American Journal of Epidemiology

Vol. 153, No. 7
Printed in U.S.A.

Estimating the Relative Incidence of Heroin Use: Application of a Method for Adjusting Observed Reports of First Visits to Specialized Drug Treatment Agencies

Matthew Hickman,[1] Shaun Seaman,[2] and Daniela de Angelis[3,4]

In this paper, the authors propose a method for estimating the incidence of heroin use by adjusting reported numbers of heroin users visiting drug treatment agencies for the time lag between onset of heroin use and first treatment request (lag distribution). The adjusted incidence is relative, since it represents the number of individuals beginning heroin use in each year whose cases will be reported within 8 years of starting use. Users with longer lag times or whose cases are never reported are excluded. Utilizing data from southeastern England (1991–1998), the authors analyzed the effects of covariates (sex, age group, ethnic group, route of consumption, and year of onset of drug use) on the lag distribution. Trends in the adjusted incidence of heroin use were very different for injectors and noninjectors: Incidence among injectors seemed to be stable, while in noninjectors it increased twofold between 1991 and 1996–1997. These results must be interpreted cautiously, especially in relation to the wider context of underlying trends in the population. Potential biases derive from underreporting and from changes in the proportion of heroin users in treatment. The lag correction method adds substantially to the value of routine treatment data, at least for heroin use, and is potentially the best method for obtaining estimates of incidence. *Am J Epidemiol* 2001;153:632–41.

epidemiologic methods; heroin; incidence; substance abuse, intravenous; substance-related disorders

Estimates of the occurrence of substance abuse are necessary to inform evidence-based policy-making and government strategies aiming at preventing drug use (1). Heroin use and injecting behavior, in particular, are of major public health importance, for two main reasons. Firstly, the risks associated with them, which include transmission of human immunodeficiency virus (HIV), hepatitis B and C, and fatal overdose, and the social costs of increased levels of crime and poverty are greater than those for many other drugs (2). Secondly, effective treatment options and interventions, such as prescription of substitute drugs and syringe exchange, are available (3, 4).

Recent research efforts in substance abuse epidemiology have been devoted to the problem of estimating prevalence through the use of indirect methods, such as capture-recapture (5, 6). Clearly, prevalence estimates are important for obtaining an overview of the impact of drug use in the population and for estimating the proportion of users who are in treatment. However, they are often not very informative, as the uncertainty associated with them is large. For example, the UK government recently estimated the number of severely dependent drug users to be between 100,000 and 200,000 nationally and between 30,000 and 70,000 in inner London (1, 7). More crucially, prevalence estimates do not provide information on trends in the spread of heroin use, which are of prime importance in deciding the appropriate balance in policy between primary and secondary prevention (8). Therefore, policy-makers need estimates of incidence to understand whether the rate of new heroin use is increasing, stable, or declining.

Heroin use was first described as an "epidemic" in the late 1960s and early 1970s, when several reports attempted to monitor the size and growth of heroin use in communities, mostly on the basis of fieldwork studies involving follow-up of patients in treatment (9–13). Hunt and Chambers (14–16) suggested some methods of estimating incidence from treatment data by adjusting such data for the delay between onset of use and first visit to a treatment agency. However, their ideas were not taken up by others. They were severely criticized at the time by several commentators who were convinced that such methods should not be pursued and that population-wide surveys would provide better information (17). The perceived need to estimate incidence receded, as it was considered constant (18).

Received for publication September 27, 1999, and accepted for publication May 4, 2000.

Abbreviations: AIDS, acquired immunodeficiency syndrome; HIV, human immunodeficiency virus.

[1] Centre for Research on Drugs and Health Behaviour, Social Science and Medicine, Imperial College School of Medicine, Reynolds Building, St. Dunstan's Road, London W6 8RP, United Kingdom.
[2] Institut National de la Santé et de la Recherche Médicale, Unité 170, 16 avenue Paul Vaillant-Couturier, 94807 Villejuif, France.
[3] MRC Biostatistics Unit, University of Cambridge, Forvie Site, Robinson Way, Cambridge CB2 2SR, United Kingdom.
[4] Communicable Disease Surveillance Centre, Public Health Laboratory Service, 61 Colindale Avenue, London NW9 5EQ, United Kingdom.
Reprint requests to Matthew Hickman, Centre for Research on Drugs and Health Behaviour, Reynolds Building, St. Dunstan's Road, London W6 8RP, United Kingdom (e-mail: m.hickman@ic.ac.uk).

However, general population surveys, although they may show trends in the use of cocaine and other drugs, in fact have not been shown to reliably detect changes in the incidence of heroin use, because it is too rare. Attempts to do so, even with large sample sizes of over 90,000 interviews, have failed because the confidence intervals are too wide to be informative (19, 20).

Mathematical modelers, when attempting to assess the transmission dynamics of HIV among injecting drug users, have restated the need for estimates of incidence, though with little success in obtaining them (21, 22). Some other indirect or modeling approaches to the estimation of incidence based on combining serial prevalence estimates and data from surveys of community-recruited drug users or constructing dynamic models also have failed to be adopted routinely (23–25). Equally, reliance on indirect indicators, such as accident and emergency records of drug-related problems or urine testing of arrestees, as markers of incidence is not completely successful. Firstly, such surveillance systems are rare outside of the United States. Secondly, problems with sampling, data completeness and accuracy, and policy changes have made trends difficult to interpret over time simply as expressions of incidence (26–29).

One of the recommendations of the US Institute of Medicine stemming from its review of drug abuse research (30) was further epidemiologic research designed to measure the extent of drug dependence. In contrast, in acquired immunodeficiency syndrome (AIDS) epidemiology, much effort has been devoted to the development of techniques aimed at interpreting and adjusting AIDS report data, since such data have been the basis for reconstructing the HIV epidemic and estimating future caseload (31). We believe that the most appropriate analogy between AIDS/HIV estimation work and the problem of estimating heroin incidence from observed reports of visits to treatment centers is the adjustment for "reporting delay." In the AIDS context, the problem of adjusting for reporting delay arises when complete information on AIDS incidence is required. Typically, the number of AIDS cases reported to surveillance centers seriously underestimates the number of recent AIDS diagnoses, because of substantial delay in reporting. Therefore, an estimate of AIDS incidence is obtained by adjusting reported data for reporting delay (32–34). The time interval between "onset of drug use" and "request for treatment" can be seen as analogous to the interval between "AIDS diagnosis" and "AIDS report," making the problem of estimating drug use incidence similar to that of estimating AIDS incidence. In this paper, we have adapted methods developed for adjusting AIDS reports to the drug use context, to demonstrate and pilot their use in estimation of the lag between onset of heroin use and first treatment request and, hence, the historical trends in heroin incidence.

MATERIALS AND METHODS

The data

The Drug Misuse Databases represent the main investment by the UK government in compiling routine statistics on drug abuse in the United Kingdom. A fuller description of this surveillance system is given elsewhere (35). Briefly, data on the sociodemographic characteristics and drug profile (including drug name, age at first use, and route of consumption for up to five illicit drugs) of people requesting treatment at a range of treatment agencies are collected using a standard reporting form. Initials, date of birth, and sex are used to identify all reports on the same individual. Reports are accepted on an ongoing basis or in batches for two 6-month reporting periods (April–September and October–March). All agencies are reminded to forward any outstanding reports before a subset of the data is passed on to the national system twice per year. Reports referring to a previous 6-month period are rare, obviating the need to also estimate and adjust for a genuine "reporting delay."

We extracted, from the North and South Thames East Regional Drug Misuse Database, all first reports mentioning heroin use. These reports included data on age at first use, date of report, date of birth, sex, route of consumption (injecting or noninjecting) at first treatment visit, and ethnic group (White (including "other"), Black (including Black Caribbean, Black African, and "Black other"), or Asian (including Pakistani, Indian, Bangladeshi, and "other Asian")). The database covers over 80 percent of Greater London (population aged 15–44 years = 3.14 million) and the surrounding counties, from Hertfordshire and Essex in the north to Kent and East Sussex in the south (population aged 15–44 years = 2 million) (36). Reports from the former South Thames West Drug Misuse Database were excluded, because the historical system did not allow identification of first reports. Only reports from specialized drug treatment agencies and National Health Service general practitioners, the main reporters to the Drug Misuse Database, were considered, in order to minimize potential bias from differential underreporting rates over time.

Table 1 shows the observed numbers of first reports of people who began using heroin between January 1, 1991, and September 31, 1998, by year of first use (onset). Year of first use was calculated using information on age at first use, age at report, and year of report. The following records were excluded from the analysis: 3,534 (11 percent) from reporters that only partially participated within the study time period; 6,635 (20 percent) with missing data on age at first use; 1,566 (5 percent) with an age at first use of <10 years or >30 years, because of concerns over the reliability of these data (see Discussion); and 13,712 (41 percent) with a year of onset prior to the year in which reporting began (i.e., 1991) (see below and Appendix).

Two alterations were made to the observed data in table 1. First, the number of reports for 1997 was adjusted to reflect the extraordinary increase in underreporting in that year arising from confusion over the closure of a statutory notification system (estimated to be an exceptional decrease in reports of approximately 16 percent) (37). Secondly, for 1998, the number of reports made through the end of September was scaled up (multiplied by 4/3) to be equivalent to the other annual figures.

TABLE 1. Observed numbers of first reports of heroin use, by year of report and year of initiation of use (onset), southeastern England, 1991–1998

Year of onset	Year of report								Total
	1991	1992	1993	1994	1995	1996	1997*	1998*	
1991	89	249	268	248	198	157	156	154	1,519
1992	0	137	248	286	262	258	191	179	1,561
1993	0	0	119	253	295	306	209	200	1,382
1994	0	0	0	132	303	356	244	215	1,250
1995	0	0	0	0	162	350	405	311	1,228
1996	0	0	0	0	0	191	387	407	985
1997	0	0	0	0	0	0	166	380	546
1998	0	0	0	0	0	0	0	143	143
Total	89	386	635	919	1,220	1,618	1,758	1,989	8,614

* Numbers of reports were adjusted for an increase in underreporting in 1977 and for 9 months of reporting (January–September) in 1998.

Statistical methods

As we noted above, the statistical problem of adjusting observed heroin incidence for the lag between a person's onset of use and his or her case's being reported is analogous to that of adjusting AIDS reports for reporting delay. We followed the methods used, in the context of AIDS reporting delay, by Brookmeyer and Liao (32). Full details are contained in the Appendix, but a brief account is given below.

We defined onset as a person's first use of heroin, and lag as the amount of time (in years) between onset of use and first report. Data were available from the Drug Misuse Database for 1991–1998. Thus, only individuals who visited a Database agency or physician between 1991 and 1998 had their cases reported. This means that there was left-truncation in the data set, because individuals who started using heroin before 1991—e.g., in the year "1991 – x"—could only be reported if they continued their drug use for at least x years. However, the data were also right-truncated, because an individual who began using heroin in "1990 + y" could only be reported if he or she first requested treatment within $9 - y$ years of beginning heroin use (i.e., before the end of 1998). The statistical technique used by Brookmeyer and Liao (32) (and by us) to estimate the lag distribution may be used with data that are right-truncated but not with data that are left-truncated. This is why we excluded from our analysis all individuals whose data were left-truncated—i.e., all individuals who began using heroin before 1991.

Additionally, individuals who started using heroin before 1991 could have had their cases reported earlier had the Drug Misuse Database been established earlier. Thus, for them, the distribution of the time between onset and first report would be artificially different from that of persons who initiated heroin use in 1991 or later. After exclusion of all individuals whose onset was prior to 1991, the maximum right-truncation time in the data set was 8 years (i.e., all observed lags were 8 years or less). Thus, the lag distribution being estimated was a conditional distribution: the distribution of the lag conditional on this lag's being 8 years or less.

The effects of age, sex, route of drug consumption at first treatment visit, and ethnic group on the lag distribution were assessed using a parametric model. The observed numbers of individuals beginning heroin use in each year $1990 + x$ were then adjusted according to the proportion of all cases reported within 8 years of onset that would be expected to be reported within $9 - x$ years of onset (i.e., the conditional lag distribution). No adjustment was required for 1991.

On the basis of results from the analysis of covariate effects on the lag distribution, we calculated the adjusted incidence of heroin use according to four models. In model 1, the same lag distribution was employed for all heroin users. In model 2, the population was divided into four strata—males and females and injectors and noninjectors—and the lag distribution was estimated separately for each stratum. Models 3 and 4 were like model 2, except that age group was also included as a three-level factor in a parametric model for the lag distribution. Model 4 also included year of onset as a continuous covariate in this parametric model.

The adjusted incidence calculated here is relative rather than absolute, since it represents the number of individuals beginning heroin use in each year who will have their cases reported within 8 years of starting. It excludes all persons who will have longer lag times or whose cases will never be reported (i.e., who will never seek treatment).

RESULTS

After ineligible reports were excluded, there were 7,824 observed cases (i.e., persons who started using heroin during or after 1991 and whose cases were reported by general practitioners and specialized treatment services in London and southeastern England between 1991 and 1998). Adjustment for extra underreporting in 1997 and for the incomplete year in 1998 increased the total number of cases to 8,614: 4,669 (54 percent) injectors and 3,945 (46 percent) noninjectors; 6,013 (70 percent) males and 2,601 (30 percent) females; 976 (11 percent) aged <20 years, 3,955 (46 percent) aged 20–24 years, and 3,683 (43 percent) aged ≥25 years at first treatment request; and 7,749 (90 percent) classified as White (or "other"), 377 (4 percent) as Black, and

488 (6 percent) as Asian (table 2). Table 1 shows the distribution of cases by year of onset and year of report (adjusted for 1997 and 1998); e.g., 89 heroin users started use in 1991 and were reported in 1991, and 356 heroin users started use in 1994 and were reported in 1996.

The overall conditional lag distribution (i.e., without stratification on or adjustment for covariates) was estimated. Let $F(x)$ denote the proportion of heroin users whose cases are reported by the end of the $(x - 1)$th calendar year after onset (conditional on being reported within 8 years). Thus, $F(1)$ is the proportion whose cases are reported in the same calendar year as onset, and $F(8) = 1.0$ by definition. The estimated values of $F(1)$, $F(2)$, …, $F(7)$ (and their standard errors) were, respectively, 0.07 (0.002), 0.23 (0.005), 0.41 (0.007), 0.56 (0.008), 0.69 (0.009), 0.80 (0.009), and 0.90 (0.008). Table 2 shows the effects of covariates on the lag distribution. A negative β-coefficient indicates that the covariate is associated with a shorter lag time, while positive β-coefficients indicate longer lags. Shorter lags mean that a greater proportion of heroin users will have had their cases reported within so many years of onset, and thus the observed incidence will need to be adjusted less, resulting in lower adjusted incidences. Year of onset was treated as a continuous covariate, since year of onset as a categorical variable was not independently significant in the model that also included it as a continuous covariate. Sex, age at first use, route of administration, and year of onset all had significant independent associations with the lag distribution. After adjustment for other covariates, the lag was shorter for females compared with males, for older age groups compared with persons aged <20 years, and for noninjectors compared with injectors. It has also shortened over time since 1991. Ethnic group was not significantly associated with lag.

Figure 1 shows the observed incidence and the adjusted incidences estimated from models 1–4. The shapes of the four estimates are similar, and the divergence between them is slight. They suggest that incidence increased gradually from 1991 to 1994, increased more rapidly between 1994 and 1996, and stabilized or started to fall slightly after 1997. Model 4 gives the lowest estimates, because it takes account of a shortening lag over time.

Figure 2 shows, for model 4, the observed and adjusted incidences by route of administration; figures 3 and 4 show incidence for males and females, respectively. These figures suggest that the trends for injectors and noninjectors are very different. Among observed cases, noninjecting—mainly "chasing" (inhalation of vapors)—became the most common route of consumption in 1993. Estimates of the numbers of new male and female injectors (who would seek treatment and be reported within 8 years) were stable over time, except for a possible fall among males in 1998. However, estimates for the most recent year must be interpreted with great caution, as the least information is available for this year. In addition, the potential for bias resulting from model misspecification is at its greatest, and this possible bias is not reflected in the confidence intervals. In contrast, noninjecting heroin use was on the increase in the 1990s. There were approximately twofold increases in the estimated numbers of new male and female noninjectors

TABLE 2. Effect of various factors on the time lag (years) between onset of heroin used and first visit to a treatment facility, southeastern England, 1991–1998*

Covariate	No.	%	Unadjusted β	Standard error	χ^2 (df)	*p* value	Adjusted† β	Standard error	χ^2 (df)	*p* value
Sex										
Male‡	6,013	70			22.5 (1)	<0.0001			35.4 (1)	<0.0001
Female	2,601	30	–0.12	0.026			–0.15	0.026		
Age group (years)										
<20‡	976	11			109.9 (2)	<0.0001			124.7 (2)	<0.0001
20–24	3,955	46	–0.22	0.027			–0.24	0.028		
≥25	3,683	43	–0.30	0.031			–0.32	0.031		
Ethnicity										
White‡,§	7,749	90			1.8 (2)	0.4	NS¶		NS	NS
Black	377	4	–0.03	0.058			NS		NS	NS
Asian	488	6	0.06	0.050			NS		NS	NS
Route of administration										
Noninjector‡	3,945	46			3.6 (1)	<0.06			5.4 (1)	<0.02
Injector	4,669	54	0.04	0.024			0.06	0.024		
Year										
1991‡	1,519	18			4.3 (1)	0.04			6.2 (1)	0.01
1991 + 1 year	7,095	82	–0.16	0.008			–0.02	0.008		

* A negative β-coefficient indicates that the covariate is associated with a shorter lag time, while a positive β-coefficient indicates a longer lag time.
† Adjusted for the other covariates in the table.
‡ Reference group.
§ Includes "other."
¶ NS, not significant.

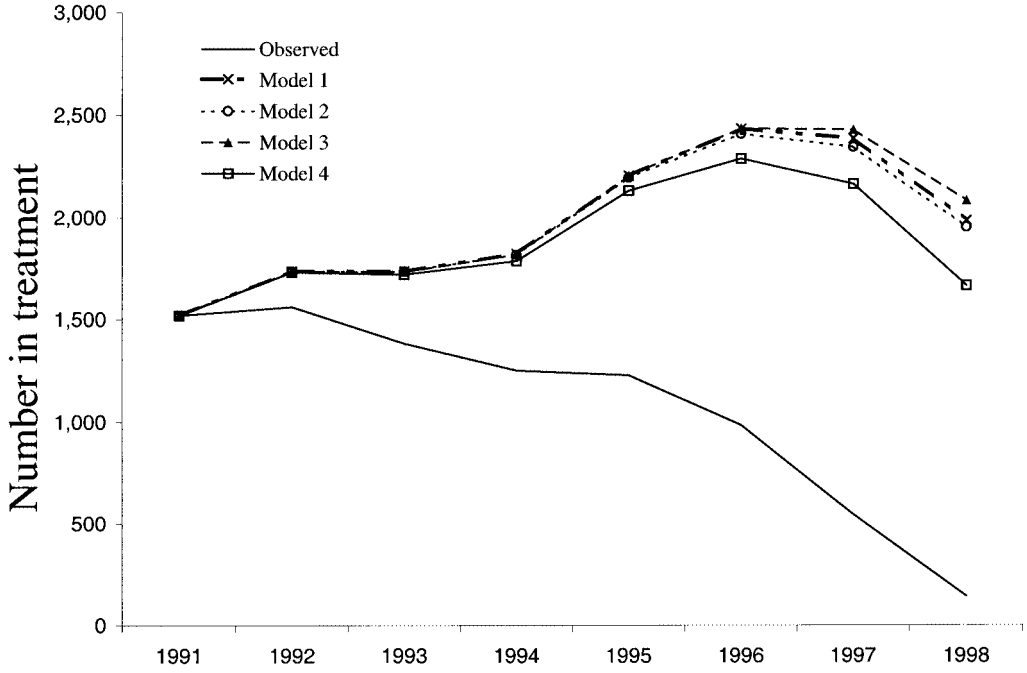

FIGURE 1. Observed (reported) incidence of heroin use and adjusted incidences calculated according to four different models, southeastern England, 1991–1998.

between 1991 and 1996–1997 (figures 3 and 4), after which the incidences appear to have stabilized in males and declined in females (although, again, this apparent decline must be treated very cautiously). Thus, it seems that the incidence of injecting has been stable but there has been a rise in the total incidence of heroin use due to new cases of noninjecting use, an incidence that may now be leveling off.

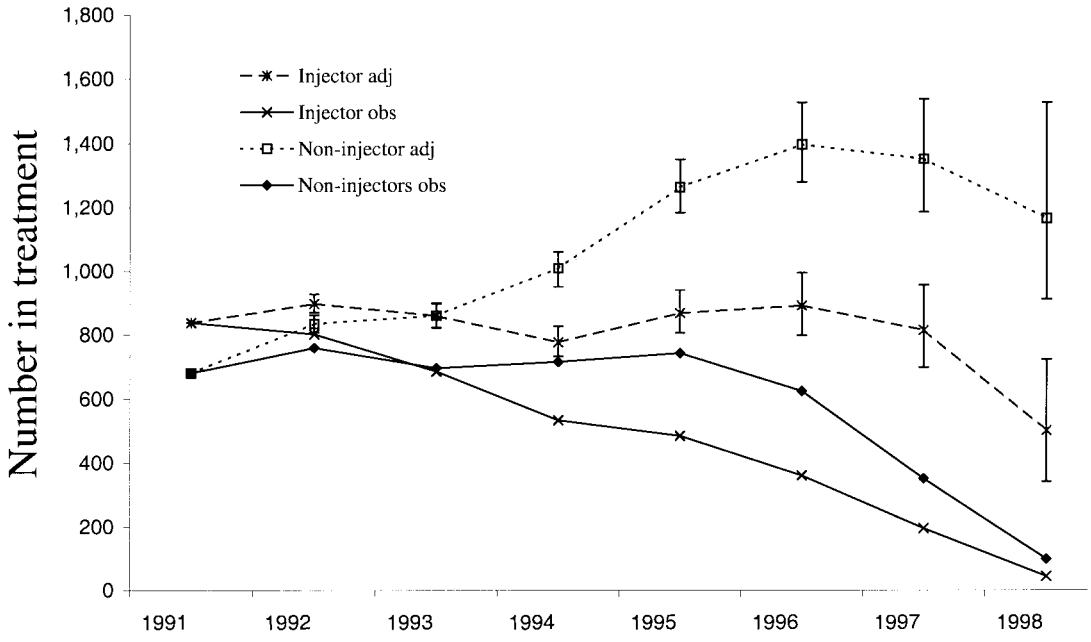

FIGURE 2. Observed and adjusted incidence of heroin use, by route of administration, southeastern England, 1991–1998. Bars, 95% confidence interval.

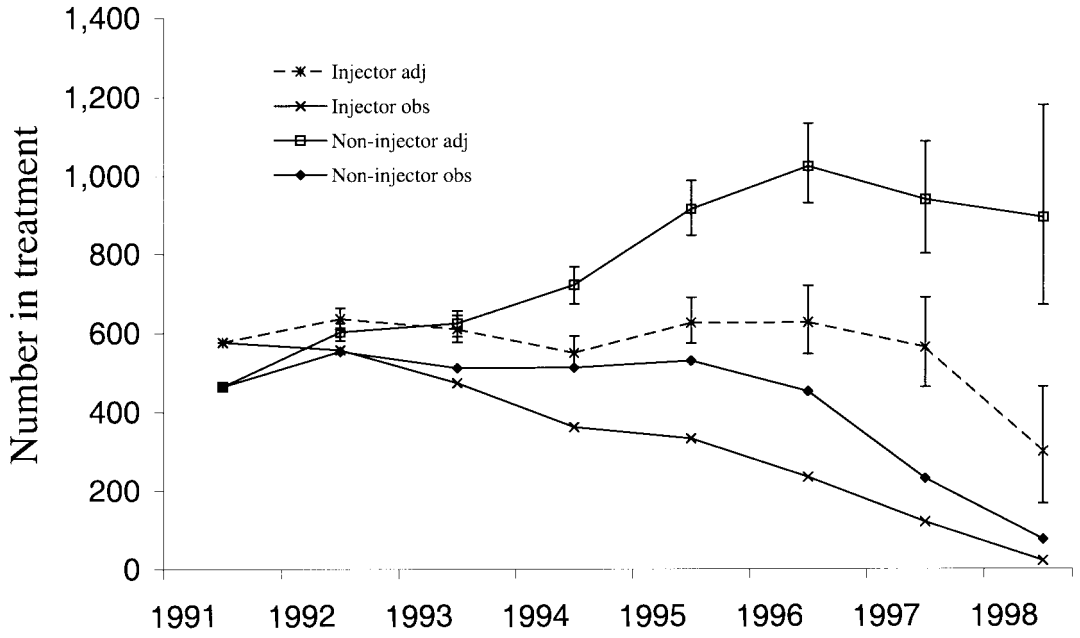

FIGURE 3. Observed and adjusted incidence of heroin use in males, by route of administration, southeastern England, 1991–1998. Bars, 95% confidence interval.

DISCUSSION

In this paper, we have proposed a "lag correction" method which adjusts observed reports of first heroin use in order to estimate the number of heroin users who will request treatment and have their cases reported within a specified time period. This is the strict interpretation that should be applied to these results. However, a wider interpretation of the results and a key purpose of using the method are that the results reveal historical trends in the "relative" incidence of heroin use. That is, they provide evidence of the general shape of the heroin epidemic. The fact that relative rather than absolute incidence is being estimated should not matter, since the trends in the epidemic are more important than its absolute size.

In the United Kingdom, cross-sectional estimates of the proportion of the heroin-using (or drug-using) population in

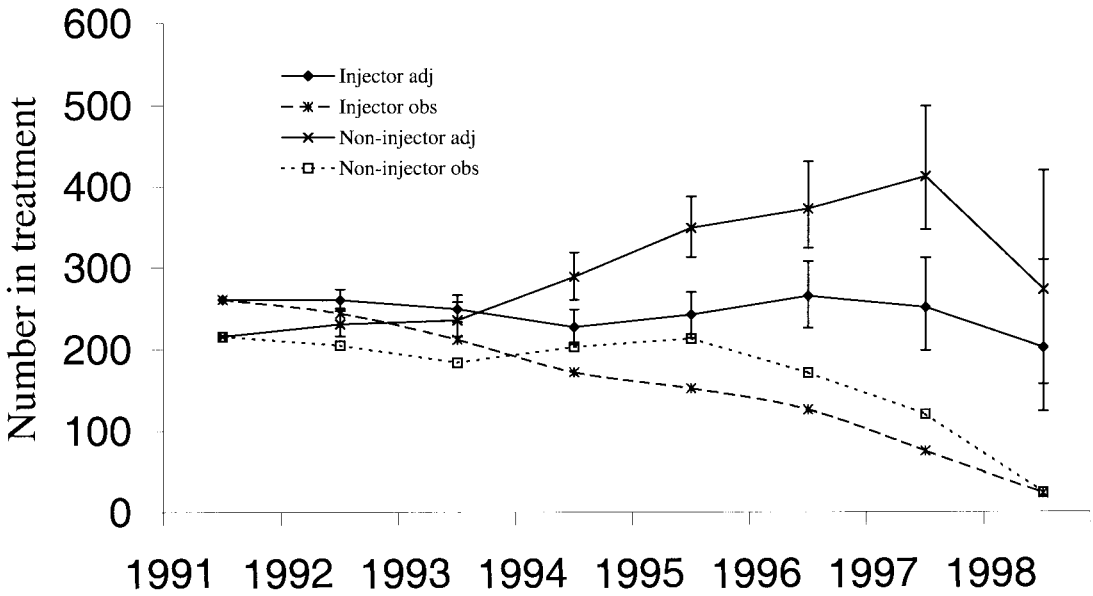

FIGURE 4. Observed and adjusted incidence of heroin use in females, by route of administration, southeastern England, 1991–1998. Bars, 95% confidence interval.

treatment have ranged from less than 15 percent to 33 percent (7). The proportion of heroin users that will ever enter treatment is not known. However, if the proportion that ever enters treatment is high and data are available for an extended period (e.g., more than 15 years), the estimates obtained through our method may approach actual incidence in the population.

Our application advances the work first suggested by Hunt and Chambers (14–16), making use of methodology borrowed from AIDS epidemiology (32–34). We have estimated the lag distribution from observed reports on individuals who started their drug use at the same time as reporting began or after reporting began. We have been able to examine and adjust for the effect of covariates on the lag distribution, and, in particular, we have been able to address changes in this distribution over calendar time. Finally, we have devised a way of providing a measure of the uncertainty associated with our estimates (32). We believe that the lag correction method is currently the best method available for routinely estimating the incidence of heroin use for cities, regions, and countries that conduct surveillance of requests for drug treatment. Alternative solutions based on population surveys or dynamic models have not proven successful (19, 20, 23–25). Equally, results obtained from the lag correction method can support the interpretation of trends in indirect indicators (if available). This is because it is often not clear whether changes (e.g., in heroin seizures, emergency room visits, or positive urine tests among arrestees) are due to changes in sampling, reporting completeness, or local and national policy or changes in the underlying occurrence of heroin use (26, 27). Kaplan (38) has proposed a method for adjusting the bias of "snapshot samples," and as an example he estimated the historical extent of heroin use in New Haven, Connecticut, from a survey of current injectors and their injecting histories. One potential problem with this method is that it relies on knowledge of the distribution of lengths of drug users' injecting careers to adjust the observed data. These quantities are typically difficult to estimate, because of a lack of relevant cohort studies, and they may also vary over time.

In the United Kingdom, there have been suggestions, based on qualitative data, of a new epidemic of heroin chasing (39). Figures 2–4 provide more substantial evidence than ever before that there has been a rise in heroin smoking in the population, while heroin injecting has remained stable. Further work is required to establish whether this rise constitutes a full-fledged "epidemic." It is also too early to tell whether numbers of new cases have reached their peak, since the most recent estimates, which suggest a leveling off, are the estimates with the greatest associated uncertainty. An increase in heroin smoking has clear public health implications for London and the United Kingdom. It would indicate a need to develop services and strategies for preventing both further spread of heroin use and transitions to injecting use, which poses greater health risks to the individual and the population (3, 14, 30). The separate estimates presented here by route of administration (injecting or noninjecting) refer to the route of administration at first visit to a treatment agency. There is good evidence to suggest that a large proportion of noninjectors, if they continue to use heroin, will switch to injecting (40).

The analysis of the lag distribution itself also has interesting public health implications for the UK government's drug policy, given that one of its aims is to encourage heroin users to seek treatment earlier (1). Our findings suggest that the lag is shortening over time. In addition, males, people under age 20, and injectors seem to have longer lags between onset of drug use and first treatment than do females, older age groups, and noninjectors, respectively, although we found no differences by ethnic group. These results offer a first description of subgroups of heroin users that may need to be targeted through outreach programs, and they provide a baseline standard with which to monitor the effect of drug policy.

In interpreting these results, especially in the wider context regarding underlying trends in the population, one must be cautious and must take into account several potential biases. Firstly, changes in the rate of underreporting can seriously affect estimates of the number of heroin users who will request treatment (thus affecting the interpretation of the results even in their strictest sense), as has been shown in the context of AIDS epidemiology (41). We tried to minimize the impact of underreporting by considering only reports made by specialized drug treatment agencies that participated in the Drug Misuse Database surveillance system throughout the study time period. Notwithstanding, we had to make an extra adjustment because of the exceptional increase in underreporting in 1997 (37). A previous validation showed that underreporting to the Drug Misuse Database was approximately 20–25 percent (42), but more recent estimates of underreporting were not available. It is likely that the completeness of reporting underwent slight improvement from 1991 to 1998 (apart from 1997), but it is unlikely that this improvement could solely explain our findings, particularly the difference between trends among heroin injectors and noninjectors.

Secondly, careful consideration must be given to the assumptions on which the lag correction relies for interpreting the results in their widest sense. Violation of these assumptions would mean that the adjusted estimates of the incidence of heroin users' requesting treatment within the truncation time (i.e., 8 years in our study) would still be correct but would not necessarily correspond to trends in the population. The first assumption is that the proportion of heroin users appearing for treatment is unchanged or, at least, does not fluctuate markedly. This might not be the case, for example, if treatment capacity increased and caused an increase in the proportion of heroin users seeking treatment. The second assumption is that the proportion of heroin users with lags longer than the truncation time (i.e., lags longer than 8 years) does not change or does not unduly affect the resulting estimates. The number and size of specialized drug treatment agencies remained stable during the study period, and there was no evidence to suggest that the treated population grew as a result of a greater proportion of heroin users' seeking treatment. This may change in the future, and this issue needs to be addressed in future development work, through sensitivity analysis and perhaps the collection of other relevant data (e.g., data on treatment capacity). Concerning the second assumption, the potential

for bias lies in the possibility that a shift in the lag distribution also moves a large number of heroin users who would have sought treatment after 8 years into the truncation period. The potential impact of this bias should be tested further, although, clearly, as more years of data accumulate, longer lag periods can be used and the effect of this bias can be reduced.

Thirdly, inaccuracy in the data may introduce biases. While dates of birth are generally accurate, "age at first use" may be less reliable. Overstatement or understatement of age at first use could lead to underestimation or inflation of trends in incidence. Most heroin users start their drug use between the ages of 17 and 20 years, but there are a significant number who start drug use earlier or later, and the true distribution of age at first heroin use is not known. Therefore, to reduce the potential bias of false data, we restricted reports to those with ages of first use between 10 and 30 years.

The use of the lag correction method adds substantially to the value of routine treatment data (at least for heroin use), which have not been utilized extensively in epidemiologic research and have been criticized for not providing any useful information (43). It provides an additional reason for collecting these data, apart from accounting purposes, and places a much needed focus in the work of policy-makers and drug agencies on the role of epidemiologic data and the importance of deriving more accurate estimates of trends. Clearly, there is a need for further sensitivity analysis to test some of the assumptions underpinning the strict and wider interpretation of the results as measures of the "relative" incidence of heroin use in the population. These issues include the impact of changes in underreporting and the choice of uplift for 1997, as well as the likely impact of changes in the proportion of heroin users requesting treatment, ever and/or within the reporting period. Future development of the method might take advantage of and direct related studies, particularly to assess current levels of underreporting; to validate the reliability of self-reports of problem drug use, routes of consumption, and age at first use; and to estimate the proportion of heroin users seeking treatment. Estimation of the relative incidence of heroin use should then be repeated on a regular basis as more data accumulate, to monitor the tentative conclusions reached in this paper.

ACKNOWLEDGMENTS

This work was partially funded by the European Monitoring Centre on Drugs and Drug Addiction (EMCDDA). The work was part of the Pilot Project to Estimate Time Trends and Incidence of Problem Drug Use in the European Union (CT.98.EP.07), and the results contribute to the European Union-funded European Network to Develop Policy Relevant Models and Socio-Economic Analyses of Drug Use, Consequences and Interventions (TSER/DGXII project ERB 4141 PL980030). The Centre for Research on Drugs and Health Behaviour acknowledges the financial support of the London Region Office of the National Health Service Executive.

The authors thank Professor Carla Rossi (Tor Vergata University, Rome, Italy), Lucas Weissing (EMCDDA, Lisbon, Portugal), and Professor Gerry Stimson (Imperial College School of Medicine, London, United Kingdom) for their helpful comments and support. They also thank the Health Authorities that fund the Thames Drug Misuse Database, which also partly funded the project.

REFERENCES

1. President of the Council. Tackling drugs to build a better Britain: the government's 10 year strategy for tackling drug misuse. London, United Kingdom: HMSO, 1998.
2. Advisory Council on the Misuse of Drugs. Drug misuse and the environment. London, United Kingdom: HMSO, 1998.
3. Advisory Council on the Misuse of Drugs. Treatment and rehabilitation. London, United Kingdom: HMSO, 1992.
4. Normand J, Vlahov D, Moses LE, eds. Preventing HIV transmission: the role of sterile needles and bleach. Washington, DC: National Academy Press, 1995.
5. Hser Y, Anglin MD, Wickens TD, et al. Techniques for the estimation of illicit drug user prevalence: an overview of relevant issues. New York, NY: National Institute of Justice, 1992.
6. European Monitoring Centre on Drugs and Drug Addiction. Estimating the prevalence of problem drug use in Europe. (EMCDDA scientific monograph series no. 1). Lisbon, Portugal: European Monitoring Centre on Drugs and Drug Addiction, 1997.
7. Hickman M, Cox S, Harvey J, et al. Estimating the prevalence of problem drug use in inner London, United Kingdom: a discussion of three capture-recapture studies. Addiction 1999;94:1653–62.
8. Hennekens CH, Buring JE. Epidemiology in medicine. Boston, MA: Little, Brown and Company, 1987.
9. Kozel HJ, Adams EH. Epidemiology of drug abuse: an overview. Science 1986;234:970–4.
10. deAlarcon R. The spread of heroin abuse in a community. Bull Narc 1969;21:17–22.
11. Hughes PH, Crawford GA. A contagious disease model for researching and intervening in heroin epidemics. Arch Gen Psychiatry 1972;27:149–55.
12. DuPont RL, Greene MH. The dynamics of a heroin addiction epidemic. Science 1973;181:716–22.
13. Levengood R, Lowinger P, Schooff K. Heroin addiction in the suburbs—an epidemiologic study. Am J Public Health 1973;63:209–14.
14. Hunt LD, Chambers CD. Heroin epidemics: a study of heroin use in the United States, 1967–75. New York, NY: Spectrum Publishing, Inc, 1976.
15. Hunt LG. Assessment of local drug abuse. New York, NY: Lexington Books, 1978.
16. Hunt LG. Recent spread of heroin use in the United States. Am J Public Health 1974;64:16–23.
17. Rittenhouse JD, ed. The epidemiology of heroin and other narcotics. (NIDA research monograph no. 16). Rockville, MD: National Institute on Drug Abuse, 1977.
18. Hughes PH, Rieche O. Heroin epidemics revisited. Epidemiol Rev 1995;17:66–73.
19. Gfroerer J, Brodsky M. The incidence of illicit drug use in the United States, 1962–1989. Br J Addict 1992;87:1345–51.
20. Johnson RA, Gerstein DR, Ghadialy R, et al. Trends in the incidence of drug use in the United States, 1919–1992. Rockville, MD: Office of Applied Studies, Substance Abuse and Mental Health Services Administration, 1996.
21. Glower S, Medley G. Modelling drugs and HIV. Br J Addict

1992;87:371–9.
22. Blower SM, Hartel D. HIV, drugs and ecology. Science 1989;246:1236.
23. Simeone RS, Nottingham WT, Holland L. Estimating the size of a heroin-using population: an examination of the use of treatment admissions data. Int J Addict 1993;28:107–28.
24. Knolle H. Incidence and prevalence of illegal drug use in Switzerland in the 1980s and early 1990s: an analytical study. Subst Use Misuse 1997;32:1349–68.
25. Frischer M, Goldberg D, Taylor A, et al. Estimating the incidence and prevalence of injecting drug use in Glasgow. Addict Res 1997;5:307–15.
26. Caukins JP, Ebener PA, McCaffrey DF. Describing DAWN's dominion. Contemp Drug Prob 1995;22:547–67.
27. Harrison LD. Trends and patterns of illicit drug use in the USA: implications for policy. Int J Drug Policy 1995;6:113–27.
28. Stimson GV, Hickman M, Turnbull PJ. Statistics on misuse of drugs have been misused. (Letter). BMJ 1998;317:1388.
29. Roberts CD. Data quality of the drug abuse warning network. Am J Drug Alcohol Abuse 1996;22:389–401.
30. Institute of Medicine, National Academy of Sciences. Pathways of addiction—opportunities in drug abuse research. Washington, DC: National Academy Press, 1996.
31. Brookmeyer R, Gail MH. AIDS epidemiology: a quantitative approach. New York, NY: Oxford University Press, 1994.
32. Brookmeyer R, Liao J. The analysis of delays in disease reporting: methods and results for the acquired immunodeficiency syndrome. Am J Epidemiol 1990;132:355–65.
33. Brookmeyer R, Damiano A. Statistical methods for short-term projections of AIDS incidence. Stat Med 1989;8:23–34.
34. Zeger SL, Lai-Chu S. Statistical methods for monitoring the AIDS epidemic. Stat Med 1989;8:3–21.
35. Donmal MC. The Drug Misuse Database: local monitoring of presenting problem drug use. London, United Kingdom: UK Department of Health, 1990.
36. Office of Population Censuses and Surveys. 1991 census county reports. London, United Kingdom: HMSO, 1993.
37. UK Department of Health. Statistics from the regional Drug Misuse Databases for six months ending September 1998. (Department of Health statistical bulletin 1999/19). London, United Kingdom: Government Statistical Service, 1999.
38. Kaplan EH. Snapshot samples. Socio-Econ Plann Sci 1997; 31:281–91.
39. Parker H, Bury C, Egginton R. New heroin outbreaks amongst young people in England and Wales. London, United Kingdom: Home Office Police Policy Directorate, 1998.
40. Griffiths P, Gossop M, Powis B, et al. Transitions in patterns of heroin administration: a study of heroin chasers and heroin injectors. Addiction 1994;89:301–9.
41. Bacchetti P. Historical assessment of some specific methods for projecting the AIDS epidemic. Am J Epidemiol 1995;141: 776–81.
42. Hickman M, Sutcliffe H, Sondhi A, et al. Validation of a regional drug misuse database: implications for future policy and surveillance of problem drug use in the UK. BMJ 1997; 315:581.
43. Sutton M, Maynard M. Are drug policies based on fake statistics? Addiction 1993;88:455–8.
44. Cox DR, Medley GF. A process of events with notification delay and the forecasting of AIDS. Philos Trans R Soc Lond B Biol Sci 1989;325:135–45.
45. Davison AC, Hinkley DV. Bootstrap methods and their application. Cambridge, United Kingdom: Cambridge University Press, 1997.

APPENDIX

The lag distribution must be analyzed as a function of "year of onset" (i.e., aggregating cases into cohorts according to the year heroin use started) rather than a function of "year of entry" (i.e., constructing cohorts according to year of first report), and only cases reported since reporting began may be included. First, heroin users who started using the drug prior to 1991 ($1991 - x$) will only be reported if they continue their heroin use for at least x years. Thus, users with very short lag times who subsequently cease drug use may not be observed. In addition, of those who began using heroin before 1991 and are observed, it is currently not possible to distinguish whether first report is also first treatment demand. Secondly, estimates of the lag distribution by "entry cohort" and by "onset cohort" will approximate each other if and only if incidence is stable (44). If this is not the case, analysis by entry cohort will bias the results, since a change in incidence cannot be distinguished from a change in the actual lag. For example, if incidence rose (while the real lag distribution remained unchanged), there would be an increase in the number of recently reported cases with short lag times. This would cause the lag distribution estimated by "entry cohort" to shorten (i.e., to have an increased proportion of short lags), and this would, in turn, lead to underestimation of the number of unobserved cases.

Let d denote lag time. For a person whose case is first reported in the same year as onset (i.e., the same year as beginning heroin use), $d = 1$. The maximum observable lag time, denoted m, was 8 years in our study (1991–1998). A person whose onset is in year $1990 + i$ has a truncation time of $m + 1 - i$; that is, he or she will not be observed unless his or her lag time is $\leq m + 1 - i$. Let

$$p_j = P[d = j | d \leq j],$$

and let

$$F(s) = P[d \leq s | d \leq m] = \begin{cases} \prod_{j=s+1}^{m}(1 - p_j) & \text{for } s = 1, \ldots, m - 1 \\ 1 & \text{for } s = m \end{cases}$$

be the conditional lag distribution function (conditional on lag being $\leq m$). The only people who can contribute information about p_j are those whose truncation times are $\geq j$ and whose reporting delays are $\leq j$. We call the set of all such individuals the "risk set" at time j. The number of persons in this risk set is denoted n_j, and the number in this risk set who have a lag time equal to j is Y_j. The nonparametric estimator of F is

$$\hat{F}(s) = \prod_{j=s+1}^{m}\left(1 - \frac{Y_j}{n_j}\right) \quad s = 1, \ldots, m - 1.$$

Let $Z_i(j)$ be the number of persons whose onset is in year i and who have lag j ($1 \leq i \leq m$; $1 \leq j \leq m$). Then $Z_i(j)$ is observed for $1 \leq j \leq m + 1 - i$ and is unobserved for $j \geq m + 2 - i$. Let $Z_i = \sum_{j=1}^{m+1-i} Z_i(j)$ be the crude incidence for year i, i.e., the total number of users whose onset is in year i and whose cases are reported by the end of year $1990 + m$. Brookmeyer and Liao (32) proposed the following estimator

for the adjusted incidence: $\Sigma_{j=1}^{m} Z_i(j)$, i.e., the number whose onset is in year i and who will seek treatment within m years of starting.

$$Z_i^* = \frac{Z_i}{\hat{F}(m + 1 - i)} \quad i = 2, \ldots, m.$$

(Z_1 and Z_1^* are equal and observed.) Brookmeyer and Liao provide formulae for the standard errors of $\hat{F}(s)$ and Z_i^*.

Brookmeyer and Liao also explain how a vector, $\underline{X}$, of covariates that may affect the conditional lag distribution can be included in the model. They suppose that K possible combinations of covariate values are possible, and that for the kth covariate combination, $\underline{X} = \underline{X}_k$, the conditional lag distribution is

$$F_k(s) = F(s)^{\exp(\underline{\beta}^T \underline{X}_k)},$$

where $\underline{\beta}$ is a vector of model parameters. A generalized linear model is used to estimate $F(s)$ and $\underline{\beta}$ and to assess the statistical significance of the latter. Analogously to $Z_i(j)$ and Z_i, let $Z_{ik}(j)$ be the number of persons whose onset is in year i and who have lag j and covariate combination $\underline{X} = \underline{X}_k$, and let $Z_{ik} = \Sigma_{j=1}^{m+1-i} Z_{ik}(j) (1 \leq k \leq K)$. Then Z_{ik} denotes the number of users who have onset in year i, are reported by the end of year m, and have covariate combination $\underline{X} = \underline{X}_k$. The adjusted incidence, estimated taking into account the covariates, is

$$\sum_{k=1}^{K} Z_{ik}^* \quad (i = 2, \ldots, m),$$

where

$$Z_{ik}^* = \frac{Z_{ik}}{\hat{F}_k(m + 1 - i)}. \qquad (1)$$

Unfortunately, the formula for the standard error of Z_{ik}^* can no longer be used, and so the bootstrap method (45) is used instead. The procedure is as follows. Having estimated $\hat{F}$ and $\hat{\underline{\beta}}$ using the generalized linear model and Z_{ik}^* using equation 1, these quantities are taken as fixed. Bootstrap iteration r consists of two steps:

1. Sample, for $i = 1, \ldots, m$ and $k = 1, \ldots, K$,

$$(Z_{ik}^{(r)}(1), \ldots, Z_{ik}^{(r)}(m)) \sim \text{multinomial } (Z_{ik}^*, (\hat{q}_{1k}, \ldots, \hat{q}_{mk})),$$

where

$$\hat{q}_{jk} = \hat{F}_k(j) - \hat{F}_k(j - 1).$$

2. Discard $Z_{ik}^{(r)}(j)$ for $j = m + 2 - i, \ldots, m$. Obtain estimates $\hat{F}^{(r)}$, $\hat{\underline{\beta}}^{(r)}$, and $Z^{*(r)}$ using data $Z_{ik}^{(r)}(j)$ for $j = 1, \ldots, m + 1 - i$. This is done in the same way as $\hat{F}$, $\hat{\underline{\beta}}$, and Z^* were obtained from data $Z_{ik}(j)$ for $j = 1, \ldots, m + 1 - i$.

Standard errors and confidence intervals are obtained from $\{Z_{ik}^{*(1)}, \ldots, Z_{ik}^{*(R)}\}$. We used $R = 1{,}000$ bootstrap samples.