OPEN

Estimating HIV Incidence, Time to Diagnosis, and the Undiagnosed HIV Epidemic Using Routine Surveillance Data

Ard van Sighem,[a] *Fumiyo Nakagawa,*[b] *Daniela De Angelis,*[c] *Chantal Quinten,*[d] *Daniela Bezemer,*[a] *Eline Op de Coul,*[e] *Matthias Egger,*[f] *Frank de Wolf,*[g] *Christophe Fraser,*[g] *and Andrew Phillips*[b]

Background: Estimates of the size of the undiagnosed HIV-infected population are important to understand the HIV epidemic and to plan interventions, including "test-and-treat" strategies.

Methods: We developed a multi-state back-calculation model to estimate HIV incidence, time between infection and diagnosis, and the undiagnosed population by CD4 count strata, using surveillance data on new HIV and AIDS diagnoses. The HIV incidence curve was modelled using cubic splines. The model was tested on simulated data and applied to surveillance data on men who have sex with men in The Netherlands.

Results: The number of HIV infections could be estimated accurately using simulated data, with most values within the 95% confidence intervals of model predictions. When applying the model to Dutch surveillance data, 15,400 (95% confidence interval [CI] = 15,000, 16,000) men who have sex with men were estimated to have been infected between 1980 and 2011. HIV incidence showed a bimodal distribution, with peaks around 1985 and 2005 and a decline in recent years. Mean time to diagnosis was 6.1 (95% CI = 5.8, 6.4) years between 1984 and 1995 and decreased to 2.6 (2.3, 3.0) years in 2011. By the end of 2011, 11,500 (11,000, 12,000) men who have sex with men in The Netherlands were estimated to be living with HIV, of whom 1,750 (1,450, 2,200) were still undiagnosed. Of the undiagnosed men who have sex with men, 29% (22, 37) were infected for less than 1 year, and 16% (13, 20) for more than 5 years.

Conclusions: This multi-state back-calculation model will be useful to estimate HIV incidence, time to diagnosis, and the undiagnosed HIV epidemic based on routine surveillance data.

(*Epidemiology* 2015;26: 653–660)

Over the past few decades, HIV has changed from an infection invariably culminating in AIDS and death to a chronic condition needing early and lifelong treatment to prevent replication of the virus and to maintain normal CD4+ T cell counts (CD4 counts). Nevertheless, as the number of individuals diagnosed with HIV as well as the number treated for their infection continues to increase in Europe, HIV infection remains a major public health problem. The exact number of people infected with the virus remains, however, unknown because many HIV-infected individuals have not yet been diagnosed. According to recent UNAIDS estimates, approximately 900,000 people are living with HIV in Western and Central Europe.[1] In European countries, recent estimates of the proportion of infected individuals who are undiagnosed ranged from 20% to 40%.[2–5]

Accurate estimates of the number of people living with HIV, including those not yet diagnosed, are of paramount importance for understanding the burden of HIV and the need for combination antiretroviral treatment.[6] Equally relevant are estimates of current trends in HIV incidence and uptake of testing, for instance when evaluating the effect on the HIV epidemic of "test-and-treat" strategies, where HIV-positive individuals are offered combination antiretroviral treatment immediately, independently of CD4 counts.[7] If the incidence of new HIV infections over time is known for the entire population, the number of individuals living with HIV may be estimated using additional data on migration and mortality.[8,9]

As many countries in Europe have implemented surveillance systems for HIV infection and AIDS, data are available on newly diagnosed HIV infections as well as on AIDS cases.[10] However, since the duration between HIV infection and HIV diagnosis is often unknown, they cannot be used as a direct measure of HIV incidence.[3,11] A growing number of countries

Submitted 23 September 2014; accepted 3 May 2015.

Editors' note: a commentary on this article appears on page 625.

[a]Stichting HIV Monitoring, Amsterdam, The Netherlands; [b]University College London, London, United Kingdom; [c]MRC Biostatistics Unit, Cambridge, United Kingdom; [d]European Centre for Disease Prevention and Control, Stockholm, Sweden; [e]National Institute for Public Health and the Environment, Bilthoven, The Netherlands; [f]Institute of Social and Preventive Medicine, University of Bern, Bern, Switzerland; and [g]Department of Infectious Disease Epidemiology, School of Public Health, Imperial College London, London, United Kingdom.

This study was funded as part of European Centre for Disease Prevention and Control Framework Contract No. ECDC/2012/050, coordinated by Chantal Quinten and supported by Anastasia Pharris. The opinions expressed are those of the authors only and do not represent ECDC's official position. The ATHENA database is supported by a grant from the Dutch Ministry of Health, Welfare and Sport and was set up and is maintained by Stichting HIV Monitoring.

The authors report no conflicts of interest.

SDC Supplemental digital content is available through direct URL citations in the HTML and PDF versions of this article (www.epidem.com). This content is not peer-reviewed or copy-edited; it is the sole responsibility of the authors.

Correspondence: Ard van Sighem, Stichting HIV Monitoring, Meibergdreef 9, 1105 AZ, Amsterdam, The Netherlands. Email:a.i.vansighem@amc.uva.nl.

ISSN: 1044-3983/15/2605-0653
DOI: 10.1097/EDE.0000000000000324

also report on CD4 counts at the time of HIV diagnosis.[10] CD4 counts not only provide information on the number of HIV-infected individuals in need of treatment but also on the duration of infection. Mathematical tools such as back-calculation methods that use data on new diagnoses and CD4 counts around diagnosis can provide some insight into the dynamics underlying these observed data by disentangling changes in HIV incidence and changes in testing patterns.[3,12–15]

We developed a method to simultaneously estimate HIV incidence, time between infection and diagnosis, and the size of the undiagnosed population, using surveillance data on reported HIV and AIDS cases and information on CD4 counts at the time of diagnosis.[13] We tested the approach on simulated data and applied it to data on men who have sex with men (MSM) in The Netherlands.

MATERIALS AND METHODS

Mathematical Model

The method used to estimate HIV incidence and time to diagnosis is an extension of the model used by Sweeting and colleagues[3] (simplified model Figure 1; full model eFigure 1; http://links.lww.com/EDE/A932).[13] In brief, the model describes HIV progression as a unidirectional flow through different stages of the infection that are characterized by CD4 counts or the presence of AIDS events.[16] Immediately after infection, all individuals first enter a phase of primary HIV infection and then, in the absence of antiretroviral treatment, progress to AIDS through up to four different CD4 strata. The proportion of patients in each CD4 cell stratum immediately after primary infection and the progression rates between CD4 strata are based on data from seroconverters in the CASCADE collaboration.[17,18] A complete description of the model and its parameters is given in the eAppendix (http://links.lww.com/EDE/A932).

During each stage of infection, patients can be diagnosed at a rate that may depend on calendar time. For simplicity, we assume that HIV-infected individuals cannot be diagnosed during primary infection when antibody responses to HIV have not fully developed yet. We considered five distinct historical periods for which CD4 stratum-specific diagnosis rates are estimated: (1) 1980–1983, during which the first AIDS cases were diagnosed; (2) 1984–1995, when serological testing for HIV became widely available; (3) 1996–1999, the start of the era of combination antiretroviral treatment; (4) 2000–2004; and (5) 2005–2012.[5,19] Diagnosis rates were approximated as a piecewise linear function of calendar time with a different slope for each of the five time intervals. Thirty different parameters were thus necessary to describe diagnosis rates, six for each stage of infection. To reduce the number of parameters, we assumed that in the first time interval 1980–1983 all diagnosis rates are zero except for d_5 because no diagnostic tests were available at that time and HIV could only be diagnosed when AIDS had developed (Figure 1). Furthermore, d_5 was fixed at a high and constant value over calendar time, reflecting the high probability of being diagnosed with HIV when AIDS symptoms appear. In addition, we assumed that diagnosis rates in the second time interval 1984–1995 were also constant over time (see also eAppendix; http://links.lww.com/EDE/A932) such that the total number of diagnosis rate parameters that needed to be estimated from the data is 16. This assumption

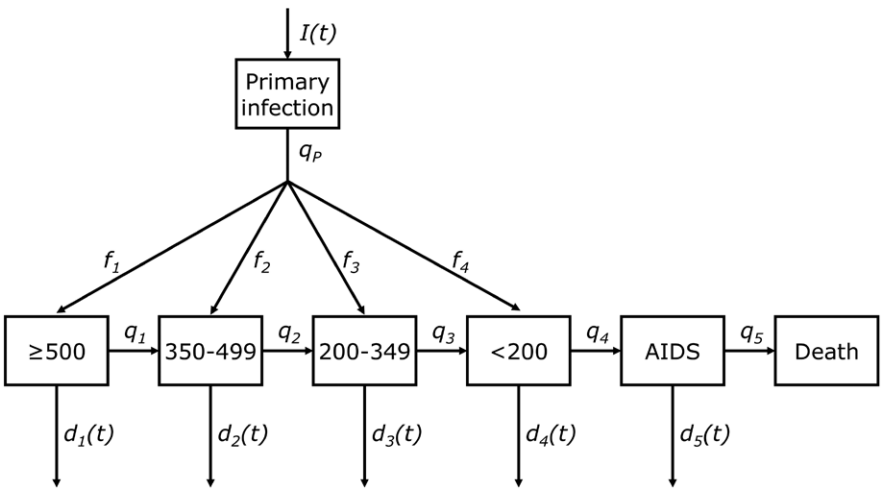

FIGURE 1. Simplified model structure. HIV incidence over calendar time *t* is denoted by $I(t)$. Immediately after infection, all individuals first enter a phase of primary infection. After primary infection, individuals enter at a rate $f_i q_P$ one of four AIDS-free CD4 compartments of undiagnosed HIV infection with $\sum_{i=1}^{4} f_i = 1$. We assume that no HIV-infected individuals immediately progress to AIDS after primary infection. In the absence of treatment, individuals progress to the next compartment at a rate $q_i\,(i = 1,\ldots,5)$ until they develop AIDS and then die because of AIDS at a rate q_5. During each stage except primary infection individuals can be diagnosed at a rate $d_i(t)$, depending on the stage and on calendar time.

was motivated by the low number of observed HIV diagnoses for the early years of the epidemic due to data truncation (see "Fitting to Surveillance Data") and by the fact that in The Netherlands HIV testing rates have traditionally been among the lowest in Europe before the introduction of combination antiretroviral treatment.[20,21]

The HIV incidence curve was approximated using cubic M-splines, which allows for high flexibility with relatively few parameters.[22] It was assumed that the incidence rate started at zero in 1980. Further details are given in the Supplementary Material (http://links.lww.com/EDE/A932).

Fitting to Surveillance Data

We fitted the above model to different sources of surveillance data from the AIDS Therapy Evaluation in The Netherlands (ATHENA) national observational HIV cohort and the National Institute of Public Health and the Environment in The Netherlands.[23,24] The ATHENA cohort includes anonymized data from all HIV-infected patients living in The Netherlands who receive care in one of the 27 HIV treatment centers. ATHENA patients are informed of data collection by their treating physician and patients can refuse further collection of clinical data according to an opt-out procedure. Written informed consent and ethical approval is not obtained, as data collection is part of HIV care. Surveillance data used in our study included the annual number of new HIV diagnoses by CD4 count stratum, the annual number of new AIDS cases, and the annual number of concurrent HIV and AIDS diagnoses (an AIDS diagnosis within 6 weeks of HIV diagnosis). The fitting procedure resulted in estimation of the diagnosis rates and HIV incidence curve (see eAppendix for further details; http://links.lww.com/EDE/A932).

Annual numbers of HIV diagnoses were available from 1984 onwards. The CD4 count at the time of diagnosis was defined as the first CD4 count within 3 months after diagnosis and before start of treatment. Overall, 75% of the patients diagnosed from 1984 onwards had a CD4 count available at the time of diagnosis. We took missing CD4 counts into account by multiplying for each year the estimated number of HIV diagnoses in each stratum by the proportion of all HIV diagnoses in the same year having a CD4 count. Thus, we implicitly assumed that patients without CD4 count measurement had the same CD4 distribution as those with a CD4 measurement (see eAppendix; http://links.lww.com/EDE/A932). The estimated annual numbers of diagnoses by CD4 count stratum were then fitted to the observed numbers. As the ATHENA cohort only started in 1998 with retrospective inclusion of patients from 1996 onwards, patients who died befovre 1996 were not included in the database. We explicitly accounted for this data truncation by calculating the probability that a patient diagnosed in each CD4 stratum survived up to 1996 without antiretroviral treatment.[19] We assumed that mono- or dual therapy, which were available before 1996 but not widely used, did not substantially alter these survival probabilities.

The model was also fitted to data from the Dutch Health Inspectorate on the annual total number of AIDS cases between 1982 and 1996.[16,24] Total numbers of AIDS diagnoses after 1996 were not used because the probability of progressing to AIDS would be affected by the use of combination antiretroviral treatment and incorporating its effect and changes in treatment guidelines on when to start antiretroviral treatment would make our model much more complicated. Instead, from 1996 onwards, we used ATHENA data on concurrent HIV/AIDS diagnoses made before treatment was started and therefore not affected by treatment. We did not use data on concurrent HIV and AIDS diagnoses before 1996, because patients in ATHENA with an HIV/AIDS diagnosis in that period are a biased group of patients who survived up to 1996. The number of HIV diagnoses and HIV/AIDS cases in 2011 and 2012 were corrected for a backlog in registration by adding 3% and 11%, respectively, to the number of cases observed so far.[23]

We constructed point wise 95% confidence intervals (CI) for the parameters and other outcomes via a bootstrap procedure (see eAppendix; http://links.lww.com/EDE/A932). Although analyses were done up to and including 2012, we report model outcomes at the end of 2011 because of wide confidence intervals in 2012. Numbers were rounded to the nearest 10, 50, or 100 if they were below 1,000, between 1,000 and 10,000, or above 10,000, respectively.

The number of HIV-infected MSM who were still alive by the end of 2011 was estimated by subtracting the cumulative number of MSM who died from the cumulative number of HIV infections. The number of MSM who died because of AIDS, while remaining undiagnosed was calculated from the mathematical model. The number of deaths among diagnosed MSM was taken from Statistics Netherlands (before 2002) or ATHENA (from 2002 onwards).[23,24] Since Statistics Netherlands has no information on transmission risk group, it was assumed that 70% of deaths among men were MSM, which is the percentage of MSM among male patients in ATHENA diagnosed before 2002.[5]

In a secondary analysis, we assumed diagnosis rates to be the same for the first three CD4 strata ($d_1(t) = d_2(t) = d_3(t)$) and assumed that $d_4(t) = d_3(t) + d_{\text{symp}}(t)$ with $d_{\text{symp}}(t)$ the rate of being diagnosed because of HIV-related symptoms. In this case, we did not use information on CD4 counts and fitted the model to the total annual number of HIV diagnoses instead of diagnoses by CD4 count stratum. We also did a multivariable sensitivity analysis to investigate the impact of assumptions on input parameters on the model outcomes (eAppendix; http://links.lww.com/EDE/A932).[25,26]

Simulated Data

We tested our approach by determining the annual number of HIV infections and the number of undiagnosed infections in three populations of simulated HIV-infected patients generated using HIV Synthesis.[8,27] In brief, HIV Synthesis is an individual-based stochastic simulation model of HIV progression and the effect of antiretroviral treatment. Model assumptions and

parameters are based on data from observational cohorts and clinical trials. Although HIV Synthesis can simulate the effect of antiretroviral treatment, for our purpose, the most important feature is its ability to generate newly diagnosed HIV and AIDS cases from prespecified incidence curves and diagnosis rates with a validated course of CD4 count changes and progression to AIDS and death during untreated HIV infection.[27] Estimated incidence curves were compared with the true curves that were used as input in the simulations.

RESULTS

Applying our method to simulated data showed that the true number of infections could be reconstructed accurately, with the majority of true values lying within the estimated 95% confidence intervals (Figure 2). However, the method could not accurately estimate the peak in the true number of infections in the mid-1980s, which is most apparent in Figure 2C. Nevertheless, the method was still able to estimate the number of undiagnosed individuals although the difference with the true number was larger in most recent years (Figure 3). In the secondary model, which did not use information on CD4 counts at the time of diagnosis, the estimated HIV infection curves and undiagnosed population looked similar (eFigures 2 and 3; http://links.lww.com/EDE/A932), although confidence intervals were wider.

Between 1980 and 2011, a cumulative number of 15,400 (95% CI = 15,000, 16,000) infections in MSM in The Netherlands were estimated to have occurred. The number of HIV infections peaked around 1985 with 870 (820, 940) new infections in that year (Figure 4A). In the 1990s, the number of infections varied between 200 and 300 per year. Since 2000, HIV incidence steadily increased and reached levels comparable with those in the mid-1980s by 2005, followed by a decreasing trend to 590 (400, 780) new infections in 2011.

The mean time between infection and diagnosis if diagnosis rates would remain the same as in the year of infection is shown in Figure 4B. Between 1980 and 1983, when HIV could only be diagnosed once AIDS symptoms appeared, the average time between infection and HIV diagnosis for people infected in this period, were conditions to have remained as they were in this period, would have been 11.6 years. This decreased to 6.1 years (95% CI = 5.8, 6.4) in the period 1984–1995. From 1996 onwards, the time to HIV diagnosis steadily decreased to 2.6 (2.3, 3.0) years on average for men infected in 2011. In contrast, the actual time between infection and diagnosis by year of diagnosis (Figure 4C) was estimated to increase from the start of the epidemic to between 7 and 8 years at the end of the 1990s, and then decreased to 3.6 (3.3, 4.0) years on average in 2011.

In total, 3150 diagnosed and 100 undiagnosed MSM had died by the end of 2011 such that 12,200 (95% CI = 11,700, 12,800) HIV-infected MSM were estimated to be alive at that time (Figure 4D). At the same time, according to ATHENA data, 6% of MSM diagnosed between 1996 and 2011 were not in care anymore because they had moved abroad or were lost to follow-up. We therefore estimated the total number of MSM living with HIV in The Netherlands, including those not yet diagnosed, to be 11,500 (11,000, 12,000).

Figure 4D shows that the total number of undiagnosed MSM has remained around 2000 for the past 15 years. Altogether, 1,750 (95% CI = 1,450, 2,200) HIV-infected MSM, or 11% (10, 14) of all those infected since the start of the HIV epidemic, were estimated to be still undiagnosed by the end of 2011 (Figure 4D). Of these undiagnosed MSM, 29% (95% CI = 22, 37) had been infected for less than 1 year, 54% (47, 60) were infected 1 to 5 years before, and 16% (13, 20) had been infected for more than 5 years. In total, 800 (650, 1,000) undiagnosed MSM, or 46% (42, 50), had CD4 cell counts below 500 cells/mm^3, whereas 28% (24, 31) had CD4 counts below 350 cells/mm^3.

The multivariable sensitivity analysis showed that parameters associated with the earliest stages of infection had the largest impact on estimated model outcomes (eTable 2;

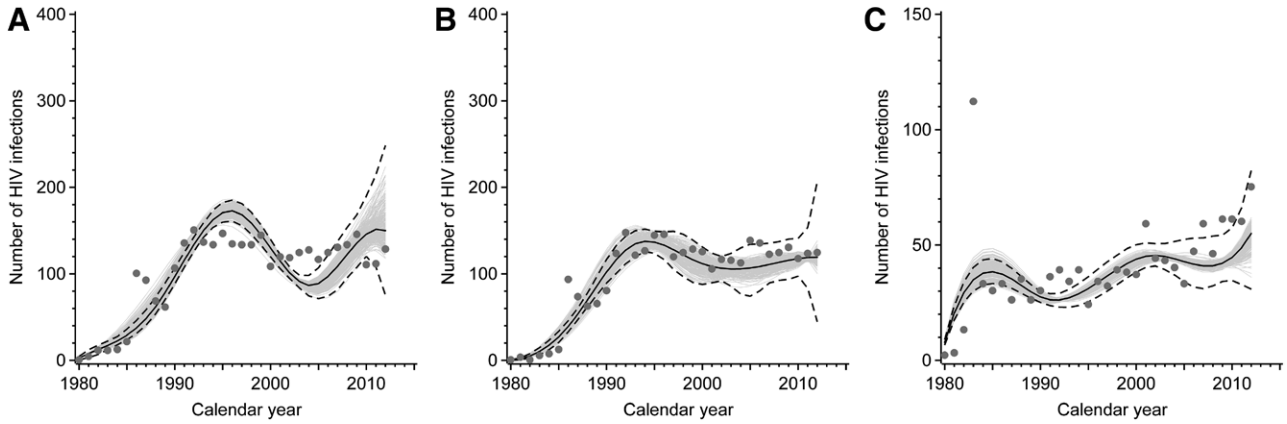

FIGURE 2. Estimated and true number of infections for three different simulated HIV epidemics. *Black solid lines* show the model estimates, and *dashed lines* are 95% confidence intervals. *Thin grey lines* show results of multivariable sensitivity analyses. *Grey dots* are the true annual number of infections.

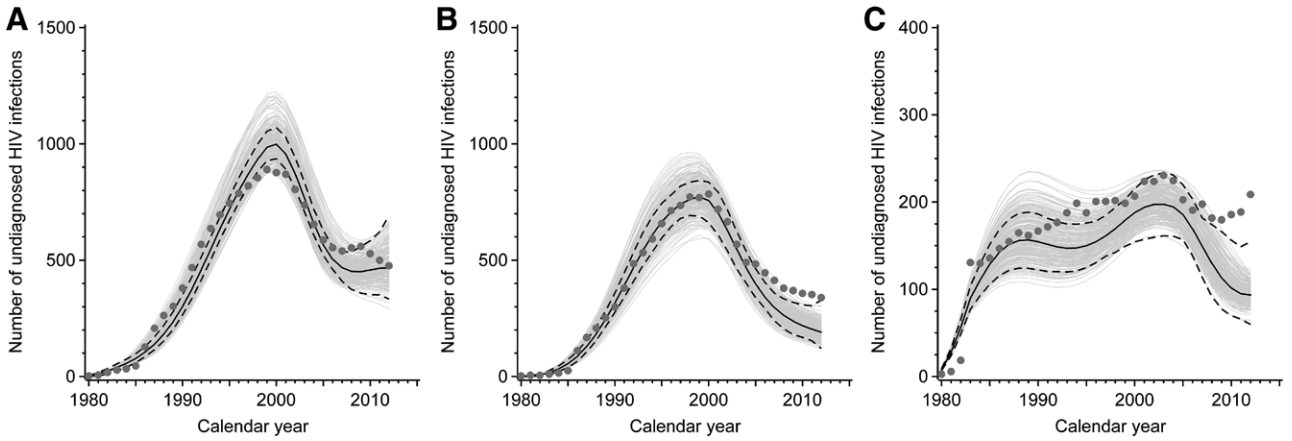

FIGURE 3. Estimated and true number of undiagnosed infections for three different simulated HIV epidemics. *Black solid lines* show the model estimates, and *dashed lines* are 95% confidence intervals. *Thin grey lines* show results of multivariable sensitivity analyses. *Grey dots* are the true annual number of undiagnosed infections.

http://links.lww.com/EDE/A932). The largest effect was seen in model outcomes before 1996, where there are less data to constrain the model fits (Figure 4).

DISCUSSION

Our study shows that by the end of 2011, 11,500 MSM were living with HIV in The Netherlands, of whom 1,750 were still undiagnosed. The annual number of infections around 2005 was estimated to be at similar levels as in the early phase of the HIV epidemic with a declining trend in more recent years. Meanwhile, testing rates have increased such that the average time from infection to diagnosis has decreased to 2.6 years for those infected in recent years.

One of the strengths of our back-calculation method is that it only uses routinely collected data on HIV and AIDS diagnoses. Our approach may therefore easily be applied in other settings with similar data availability. Furthermore, using information on the stage of the infection at the time of diagnosis, either measured by CD4 counts or by a concurrent AIDS diagnosis, our method is able to distinguish changes in HIV incidence from changes in the probability of being diagnosed.[3,5,6,13–15] As our method integrates HIV incidence, natural disease progression, and HIV diagnosis, estimates of the number of undiagnosed infections by disease stage and CD4 stratum are automatically generated. The multivariable sensitivity analysis and results from simulated data showed that generally our findings were robust to changes in input parameters, although less so in calendar years for which there were less data to constrain the model fits. Also, estimates of HIV incidence and the undiagnosed proportion were similar when only using total annual numbers of HIV and AIDS diagnoses and information on simultaneous HIV/AIDS diagnosis. Our approach may therefore still be applicable in settings with no or limited (historical) data on CD4 counts.

There are, however, also a number of limitations to our study. First, our analysis relies on historical data on AIDS cases and HIV diagnoses, which may not be available in all countries or for all risk groups. Second, the method uses data on case reports, which may be subject to underreporting, misclassification of transmission risk group or stage of infection, or incomplete information. Underreporting of HIV cases to the national surveillance system, which may be quite likely in some settings, will lead to underestimates of HIV incidence and may also affect estimates of diagnosis rates, for instance, when HIV/AIDS diagnoses are less likely to be reported. In The Netherlands, case reports for MSM diagnosed with HIV before 1996 were incomplete, because there is no national reporting of HIV diagnoses since the start of the epidemic and ATHENA only started collecting data from 1996 onwards. Also, CD4 counts at the time of diagnosis were not available for all patients, although in more recent calendar years they were missing in less than 15% of newly diagnosed patients. In particular, when CD4 counts are not missing at random, for instance when symptomatic HIV diagnoses are less likely to have a CD4 count than nonsymptomatic diagnoses, the time to diagnosis can be underestimated. Furthermore, the mathematical model only takes into account AIDS-related deaths such that HIV incidence may be underestimated when there is substantial non-AIDS-related mortality in undiagnosed individuals. Emigration of undiagnosed HIV-infected individuals would also yield lower estimates of the annual number of new HIV infections. However, we expect only a limited effect of prediagnosis mortality and emigration on the estimated number living with HIV, because both cumulative incidence and cumulative number of reported deaths will be affected in the same way. Our approach does not explicitly take into account migration and may therefore be less suitable for migrant populations, many of whom will have been infected in the country of origin but cannot be diagnosed in The Netherlands before arriving.

The rates of CD4 cell decline and onset of AIDS in our analysis were based on data from an international cohort of HIV-infected individuals with a well-estimated time of

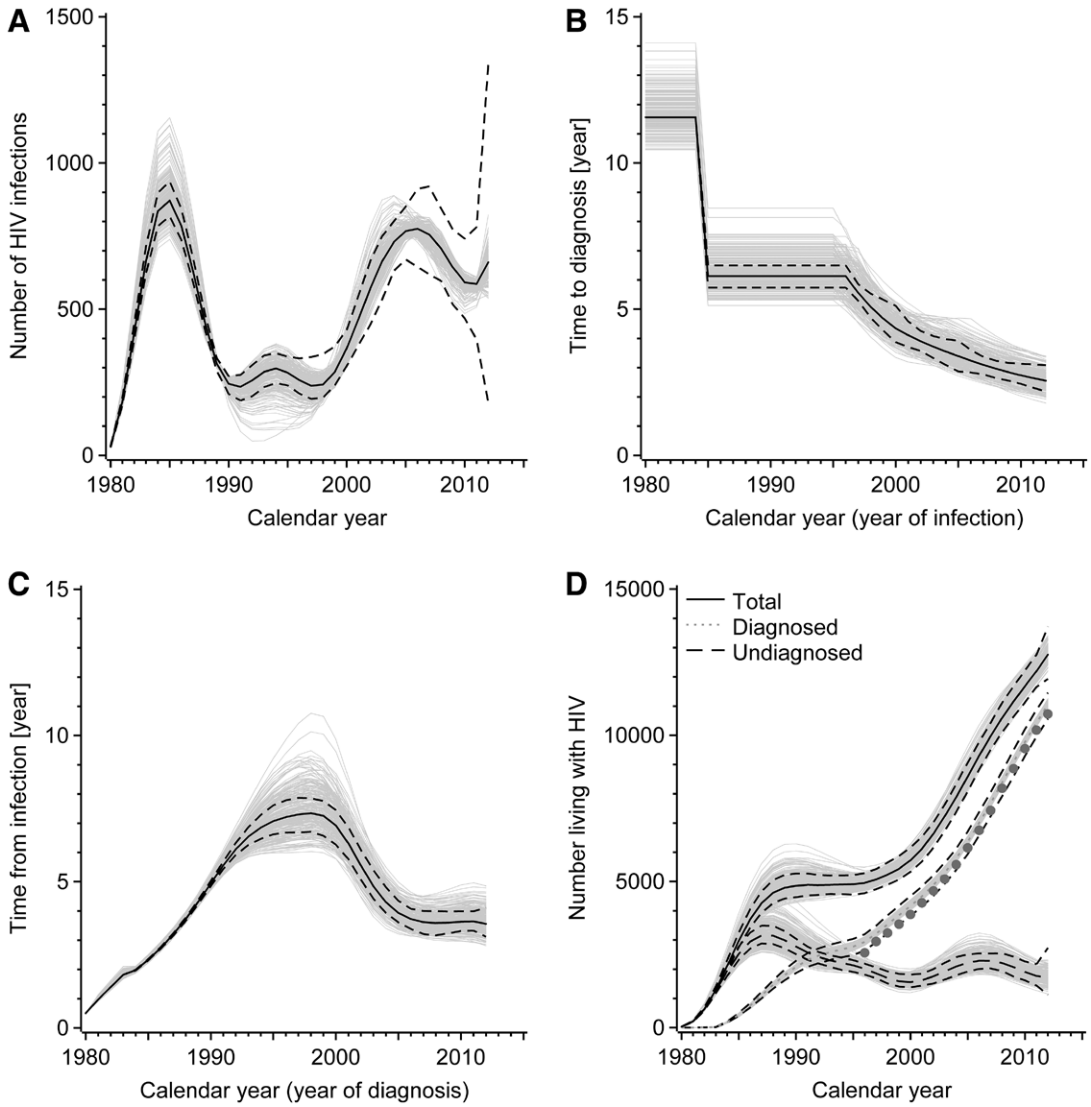

FIGURE 4. Model outcomes for men who have sex with men (MSM) in The Netherlands. A, Annual number of new HIV infections; (B) average time from HIV infection to diagnosis by year of infection if diagnosis rates would remain the same as in the year of infection; (C) average time from HIV infection to diagnosis by year of diagnosis; (D) total number of individuals living with HIV and number of diagnosed and undiagnosed HIV infections, with dots representing the number of diagnosed MSM living with HIV according to the ATHENA database. *Thin grey lines* show results of multivariable sensitivity analyses.

infection.[17] Whereas CD4 cell decline in these seroconverters was similar to that in seroprevalent HIV patients, our method ignored differences in progression rates by age, ethnicity, viral load, and, possibly, increasing virulence in recent calendar years.[28–32] Patients may also have received prophylaxis for opportunistic infections thus prolonging the time to onset of AIDS. We may therefore overestimate the time to AIDS in untreated patients. This is however unlikely to invalidate our results as in recent years the majority of HIV patients are diagnosed before onset of AIDS. The average time to onset of AIDS in the absence of treatment was also longer than in a previous analysis in The Netherlands, which may explain, at least in part, why we also estimated a longer time to diagnosis.[5,19]

Previously, using the Multi-Parameter Evidence Synthesis approach, the number of MSM in The Netherlands living with HIV was estimated to be 11,758 (95% credibility interval = 8,670–17,573) at the beginning of 2008.[2,33] Our method estimates a lower number of 9,950 (95% CI = 9,550, 10,400) infected MSM. We think that part of this difference is explained by MSM who opt out of further data collection in ATHENA, approximately 1% of the male patients registered in the cohort. As data on transmission risk group were unavailable for these men, they were not part of our analysis and this may have led to a downward bias in the estimated number of infections. Also, about 5% of HIV-positive MSM participating in a sexual health survey were not yet in care in an HIV treatment centre,

and would therefore not be in the ATHENA database, although this mainly concerned MSM who were recently diagnosed.[34]

This study confirms previous findings that in The Netherlands, a country with free HIV testing for high-risk groups like MSM and universal access to care, a substantial proportion of infected MSM are still unaware of their infection.[2,5,19] We found that 16% of undiagnosed infections have been living with HIV for more than 5 years, a proportion similar as found in a recent study in France.[35] Approximately, half of the undiagnosed patients had CD4 counts below 500 cells/mm^3 and would thus be eligible for treatment, whereas more than a quarter of the undiagnosed infections were already in immediate need of treatment (CD4 < 350 cells/mm^3).[36] Most likely, these MSM did not yet suffer from clinical symptoms that were severe or specific enough to get tested for HIV. Diagnosis and start of combination antiretroviral treatment at lower CD4 cell counts have been shown to be associated with poorer immunological recovery and increased rates of AIDS or death.[37,38]

The most likely reason for the increase in the number of HIV infections is continuing and increasing risky sexual behavior, in particular by those unaware that they are infected with HIV.[5,19,39,40] Although MSM who have been diagnosed with HIV decrease their sexual risk behavior, this decrease has been shown to be of a transient nature, also in MSM not yet treated with combination antiretroviral treatment.[41,42] Even though 79% of the diagnosed MSM currently in care have a suppressed viral load, and more than a third of MSM diagnosed in recent years were infected at most 1.5 years before, this has apparently not been enough to lead to a substantial decline in HIV incidence in this group.[23] To fully curb the HIV epidemic in MSM, efforts must therefore continue to increase awareness of HIV and promote testing such that the beneficial effect of cART, both for the individual patient as well as for preventing transmission, can be fully exploited.

In conclusion, we have shown that our method can be used to simultaneously estimate HIV incidence, the time between infection and diagnosis, and the size and characteristics of the undiagnosed population, using only routine surveillance data. Our method can easily be applied by other European countries and would be a useful tool to gauge how successful interventions like "test-and-treat" strategies are in reaching all or almost all undiagnosed HIV infections.

ACKNOWLEDGMENTS

We thank Dr. Andrew Amato-Gauci of ECDC for his input into the concept development and constructive review comments.

REFERENCES

1. UNAIDS. Global report: UNAIDS report on the global AIDS epidemic 2013. UNAIDS/JC2502/1/E. Joint United Nations Programme on HIV/AIDS (UNAIDS), 2013.
2. van Veen MG, Presanis AM, Conti S, et al. National estimate of HIV prevalence in the Netherlands: comparison and applicability of different estimation tools. *AIDS*. 2011;25:229–237.
3. Birrell PJ, Gill ON, Delpech VC, et al. HIV incidence in men who have sex with men in England and Wales 2001-10: a nationwide population study. *Lancet Infect Dis*. 2013;13:313–318.
4. Hamers FF, Phillips AN. Diagnosed and undiagnosed HIV-infected populations in Europe. *HIV Med*. 2008;9(Suppl 2):6–12.
5. Bezemer D, de Wolf F, Boerlijst MC, et al. A resurgent HIV-1 epidemic among men who have sex with men in the era of potent antiretroviral therapy. *AIDS*. 2008;22:1071–1077.
6. Working Group on Estimation of HIV Prevalence in Europe. HIV in hiding: methods and data requirements for the estimation of the number of people living with undiagnosed HIV. *AIDS*. 2011;25:1017–1023.
7. Kretzschmar ME, Schim van der Loeff MF, Birrell PJ, De Angelis D, Coutinho RA. Prospects of elimination of HIV with test-and-treat strategy. *Proc Natl Acad Sci USA*. 2013;110:15538–15543.
8. Phillips AN, Cambiano V, Nakagawa F, et al. Increased HIV incidence in men who have sex with men despite high levels of ART-induced viral suppression: analysis of an extensively documented epidemic. *PLoS One*. 2013;8:e55312.
9. Stover J, Brown T, Marston M. Updates to the Spectrum/Estimation and Projection Package (EPP) model to estimate HIV trends for adults and children. *Sex Transm Infect*. 2012;88(Suppl 2):i11–i16.
10. European Centre for Disease Prevention and Control/WHO Regional Office for Europe. HIV/AIDS surveillance in Europe 2012. Stockholm, European Centre for Disease Prevention and Control, 2013.
11. Smith MK, Powers KA, Muessig KE, Miller WC, Cohen MS. HIV treatment as prevention: the utility and limitations of ecological observation. *PLoS Med*. 2012;9:e1001260.
12. Campsmith ML, Rhodes PH, Hall HI, Green TA. Undiagnosed HIV prevalence among adults and adolescents in the United States at the end of 2006. *J Acquir Immune Defic Syndr*. 2010;53:619–624.
13. Sweeting MJ, De Angelis D, Aalen OO. Bayesian back-calculation using a multi-state model with application to HIV. *Stat Med*. 2005;24:3991–4007.
14. Ndawinz JD, Costagliola D, Supervie V. New method for estimating HIV incidence and time from infection to diagnosis using HIV surveillance data: results for France. *AIDS*. 2011;25:1905–1913.
15. Becker NG, Lewis JJ, Li Z, McDonald A. Age-specific back-projection of HIV diagnosis data. *Stat Med*. 2003;22:2177–2190.
16. Centers for Disease Control and Prevention. 1993 revised classification system for HIV infection and expanded surveillance case definition for AIDS among adolescents and adults. *MMWR Morb Mortal Wkly Rep*. 1992;41:1–19.
17. Lodi S, Phillips A, Touloumi G, et al.; CASCADE Collaboration in EuroCoord. Time from human immunodeficiency virus seroconversion to reaching CD4+ cell count thresholds <200, <350, and <500 Cells/mm^3: assessment of need following changes in treatment guidelines. *Clin Infect Dis*. 2011;53:817–825.
18. Cori A, Ayles H, Beyers N, et al.; HPTN 071 PopART Study Team. HPTN 071 (PopART): a cluster-randomized trial of the population impact of an HIV combination prevention intervention including universal testing and treatment: mathematical model. *PLoS One*. 2014;9:e84511.
19. Bezemer D, de Wolf F, Boerlijst MC, van Sighem A, Hollingsworth TD, Fraser C. 27 years of the HIV epidemic amongst men having sex with men in the Netherlands: an in depth mathematical model-based analysis. *Epidemics*. 2010;2:66–79.
20. Bochow M, Chiarotti F, Davies P, et al. Sexual behaviour of gay and bisexual men in eight European countries. *AIDS Care*. 1994;6:533–549.
21. Mikolajczak J, Hospers HJ, Kok G. Reasons for not taking an HIV-test among untested men who have sex with men: an Internet study. *AIDS Behav*. 2006;10:431–435.
22. Alioum A, Commenges D, Thiebaut R, Dabis F. A multistate approach for estimating the incidence of human immunodeficiency virus by using data from a prevalent cohort study. *Appl Statist*. 2005;54:739–752.
23. van Sighem A, Gras L, Kesselring A, Smit C, et al; on behalf of the Netherlands collaborative HIV treatment centres. Monitoring Report 2013. Human Immunodeficiency Virus (HIV) Infection in the Netherlands. Amsterdam, The Netherlands: Stichting HIV Monitoring, 2013.
24. Soetens LC, Koedijk FDH, van den Broek IVF, et al. Sexually transmitted infections, including HIV, in the Netherlands in 2012. RIVM report 150002003/2013. Bilthoven, The Netherlands: National Institute for Public Health and the Environment, Ministry of Health, Welfare and Sport, 2013.

25. Sanchez MA, Blower SM. Uncertainty and sensitivity analysis of the basic reproductive rate. Tuberculosis as an example. *Am J Epidemiol*. 1997;145:1127–1137.
26. van Sighem A, Vidondo B, Glass TR, et al.; Swiss HIV Cohort Study. Resurgence of HIV infection among men who have sex with men in Switzerland: mathematical modelling study. *PLoS One*. 2012;7:e44819.
27. Phillips AN, Sabin C, Pillay D, Lundgren JD. HIV in the UK 1980-2006: reconstruction using a model of HIV infection and the effect of antiretroviral therapy. *HIV Med*. 2007;8:536–546.
28. CASCADE (Concerted Action on SeroConversion to AIDS and Death in Europe) Collaboration. Time from HIV-1 seroconversion to AIDS and death before widespread use of highly-active antiretroviral therapy: a collaborative re-analysis. *Lancet* 2000; 355:1131–1137.
29. Lodi S, Phillips A, Touloumi G, et al.; CASCADE Collaboration. CD4 decline in seroconverter and seroprevalent individuals in the precombination of antiretroviral therapy era. *AIDS*. 2010;24:2697–2704.
30. Natural History Project Working Group for the Collaboration of Observational HIV Epidemiological Research Europe (COHERE) in EuroCoord. Factors associated with short-term changes in HIV viral load and CD4+ cell count in antiretroviral-naive individuals. *AIDS* 2014; 28:1351–1356.
31. Müller V, Maggiolo F, Suter F, et al. Increasing clinical virulence in two decades of the Italian HIV epidemic. *PLoS Pathog*. 2009;5:e1000454.
32. Gras L, Geskus RB, Jurriaans S, et al.; ATHENA National Observational Cohort. Has the rate of CD4 cell count decline before initiation of antiretroviral therapy changed over the course of the Dutch HIV epidemic among MSM? *PLoS One*. 2013;8:e64437.
33. Goubar A, Ades AE, De Angelis D, et al. Estimates of human immunodeficiency virus prevalence and proportion diagnosed based on Bayesian multiparameter synthesis of surveillance data. *J R Statist Soc A*. 2008;171:541–567.
34. van Empelen P, van Berkel M, Roos E, Zuilhof W. *Schorer Monitor 2011*. Amsterdam, The Netherlands: Schorer; 2011.
35. Supervie V, Ndawinz JD, Lodi S, Costagliola D. The undiagnosed HIV epidemic in France and its implications for HIV screening strategies. *AIDS*. 2014;28:1797–1804.
36. Panel on Antiretroviral Guidelines for Adults and Adolescents. Guidelines for the Use of Antiretroviral Agents in HIV-1-Infected Adults and Adolescents. 2014. Available at: http://aidsinfo.nih.gov/contentfiles/lvguidelines/adultandadolescentgl.pdf. Accessed May 7, 2014.
37. Mocroft A, Lundgren JD, Sabin ML, et al.; Collaboration of Observational HIV Epidemiological Research Europe (COHERE) study in EuroCoord. Risk factors and outcomes for late presentation for HIV-positive persons in Europe: results from the Collaboration of Observational HIV Epidemiological Research Europe Study (COHERE). *PLoS Med*. 2013;10:e1001510.
38. Gras L, Kesselring AM, Griffin JT, et al.; ATHENA, Netherlands National Observational Cohort Study. CD4 cell counts of 800 cells/mm^3 or greater after 7 years of highly active antiretroviral therapy are feasible in most patients starting with 350 cells/mm3 or greater. *J Acquir Immune Defic Syndr*. 2007;45:183–192.
39. Jansen IA, Geskus RB, Davidovich U, et al. Ongoing HIV-1 transmission among men who have sex with men in Amsterdam: a 25-year prospective cohort study. *AIDS*. 2011;25:493–501.
40. van Sighem A, Jansen I, Bezemer D, et al. Increasing sexual risk behaviour among Dutch men who have sex with men: mathematical models versus prospective cohort data. *AIDS*. 2012;26:1840–1843.
41. Marks G, Crepaz N, Senterfitt JW, Janssen RS. Meta-analysis of high-risk sexual behavior in persons aware and unaware they are infected with HIV in the United States: implications for HIV prevention programs. *J Acquir Immune Defic Syndr*. 2005;39:446–453.
42. Heijman T, Geskus RB, Davidovich U, Coutinho RA, Prins M, Stolte IG. Less decrease in risk behaviour from pre-HIV to post-HIV seroconversion among MSM in the combination antiretroviral therapy era compared with the pre-combination antiretroviral therapy era. *AIDS*. 2012;26:489–495.