ($P<0.05$ Mann-Whitney test) when their intercurrent sexually transmitted diseases were treated; there were no significant changes in the blood.

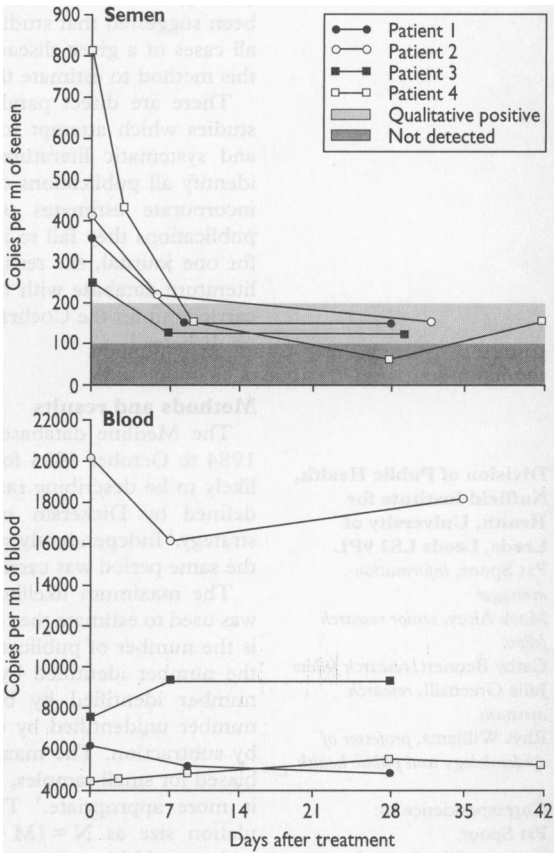

Fig 1—*Fluctuations in HIV proviral loads in serial semen and blood samples from four patients*

Department of Virology, Royal Free Hospital School of Medicine, London NW3 2QG
M C Atkins, *research fellow*
V C Emery, *reader in virology*
P D Griffiths, *professor of virology*

Department of Genitourinary Medicine, Royal Chelsea and Westminster Hospital, London SW10 9TH
E M Carlin, *senior registrar*
F Boag, *consultant physician*

Correspondence to:
Dr Boag.

BMJ 1996;313:341–2

Comment

The presence of HIV in semen has been well documented,[2 3] but the relation between the viral load in semen and peripheral blood CD4 counts is not a simple one.[3] No studies have looked at the serial viral load of genital fluids during treatment for other sexually transmitted diseases, although a single case report has suggested that chlamydial urethritis may increase shedding of HIV-1 in the semen.[4] Our results help explain how transmission of HIV may be facilitated by concomitant sexually transmitted diseases and add further support for an aggressive approach to treating sexually transmitted diseases in HIV infected patients, as a means of reducing transmission of HIV and for reinforcing the benefits of using condoms.[5]

Funding: Part of this work was supported by the UK Medical Research Council and by the St Stephens AIDS Trust.
Conflict of interest: None.

1 Gaynor R. Cellular transcription factors involved in the regulation of HIV-1 gene expression. *AIDS* 1992; 6:347-63.
2 Zagury D, Bernard J, Leibowitch J, Groopman JE, Feldman M, Gallo RC. HTLV-III in cells cultured from semen of two patients with AIDS. *Science* 1984; **226**:449-51.
3 Vernazza PL, Eron JJ, Cohen MS, van der Horst CM, Troiani L, Fiscus SA. Detection and biological characterization of infectious HIV-1 in semen of seropositive men. *AIDS* 1994;8:1325-9.
4 Eron J, Gilliam B, Fiscus S, Dyer J, Cohen MS. HIV-1 shedding and chlamydia urethritis. *JAMA* 1996;275:38.
5 Grosskurth H, Mosha F, Todd E, Murijarubi E, Klokke A, Senkoro K, *et al.* Impact of improved treatment of sexually-transmitted diseases on HIV-1 infection in rural Tanzania: randomized controlled trial. *Lancet* 1995;**346**:530-6.

(Accepted 24 April 1996)

Use of the capture-recapture technique to evaluate the completeness of systematic literature searches

Pat Spoor, Mark Airey, Cathy Bennett, Julie Greensill, Rhys Williams

Capture-recapture methods were pioneered in ecology and derive their name from censuses of wildlife in which several animals are captured, marked, released, and subject to recapture. In epidemiology the technique examines the degree of overlap between two (or more) methods of ascertainment and uses a simple formula to estimate the total size of the population. When the number already identified is subtracted from this estimate the number of cases not ascertained by either (or any) of the methods can then be calculated. It has been suggested that studies which attempt to ascertain all cases of a given disease in a population should use this method to estimate the number of missing cases.[1 2]

There are direct parallels between epidemiological studies which attempt to ascertain all available cases and systematic literature searches which attempt to identify all publications on a given topic: both should incorporate estimates of the number of cases or publications they fail to identify. Our study compared, for one journal, the results of searching an electronic literature database with those of hand searching, both carried out for the Cochrane collaborative review group on diabetes.[3]

Methods and results

The Medline database was searched from January 1984 to October 1994 for articles in *Diabetic Medicine* likely to be describing randomised controlled trials, as defined by Dickersin *et al* and using their search strategy.[4] Independently, a handsearch of the journal for the same period was carried out, with the same aim.

The maximum likelihood estimator, $N = M(n/m)$, was used to estimate the total population size,[2] where M is the number of publications identified by Medline, n the number identified by hand searching, and m the number identified by both sources. The estimated number unidentified by either method was calculated by subtraction. The maximum likelihood estimator is biased for small samples, for which Chapman's method is more appropriate.[5] This estimates the total population size as $N = (M + 1)(n + 1)/(m + 1) - 1$. The variance of N is estimated as $\mathrm{Var}(N) = (M + 1)(n + 1)(M - m)(n - m)/((m + 1)2(m + 2))$, from which 95% confidence intervals can be constructed.

Table 1—*Extent of overlap in the number of publications found by Medline search of "Diabetic Medicine" (January 1984 to October 1994) and by hand searching*

		Medline search	
		Found	Not found
Hand search	Found	115	35
	Not found	8	2*

*Estimated by capture-recapture technique, rounded to the nearest whole number.

Table 1 shows the number of publications identified by each method and the overlap. The articles missed by the hand search are attributed to human error; those not identified by the Medline search were improperly indexed, either because until recently no appropriate methodological subject heading existed or because the abstract failed to describe the study design. For our data both the maximum likelihood estimator and Chapman's method gave the same estimate of total population size (160, 95% confidence interval 158 to 164) rounded to the nearest whole number. The number of articles "missed" was 2 (0 to 6).

Comment

A caveat to the application of these methods is that if there is positive dependency between the two sources—that is, if an article identified by hand searching is more likely to be ascertained in Medline than one not so identified—then the estimates will underestimate the true population. If, however, Medline and the hand search are negatively dependent then the estimates will overestimate the true population.[2] Log-linear modelling offers an alternative approach to modelling dependency among data, where it is present.

The term capture-recapture is not so appropriate for the technique's use in epidemiology or literature searches since, while cases and publications may be said to be "captured," nothing is being "recaptured." As applied in epidemiology the method has been termed "ascertainment intersection."[2] However, we suggest the more informative descriptor "comparison of multiple methods of ascertainment" (or COMMA) for this useful technique, which we advocate for all systematic literature searches.

We thank Mr Alan S Rigby for statistical advice.
Funding: Northern and Yorkshire Regional Health Authority and the British Diabetic Association.
Conflict of interest: None.

Division of Public Health, Nuffield Institute for Health, University of Leeds, Leeds LS2 9PL
Pat Spoor, *information manager*
Mark Airey, *senior research fellow*
Cathy Bennett, *research fellow*
Julie Greensill, *research assistant*
Rhys Williams, *professor of epidemiology and public health*

Correspondence to:
Pat Spoor.
p.a.spoor@leeds.ac.uk

BMJ 1996;**313**:342–3