Estimating the coverage of a targeted mobile tuberculosis screening programme among illicit drug users and homeless persons with truncated models

N. A. H. VAN HEST[1,2]*, G. DE VRIES[1,2], F. SMIT[3,4], A. D. GRANT[5]
AND J. H. RICHARDUS[2,6]

[1] *Tuberculosis Control Section, Municipal Public Health Service Rotterdam-Rijnmond, Rotterdam, The Netherlands*
[2] *Department of Public Health, Erasmus MC, University Medical Centre Rotterdam, Rotterdam, The Netherlands*
[3] *Trimbos Institute* (*Netherlands Institute of Mental Health and Addiction*), *Utrecht, The Netherlands*
[4] *Department of Clinical Psychology, Vrije Universiteit Amsterdam, Amsterdam, The Netherlands*
[5] *Statistics, Modelling and Bioinformatics Department, Centre for Infections, Health Protection Agency, London, UK*
[6] *Division of Infectious Disease Control, Municipal Public Health Service Rotterdam-Rijnmond, Rotterdam, The Netherlands*

SUMMARY

Truncated models are indirect methods to estimate the size of a hidden population which, in contrast to the capture–recapture method, can be used on a single information source. We estimated the coverage of a tuberculosis screening programme among illicit drug users and homeless persons with a mobile digital X-ray unit between 1 January 2003 and 31 December 2005 in Rotterdam, The Netherlands, using truncated models. The screening programme reached about two-third of the estimated target population at least once annually. The intended coverage (at least two chest X-rays per person per year) was about 23%. We conclude that simple truncated models can be used relatively easily on available single-source routine data to estimate the size of a population of illicit drug users and homeless persons. We assumed that the most likely overall bias in this study would be overestimation and therefore the coverage of the targeted mobile tuberculosis screening programme would be higher.

INTRODUCTION

The epidemiological pattern of tuberculosis (TB) in low-incidence countries is changing, with an increasing number of TB patients living in urban areas [1–3]. This is due to overrepresentation of immigrants from countries with a high incidence of TB in large cities, and to urban risk groups for TB such as illicit drug users and homeless persons [4–6]. Conventional TB control methods such as contact-tracing and preventive treatment are inadequate among marginalized care-avoiders [5, 7, 8]. As an alternative, radiological screening programmes for illicit drug users and homeless persons have been recommended in European cities [9–12].

TB re-emerged among illicit drug users and homeless persons in Rotterdam (population ~600 000) in 2001, after periodic radiological screening was

* Author for correspondence: N. A. H. van Hest, M.D., M.Sc., Consultant Tuberculosis Control Physician/Epidemiologist, Tuberculosis Control Section, Division of Infectious Disease Control, Municipal Public Health Service Rotterdam-Rijnmond, PO Box 70032, 3000 LP Rotterdam, The Netherlands.
(Email: vanhestr@ggd.rotterdam.nl)

discontinued in 1996. In response, a periodic radiological screening programme was re-introduced in May 2002, using a mobile digital X-ray unit (MDXU) and visiting day and night shelters and hostels for homeless persons, methadone-dispensing centres and safe drug consumption rooms for opiate users, as well as the street prostitution zone in Rotterdam. The programme aimed to screen clients of these facilities and services bi-annually [5, 13].

For priority setting, service planning and resource allocation it is necessary to know the number of persons in a targeted group. This number can also be used to assess the coverage of an intervention [14]. Often direct (enumeration) techniques are not feasible to estimate the size of hidden populations and indirect techniques have to be used. One such indirect technique, capture–recapture analysis [15–17], has been used to estimate the size of hidden populations, including illicit drug users [18, 19], and homeless persons [20, 21]. However, capture–recapture analysis preferably needs at least three linked data sources, which are not always available for hidden populations. As an alternative, truncated models are described in the literature [17, 22, 23]. Contrary to conventional capture–recapture analysis, truncated models can use frequency data from a single source of information. These models have been applied to estimate the size of hidden populations such as criminals [24, 25], illegal residents [26], and illicit drug users and homeless persons [14, 27–30].

The objective of this study is to estimate the coverage of a mobile TB screening programme among illicit drug users and homeless persons in Rotterdam, using simple truncated models.

METHODS

Ethics committee approval was not required for this study.

Study design, participants and study years

Participants in this descriptive study are individuals that use the services of shelters and hostels for homeless persons, methadone-dispensing centres or safe drug-consumption rooms for opiate users, or work in the street prostitution zone in Rotterdam, having at least one chest X-ray taken in the MDXU of the mobile TB screening programme between 1 January 2003 and 31 December 2005. Because 2002 was an incomplete year of screening and not all facilities were

visited twice by the MDXU these data were excluded. A proportion of individuals in the target group use multiple facilities and their chest X-ray can be taken at different locations, sometimes more than twice yearly. Chest X-rays were read by public health TB physicians on location or within a few working days at the Public Health Service.

Data collection and validation

Data on participants of the MDXU screening programme, such as name, date of birth, sex, date of chest X-ray and chest X-ray result, are routinely entered into the electronic Client Information System of the Tuberculosis Control Section of the Municipal Public Health Service Rotterdam-Rijnmond, using a unique personal identification number. To avoid misclassification of individuals due to clerical errors such as misspelling of names or typing errors, all names and dates of birth of the participants were double-checked in the Client Information System during data entry. Since 2005 the Universal Mobile Telecommunications System (UMTS) provides wireless connection between the MDXU and the Client Information System facilitating checking personal data of participants on location. The number of individuals participating in the TB screening programme and the frequency of their visits per year and for the total study period were extracted from the Client Information System.

Truncated models

The number of illicit drug users and homeless persons in the target group for the mobile TB screening programme, and hence the coverage of the programme, was estimated through simple truncated models. Although their results were expected to be similar, as two examples we used Zelterman's truncated Poisson mixture model and Chao's truncated heterogeneity model, which can be applied to frequency counts of observations of individuals in a single register [31–33]. Truncated models aim to estimate the number of unobserved persons in the (truncated) zero-frequency class based upon information of the lower observed frequency classes, assuming a specific truncated distribution of the observed data, e.g. Poisson, binomial or a mixture [17, 31–34]. Observed frequency distributions may not be strictly Poisson and to relax this assumption Zelterman and Chao based their models on a Poisson mixture distribution. This allows for

Table 1. *Total number of screened individuals per frequency class and number of chest X-rays taken in the mobile radiological tuberculosis screening programme among illicit drug users and homeless persons in 2003–2005 in Rotterdam, The Netherlands*

Frequency class	Number of individuals	Percentage	Number of chest X-rays
1x	1384	45·6	1384
2x	585	19·3	1170
3x	397	13·1	1191
4x	267	8·8	1068
5x	218	7·2	1090
>6x	183	6·0	1172
Total	3034	100	7075

greater flexibility and applicability on real life data because the models explicitly cater for departures from the strict Poisson assumption. Zelterman's Poisson mixture model of the estimated total population size, est(N), is given by

$$\text{est}(N) = \text{obs}(N)/[1 - \exp(-2f_2/f_1)]$$

and Chao's heterogeneity model by

$$\text{est}(N) = \text{obs}(N) + (f_1)^2/2f_2,$$

where f_1 denotes the number of persons falling in the first frequency class, f_2 denotes the number of persons falling in the second frequency class, obs(N) denotes the number of all observed individuals and exp is the exponential.

The simple truncated models do not need statistical packages and have performed well when compared to log-linear capture – recapture analysis [35]. They supposedly perform well even when data are sparse. Frequency data are less sensitive to privacy regulations. The truncated models of Zelterman and Chao were previously used to estimate the number of problematic illicit drug users in Rotterdam and detailed conceptual aspects of these models have been described [14, 28, 30]. An overview of a range of truncated models, is given elsewhere [22]. The underlying assumptions and limitations of truncated models will be discussed later.

Coverage

The annual coverage is defined as the number of individuals screened at least once per year [obs(N) or the annual case-ascertainment] divided by the estimated annual number of illicit drug users and homeless persons in the target group for periodic TB screening [est(N)], expressed as a percentage [obs(N)/[est(N)] × 100]. This definition is different from the use of the word coverage by Chao in her heterogeneity model article [31], which is related to the proportion of times that the confidence interval includes the true number of cases in a simulation study, or another well-known publication of Chao's, in which it is related to a measure to quantify the source overlap information [36].

RESULTS

Between 1 January 2003 and 31 December 2005 a total of 7075 chest X-rays were made of 3034 individuals. Table 1 shows the total number of screened individuals per frequency class and number of chest X-rays taken. Nearly half of the individuals screened (45·6%) entered the programme only once.

Table 2 shows the annual number of screened individuals, people not previously screened and number of X-rays taken, per frequency class and in total. The annual number of individuals screened gradually decreased over the years. The annual number of people not previously screened strongly decreased but in 2004 and 2005 a considerable number of these persons still entered the programme. The annual number of individuals in the first frequency class (seen once), second frequency class (seen twice) and total number of individuals screened respectively represent f_1, f_2 and obs(N) in the formula of the truncated models.

Table 3 shows the annual observed and estimated number of illicit drug users and homeless persons in the target group for periodic TB screening for the two truncated models, as well as the estimated coverage of the mobile TB screening programme. The estimates of Chao's model are slightly higher but in the same range as Zelterman's model. The radiological mobile targeted TB screening programme reaches about 63% of the estimated target population at least once per year. The intended coverage of the screening programme (at least two chest X-rays per person per year) was about 22%, 25% and 21% in 2003, 2004 and 2005, respectively.

DISCUSSION

Main findings

This study demonstrates that truncated models can be used relatively easily on available single-source

Table 2. *Annual number of individuals screened, people not previously screened and number of X-rays taken, per frequency class and in total in the mobile radiological tuberculosis screening programme among illicit drug users and homeless persons in 2003–2005 in Rotterdam, The Netherlands*

Frequency class	2003			2004*			2005*		
	Number of individuals screened	People not previously screened (%)	Number of chest X-rays	Number of individuals screened	People not previously screened (%)	Number of chest X-rays	Number of individuals screened	People not previously screened (%)	Number of chest X-rays
1x	1162	1162 (100%)	1162	1058	594 (56%)	1058	997	405 (41%)	997
2x	555	555 (100%)	1110	597	144 (24%)	1194	489	56 (11%)	978
⩾3x	107	107 (100%)	333	57	8 (14%)	179	21	3 (14%)	64
Total	1824	1824 (100%)	2605	1712	746 (44%)	2431	1507	464 (31%)	2039

* Corrected for screening of a large shelter in January 2005 planned for December 2004.

Table 3. *Annual number of observed and estimated individuals and coverage of the mobile radiological tuberculosis screening programme among illicit drug users and homeless persons in 2003–2005 in Rotterdam, The Netherlands*

Year	Zelterman's Poisson mixture model [32]				Chao's heterogeneity model [31, 33]			
	obs(N)	est(N)	95% CI	% covered	obs(N)	est(N)	95% CI	% covered
2003	1824	2964	(2803–3152)	(62%)	1824	3040	(2868–3241)	(60)
2004	1712	2531	(2411–2671)	(68%)	1712	2649	(2512–2810)	(65)
2005	1507	2411	(2274–2572)	(63%)	1507	2523	(2369–2705)	(60)

obs(N), Number of individuals observed; est(N), number of individuals estimated; CI, confidence interval.

routine data to estimate the size of a hidden population of illicit drug users and homeless persons. Our results show that a radiological mobile targeted TB screening programme among illicit drug users and homeless persons in Rotterdam reaches about two-third of the estimated target population at least once per year. Between 21% and 25% of the estimated target population meets the objective of the programme and has two or more chest X-rays taken per year.

Limitations

As with capture–recapture analysis, the validity of the estimates of truncated models depends on the possible violation of the underlying assumptions. These assumptions are perfect identification (i.e. no misclassification of the number of visits of one client), a closed population (i.e. no in-migration or out-migration in the time period studied), ideally but not necessarily a homogeneous population (i.e. no subgroups with markedly different probabilities to be observed and re-observed), a constant probability of being observed (i.e. there should be no individual behavioural response and the probability of being re-observed should not be influenced by the experience of a previous visit) and, as explained earlier in the Methods section, a specific truncated distribution of the observed data [14, 30].

Perfect identification assumption

In this programme individuals were assigned unique identification numbers in the Client Information System and personal identifiers were double-checked upon data entry to avoid misclassification. The staff of the facilities visited assisted the programme by providing a list of names and dates of birth of clients eligible for screening. Most clients had personal identification cards which were checked at screening. Furthermore, social workers from the services assisted on the day of screening which also reduced the possibility of misclassifying individuals. Violation of the perfect record-linkage assumption is therefore considered minimal.

Closed population assumption

To reduce bias as a result of violation of the closed population assumption we divided the study in 1-year periods. The MDXU visits each location for one day twice a year. This limits the opportunity for passers-by and short-term clients to be observed. Tables 1 and 2, however, show that every year a substantial number of people not previously screened enter the programme. These can be individuals belonging to the target group but not yet captured by the screening programme, individuals not belonging to the target group or individuals that recently joined the target group. Influx of the last two categories will result in annual estimates of the target population of long-term illicit drug users and homeless persons being too high and hence the estimate of the screening programme coverage being too low.

Homogeneity assumption

Some problematic illicit drug users and homeless persons, such as cocaine users or persistent rough sleepers, will never be reached. Their likelihood of attending the TB screening programme is zero because they never utilize the facilities and services. This group is not included in the truncated model estimate [28].

We cannot exclude individuals entering the screening programme, e.g. among individuals entering the programme only once, that do not belong to the group of long-term illicit drug user and homeless persons. In a previous conventional log-linear capture–recapture estimation of the number of clients of a methadone maintenance programme it was demonstrated that differences in capture probabilities of the population of interest, problematic drug users, and the sampled population, also including non-problematic drug users, could considerably overestimate the size of the population of interest [19].

We cannot exclude heterogeneity among individuals belonging to the target group entering the screening programme but the opportunities to participate in the screening (opting-out strategy) or not to participate (not attending the facility or service on the day of screening) are assumed to be largely similar for the majority. The truncated models are arguably more robust to violation of the homogeneity assumption because they are partly based upon the lower frequency classes, and assumed to have more resemblance to the zero frequency class. The relative insensitivity to violation of the homogeneity assumption of Zelterman's and Chao's models is also supported mathematically and through simulation studies [14, 22, 32]. However, in the presence of heterogeneity they can underestimate the population.

An alternative approach to estimating a heterogeneous population would be to use a population mixture model. Such a model (for the data in Table 1) regards the eligible population for each visit as a mixture of 'local clients at the facilities', having six opportunities to be observed and 'roaming clients' from other facilities, visiting more places than their own facility and can be captured at other facilities by the MDXU as well. They possibly have more than a total of six opportunities to be observed. For each visit the capture of the local clients could be modelled as binomial (6, p1) and the capture of roaming clients as Poisson (lambda), where lambda is probably less than 6 times p1. However, in our population of homeless persons and illicit drug users a clear distinction between local and roaming clients is arbitrary as many clients use multiple services, e.g. methadone-dispensing centres due to their addiction and day- and night-care facilities due to their homelessness, and their need for specific services may change over time. Furthermore, we have not considered such a population mixture model or E-M algorithm because, although more accurate, their complexity disagrees with the appealing ease of use of the simple truncated models. For the purpose of our study more exact, but complex to calculate, estimates were subsidiary to the simplicity of a method which should be close enough. As described for capture–recapture analysis simple truncated models are useful under certain circumstances, e.g. when the probable direction of the bias caused by violation of the underlying assumptions can be predicted and plausible lower and upper boundaries of the prevalence or incidence of a disease or the coverage of a community health-care intervention can be estimated [17, 37, 38].

Constant (re)observation probability assumption

For the majority of the individuals in the target group of the mobile TB screening programme the facilities and services where screening took place provide important needs, namely methadone and shelter. These needs are probably constant over time, creating a considerable probability of attending the services. Frequent users have the highest risk of TB but are also most likely to be screened. Although incentives, e.g. chocolate bars, were given to participants at some

locations, it is unlikely that this creates an important positive behavioural response to participate again. This also applies to clients with radiographic abnormalities inconsistent with TB as they are referred to a chest physician in one of the general hospitals in Rotterdam where further analysis and follow-up is performed. The opting-out strategy and (strong) persuasion by the staff of the social and medical services to participate prevents a negative behavioural response. The pressure particular institutions put on their clients to participate in the screening programme is considered relatively constant on each screening day. The coverage of the screening programme will never be perfect as each year a proportion of the target group will temporarily have a low or zero probability to attend, e.g. due to admission in a rehabilitation clinic or prison sentence. Finally, is has been explained elsewhere that the probability of being observed does not have to be constant as long as a capture or non-capture does not influence a possible change in probability [26].

Poisson distribution of the observed data assumption

Zelterman and Chao based their model on a Poisson mixture distribution, catering for departures from the strict Poisson assumption. We have examined whether the Zelterman model used tolerates the departures from the Poisson distribution observed in our data. We have performed negative binomial regression, with number of times screened as the covariate and number of individuals as the outcome, on the Table 1 data (counting >6 as 6). The variance of the data is larger than that of a Poisson distribution. This overdispersion is statistically significant ($P = 0.11$), but small (alpha $= 0.024$), and so does not invalidate the use of Zelterman's estimator [32]. Therefore it seems reasonable to use this simple model in the context of our study, as explained earlier.

A further limitation is that persons in the target group could have indicated on the day of screening that a chest X-ray was recently taken in the MDXU, in a general hospital, upon detention in prison or at the Tuberculosis Control Section upon referral, exempting them from the screening exercise. This information, together with improved experience, better coordination and UMTS access over the years, would prevent some clients from being recorded twice or more than twice yearly in the screening programme, as reflected in Table 2, leading to overestimation, but we assumed this effect to be limited.

Cross-validation of the estimates of the target group

The number of problematic illicit drug users in Rotterdam, already including many homeless persons, was most recently estimated in 2003 with two-source capture–recapture analysis, using a similar case-definition, which observed and estimated 1910 and 2856 clients respectively [39]. These numbers are similar to our results in 2003.

Alternative simple truncated models

Although we used truncated Poisson mixture models, an alternative is to use a truncated binomial model such as $\mathrm{est}(N) = \mathrm{obs}(N) + (f_1)^2/4f_2$. This model, close to Chao's model, estimates a lower number of 2432, 2181 and 2015 illicit drug users and homeless persons in 2003, 2004 and 2005 respectively, resulting in a slightly higher estimated coverage of the screening programme.

CONCLUSION

Although the limitations of the simple single-source truncated models should be appreciated and bias cannot be excluded, alternative methods for estimating illicit drug users and homeless persons have their own restrictions. Conventional two-source and three-source capture–recapture analysis have similar underlying assumptions and hence limitations, and for hidden populations sufficient adequate registers for record-linkage may not be available. Compared to alternative estimators the ease of use of the truncated models is appealing. We could extract, check and prepare the required data from an existing routine dataset in 2 days and calculate the point estimates on a pocket calculator. We assumed the probable overall bias in this study to be overestimation and therefore the coverage of the targeted mobile TB screening programme among problematic illicit drug users and homeless persons in Rotterdam would be higher than the 63% one chest X-ray per year and 21–25% for at least two chest X-rays per year, especially among those with the highest risk.

ACKNOWLEDGEMENTS

DECLARATION OF INTEREST

None.

REFERENCES

1. **Fujiwara PI, Frieden TR.** Tuberculosis epidemiology and control in the inner city. In: Rom W, Garay S, eds. *Tuberculosis*. New York: Little, Brown & Co., 1996, pp. 99–112.

2. **Moore-Gillon JC.** Tuberculosis and poverty in the developed world. In: Davis PD, ed. *Clinical Tuberculosis*. London: Chapman and Hall, 1998, pp. 383–396.

3. **Hayward AC, et al.** Epidemiology and control of tuberculosis in Western European cities. *International Journal of Tuberculosis and Lung Disease* 2003; **7**: 751–757.

4. **Valin N, et al.** Outbreak of tuberculosis in a migrants' shelter, Paris, France, 2002. *International Journal of Tuberculosis and Lung Disease* 2005; **9**: 528–533.

5. **De Vries G, Van Hest RA.** From contact investigation to tuberculosis screening of drug addicts and homeless persons in Rotterdam. *European Journal of Public Health* 2006; **16**: 33–36.

6. **Story A, Van Hest R, Hayward A.** Tuberculosis and social exclusion. *British Medical Journal* 2006; **333**: 57–58.

7. **Barnes PF, et al.** Patterns of tuberculosis transmission in Central Los Angeles. *Journal of the American Medical Association* 1997; **278**: 1159–1163.

8. **Bock NN, et al.** A tuberculin screening programme and isoniazid preventive therapy programme in an inner-city population. *American Journal of Respiratory and Critical Care Medicine* 1999; **159**: 295–300.

9. **Kumar D, et al.** Tuberculosis among the homeless at a temporary shelter in London: report of a chest x ray screening programme. *Journal of Epidemiology and Community Health* 1995; **49**: 629–633.

10. **Southern A, et al.** Tuberculosis among homeless people in London: an effective model of screening and treatment. *International Journal of Tuberculosis and Lung Disease* 1999; **3**: 1001–1008.

11. **Solsona J, et al.** Screening for tuberculosis upon admission to shelters and free-meal services. *European Journal of Epidemiology* 2001; **17**: 123–128.

12. **De Vries G, Van Hest NA, Šebek MM.** Active tuberculosis screening with mobile digital X-ray units among drug addicts and homeless people in Rotterdam. In: *Abstracts of the 3rd Congress of European Region International Union against Tuberculosis and Lung Disease*. Moscow, Russia: Russian Respiratory Society, 2004, p. 8.

13. **De Vries G, Van Hest R, Richardus JH.** Impact of mobile radiographic screening on tuberculosis among drug users and homeless persons. *American Journal of Respiratory and Critical Care Medicine* 2007. Published online: 5 April 2007. doi:10.1164/rccm.200612-1877C.

14. **Smit F, Reinking D, Reijerse M.** Estimating the number of people eligible for health service use. *Evaluation and Program Planning* 2002; **25**: 101–105.

15. **International Working Group for Disease Monitoring and Forecasting.** Capture-recapture and multiple-record estimation I: History and theoretical development. *American Journal of Epidemiology* 1995; **142**: 1047–1058.

16. **International Working Group for Disease Monitoring and Forecasting.** Capture-recapture and multiple-record estimation II: Applications in human diseases. *American Journal of Epidemiology* 1995; **142**: 1059–1068.

17. **Hook EB, Regal RR.** Capture-recapture methods in epidemiology: methods and limitations. *Epidemiologic Reviews* 1995; **17**: 243–263.

18. **Hay G, McKeganey N.** Estimating the prevalence of drug misuse in Dundee, Scotland: an application of capture-recapture methods. *Journal of Epidemiology and Community Health* 1996; **50**: 469–472.

19. **Buster MC, van Brussel GH, van den Brink W.** Estimating the number of opiate users in Amsterdam by capture-recapture: the importance of case definition. *European Journal of Epidemiology* 2001; **17**: 935–942.

20. **Fisher N, et al.** Estimating numbers of homeless and homeless mentally ill people in north east Westminster by using capture-recapture analysis. *British Medical Journal* 1994; **308**: 27–30.

21. **Gurgel RQ, et al.** Capture-recapture to estimate the number of street children in a city in Brazil. *Archives of Disease in Childhood* 2004; **89**: 222–224.

22. **Wilson RM, Collins MF.** Capture-recapture estimation with samples of size one using frequency data. *Biometrika* 1992; **79**: 543–553.

23. **Van Hest NA, et al.** Estimating infectious disease incidence: validity of capture–recapture analysis and truncated models for incomplete count data. *Epidemiology and Infection* Published online: 12 March 2007. doi:10.1017/S0950268807008254.

24. **Rossmo DK, Routledge R.** Estimating the size of criminal populations. *Journal of Quantitative Criminology* 1990; **6**: 293–314.

25. **Van der Heijden PG, Cruyff MJ, Van Houwelingen, H.** Estimating the size of a criminal population from police registrations using the truncated Poisson regression model. *Statistica Neerlandica* 2003; **57**: 289–304.

26. **Van der Heijden PG, et al.** Point and interval estimation of the truncated Poisson regression model. *Statistical Modelling* 2003; **3**: 305–322.

27. **Hser YI.** Population estimation of illicit drug users in Los Angeles County. *Journal of Drug Issues* 1993; **23**: 323–334.

28. **Smit F, Toet J, Van der Heijden PG.** Estimating the number of opiate users in Rotterdam using statistical models for incomplete count data. In: Hay G, McKeganey N, Birks E, eds. *Final report EMCDDA project Methodological Pilot Study of Local Level Prevalence Estimates*. Lisbon: EMCDDA, 1997, pp. 47–66.

29. **Bohning D, et al.** Estimating the number of drug users in Bangkok 2001: a capture-recapture approach using repeated entries in one list. *European Journal of Epidemiology* 2004; **19**: 1075–1083.

30. **Hay G, Smit F.** Estimating the number of hard drug users from needle-exchange data. *Addiction Research and Theory* 2003; **11**: 235–243.

31. **Chao A.** Estimating the population size for capture-recapture data with unequal catchability. *Biometrics* 1987; **43**: 783–791.

32. **Zelterman D.** Robust estimation in truncated discrete distributions with application to capture-recapture experiments. *Journal of Statistical Planning and Inference* 1988; **18**: 225–237.

33. **Chao A.** Estimating animal abundance with capture frequency data. *Journal of Wildlife Management* 1988; **52**: 295–300.

34. **Chao A.** Estimating population size for sparse data in capture-recapture experiments. *Biometrics* 1989; **45**: 427–438.

35. **Hook EB, Regal RR.** Validity of Bernouilli census, log-linear and truncated binomial models for correcting for underestimates in prevalence studies. *American Journal of Epidemiology* 1982; **116**: 168–76.

36. **Chao A, et al.** The applications of capture-recapture models to epidemiological data. *Statistics in Medicine* 2001; **20**: 3123–3157.

37. **Hook EB, Regal RR.** Capture-recapture methods. *Lancet* 1992; **339**: 742.

38. **Hook EB, Regal RR.** The value of capture-recapture methods even for apparent exhaustive surveys. The need for adjustment for source of ascertainment intersection in attempted complete prevalence studies. *American Journal of Epidemiology* 1992; **135**: 1060–1067.

39. **Biesma S, Snippe J, Bieleman B.** *Illicit Drug Users Registered.* Groningen: Intraval, 2004, p. 28.