Inferring HIV incidence from case surveillance with CD4$^+$ cell counts

James Jansson[a], Cliff C. Kerr[a,b], Kylie-Ann Mallitt[a], Jianyun Wu[a], Richard T. Gray[a] and David P. Wilson[a]

Background: In some countries, HIV surveillance is based on case-reporting of newly diagnosed infections. We present a new back-projection method for estimating HIV-incidence trends using individuals' CD4$^+$ cell counts at diagnosis.

Methods: On the basis of a review of CD4$^+$ cell count distributions among HIV-uninfected people, CD4$^+$ cell count following primary infection, and rates of CD4$^+$ cell count decline over time among people with HIV, we simulate the expected distribution in time between infection and diagnosis. Applying this to all diagnosed individuals provides a distribution of likely infection times and estimates for population incidence, level of undiagnosed HIV, and the average time from infection to diagnosis each year. We applied this method to the national HIV case surveillance data of Australia for 1983–2013.

Results: The estimated number of new HIV infections in Australia in 2013 was 912 (95% uncertainty bound 835–1002). We estimate that 2280 (95% uncertainty bound 1900–2830) people were living with undiagnosed HIV at the end of 2013, corresponding to approximately 9.4% (95% uncertainty bound 7.8–10.1%) of all people living with HIV. With increases in the average CD4$^+$ count at diagnosis, the inferred HIV testing rate has been increasing over time and the estimated mean and median times between infection and diagnosis have decreased substantially. However, the estimated mean time between infection and diagnosis is considerably greater than the median, indicating that some people remain undiagnosed for long periods. Differences were found between cases attributable to male homosexual exposure versus other cases.

Conclusion: This methodology provides a novel way of estimating population incidence by combining diagnosis dates and CD4$^+$ cell counts at diagnosis.

AIDS 2015, **29**:1517–1525

Keywords: Australia, back projection, case surveillance, HIV, incidence, surveillance

Introduction

Public health surveillance systems provide information about the state of epidemics, informing decision-making around prevention, and clinical responses and impact evaluations of interventions. In most high-income countries, routine surveillance for HIV is based on case-reporting of all new HIV diagnoses. However, these surveillance practices imprecisely capture changes in HIV-incidence patterns due to the time lag between infection and diagnosis. Incidence is a critical measure for making up-to-date public health decisions, but it is difficult to capture accurately. Direct observation of incidence through ongoing cohort studies is infeasible for population-wide estimates, especially in low-incidence countries, as a large number of observations are necessary to provide sufficient accuracy. Cohort studies of at-risk groups can reduce the number of observations necessary,

[a]Kirby Institute, UNSW Australia, and [b]Complex Systems Group, School of Physics, University of Sydney, Sydney, New South Wales, Australia.

Correspondence to David P. Wilson, The Kirby Institute, University of New South Wales, Sydney, NSW 2052, Australia.
Tel: +61 2 9385 0900; fax: +61 2 9385 0920; e-mail: dwilson@unsw.edu.au
Received: 19 August 2014; revised: 27 February 2015; accepted: 18 March 2015.

DOI:10.1097/QAD.0000000000000679

ISSN 0269-9370

but continual recruitment to such studies is expensive to maintain and may not be representative of the general population. Most high-income countries base epidemic assessments on numbers of diagnoses, sometimes accompanied by statistical back-projection methods that estimate population incidence. In countries that base their population HIV surveillance on case notification, data are collected on factors such as demography, likely route of HIV acquisition, and $CD4^+$ cell count. $CD4^+$ cell count is collected because it provides an indication of progression in disease and thus also eligibility to initiate antiretroviral therapy (ART). Although $CD4^+$ measurements vary between and within individuals [1,2], because they provide indications of disease progression, they can be used to estimate a range in likely time from infection to diagnosis. We posit that applying this concept to surveillance data in populations can provide estimates of distributions of time from infection to diagnosis across all diagnosed cases, from which trends in population incidence can also be inferred.

The Working Group on Estimation of HIV Prevalence in Europe produced a review paper of methods [3–9] that combine diagnosis data with additional information such as disease progression and testing rates to estimate the incidence and number of people living with undiagnosed HIV [10]. Some of these methods have been complicated over time by the introduction of ART which has changed the expected disease history of people living with HIV (PLHIV), whereas new tests such as incidence assays and the routine measurement of $CD4^+$ cell counts have improved the available data.

$CD4^+$ cell counts at diagnosis have previously informed estimates of HIV incidence in England and Wales through their use in mathematical models [11,12]. These multistate compartmental models classify the population of PLHIV into bands of $CD4^+$ cell counts, and use Bayesian methods to calculate probabilistic transfer between bands of $CD4^+$ cell counts and the probability of diagnosis. Other methods have used $CD4^+$ cell count at diagnosis and simultaneous HIV/AIDS diagnosis to estimate the number of people with undiagnosed HIV infections [13].

Here, we describe a novel method to infer the time between infection and diagnosis for individuals based on $CD4^+$ cell count at diagnosis. We apply the method to data from Australia's National HIV Registry to estimate population incidence. Our approach differs from the previous approaches because it calculates the continuous expected decline and variability in $CD4^+$ cell count within individuals in the population using results from observational studies, including starting $CD4^+$ cell count, $CD4^+$ decline, and within-individual $CD4^+$ variability. This method not only provides estimates for incidence but also estimates for average time between infection and diagnosis, annual testing rates in the population, the total number of undiagnosed people, and associated uncertainties in these measures.

Methods

Data source

We demonstrated our back-projection methodology using diagnosis data from the Australian National HIV Registry [14], which contains data for all 35 286 individuals diagnosed with HIV in Australia up to 31 December 2013. Of these records, 16 568 had $CD4^+$ cell counts at diagnosis (taken within 3 months of diagnosis), with increasing testing and reporting over time such that the vast majority of cases in recent years have an associated $CD4^+$ cell count around diagnosis. Missing $CD4^+$ cell counts at diagnosis were imputed using Markov Chain Monte Carlo multiple imputation over the variables of sex, year of diagnosis, age, and square root $CD4^+$ cell count at diagnosis. A total of 18 718 (53%) $CD4^+$ cell counts were imputed, with 9882 $CD4^+$ cell counts (93.1%) for the 10 614 diagnoses prior to 1990. $CD4^+$ cell counts were recorded in 82% of cases after 2000. We excluded 1226 people who were previously diagnosed overseas, leaving 34 060 local diagnoses. Multiple reporting of the same diagnosis is accounted for by using a method proposed by Larsen [15], in which the number of unique birthdates are used to estimate the expected number of diagnoses of people born in that year. This method has been used previously to account for multiple reporting in Australia [16]. We estimated that 3113 of the local diagnoses were duplicates, leaving 30 947 diagnoses for the calculation of incidence in Australia.

Simulation overview

To back-project incidence from diagnoses, we used the following variables: evidence of previous HIV-negative date, indeterminate western blot, recent illness indicating seroconversion sickness, $CD4^+$ cell count at diagnosis, and date of diagnosis. To estimate the time between infection and diagnosis, we simulated $CD4^+$ cell count change among treatment-naive PLHIV.

Simulating $CD4^+$ progression within individuals

In the simulation, $CD4^+$ cell counts start from a HIV-uninfected $CD4^+$ distribution. From a variety of reported $CD4^+$ cell count distributions [17–33], we created a bootstrapped dataset through simulation and random resampling (weighted according to the study sample size). We used the median and SD of the log-normal distribution of this bootstrapped data for our simulations of $CD4^+$ decline.

Following infection, $CD4^+$ cell count rapidly declined to a median of 418 cells/μl at day 17 [34] and subsequently rebounded to a median of 636 (95% uncertainty bound 586–686) at day 40 during the primary infection stage

[17–33] (see Appendices 2 and 3 for justifications, http://links.lww.com/QAD/A690). The simulated CD4$^+$ progression then followed a square root decline of 1.6 square root cells/μl (95% uncertainty bound 1.4–1.8) per year [35–47] until the CD4$^+$ cell count reached an average of 61 cells/μl, from which point the decline progressed linearly (see Appendix 4 for justifications, http://links.lww.com/QAD/A690). CD4$^+$ cell counts of individuals between measurements can vary greatly [2,48]. We incorporated variability in the measured CD4$^+$ cell count about the expected mean for the individual at that point in time using the equation from Hughes *et al.* ($\sigma = 0.930 - 0.110 \log_e \mu$, where σ is the SD about the expected mean μ of the CD4$^+$ cell count at that point of time in the $\log_e$ scale [49]). This accounts for individual variability in repeated CD4$^+$ measurements.

Estimation of average testing rate

We generated a sample population by simulating CD4$^+$ cell count decline with time and applying an annual testing probability, which increases exponentially as CD4$^+$ cell counts approach zero due to symptomatic presentation (the rate of change in testing rate with CD4$^+$ cell count is fitted to data in the registry). The testing probability applied to the simulated population is adjusted until the distribution of the CD4$^+$ cell count at diagnosis of the simulated sample population matches the distribution of the CD4$^+$ cell count at diagnosis of the empirical diagnosis data of the HIV Registry. This optimization method is repeated for 200 different parameterizations (based on uncertainty in the parameter estimates) and for each year in the simulation to account for stochastic variability and changes in testing patterns over time.

Selecting time of infection for each diagnosis

To estimate the time since infection for each individual diagnosis in the Australian National HIV Registry, the model matches the CD4$^+$ cell count at diagnosis with simulated CD4$^+$ cell count declines. For example, if an individual has a CD4$^+$ cell count of 350 cells/μl at diagnosis, then the model randomly selects a simulated CD4$^+$ cell count at diagnosis from our set of simulations, with CD4$^+$ cell counts at diagnosis of 350 ± 10 cells/μl. The time until diagnosis of the selected simulation is then the estimated time until diagnosis of the diagnosed individual. This linking process occurs for each of the 200 parameterizations.

Additional adjustments due to data acquired in recent years

Individual data recording recent indeterminate Western blot or a recent illness indicating seroconversion is used to estimate whether an infection occurred within 40 days of these observations. In cases when the back-projected estimated date of infection occurs prior to the last negative HIV test result, a random (uniform) time between the last negative date and the date of HIV diagnosis was selected.

Estimating incidence of currently undiagnosed infections

The back-projection of diagnoses to incidence allows estimation of a range in the date of infection of already diagnosed infections. However, the majority of infections occurring in 2013 remained undiagnosed by the end of 2013. Hence, to determine the total incidence, numbers of undiagnosed infections need to be estimated. To estimate the number of incident but undiagnosed cases, the model first pools the results for the estimated times between infection and diagnosis (up to 20 years) for diagnoses over the final 5 years of diagnoses data for sampling (the model uses this pool for all intervals). It applies this distribution of duration of time from infection to diagnosis to infer the extent of undiagnosed cases. A detailed description of the resampling method is provided in the Appendix Section 8 (http://links.lww.com/QAD/A690).

Uncertainty

The simulation produces a distribution for the time between infection and diagnosis to incorporate the uncertainty associated with CD4$^+$ measurements and declines. The approach also includes variability in CD4$^+$ cell counts for uninfected individuals [17–33], uncertainty in the decline in CD4$^+$ cell count during primary infection [34,44,46,50], uncertainty in the mean rate of decline in CD4$^+$ cell counts [35–46], variability in individual rates of CD4$^+$ decline [42], and variability of individual CD4$^+$ cell counts with repeat measurement [49]. Appendix 1 (http://links.lww.com/QAD/A690) provides further information on the methodology, simulation parameter values, and uncertainty ranges. Throughout this study, the 95% uncertainty bound is the result produced by, or used in, the model in which the central 95% of results lie. This methodology was implemented in MATLAB and the source code is publicly available [51].

Results

In Fig. 1, the simulated CD4$^+$ progression of a population of untreated PLHIV is shown. The HIV-uninfected CD4$^+$ cell count distribution has a median of 890 [interquartile range (IQR) 707–1119] cells/μl. Upon HIV infection, the CD4$^+$ cell count decreases and then recovers quickly during the primary infection stage. The estimated time for the median CD4$^+$ cell count of untreated PLHIV to reach 500 cells/μl is 2.0 (95% uncertainty bound 1.35–2.7) years. To reach 350 cells/μl, it requires 4.4 (median, 95% uncertainty bound 3.6–5.3) years and to reach 200 cells/μl, it requires 7.0 (median, 95% uncertainty bound 6.2–8.1) years. Note

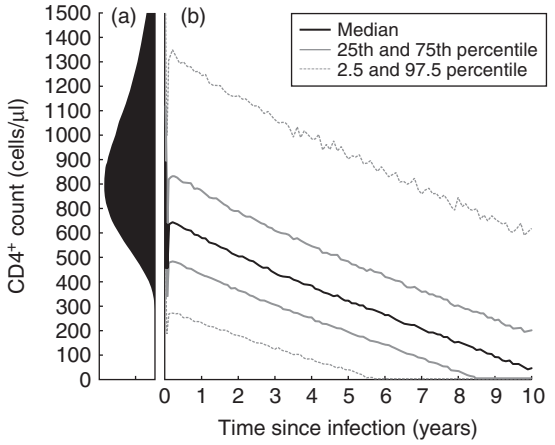

Fig. 1. The (a) probability distribution function of HIV-uninfected $CD4^+$ cell counts and (b) the simulation of $CD4^+$ cell count rapid decline, rebound and then steady decline following HIV.

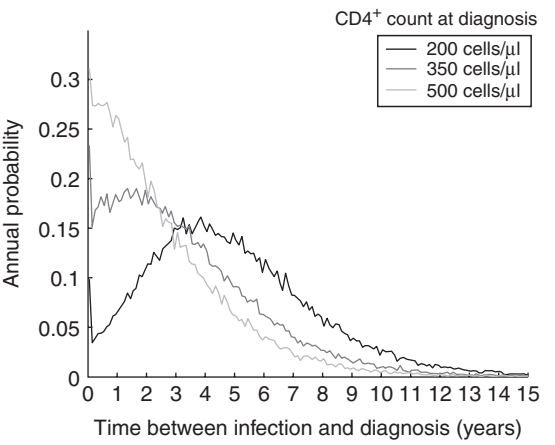

Fig. 2. Sample distribution of time from infection to diagnosis by $CD4^+$ cell count at diagnosis in 2013 for Australian HIV surveillance data.

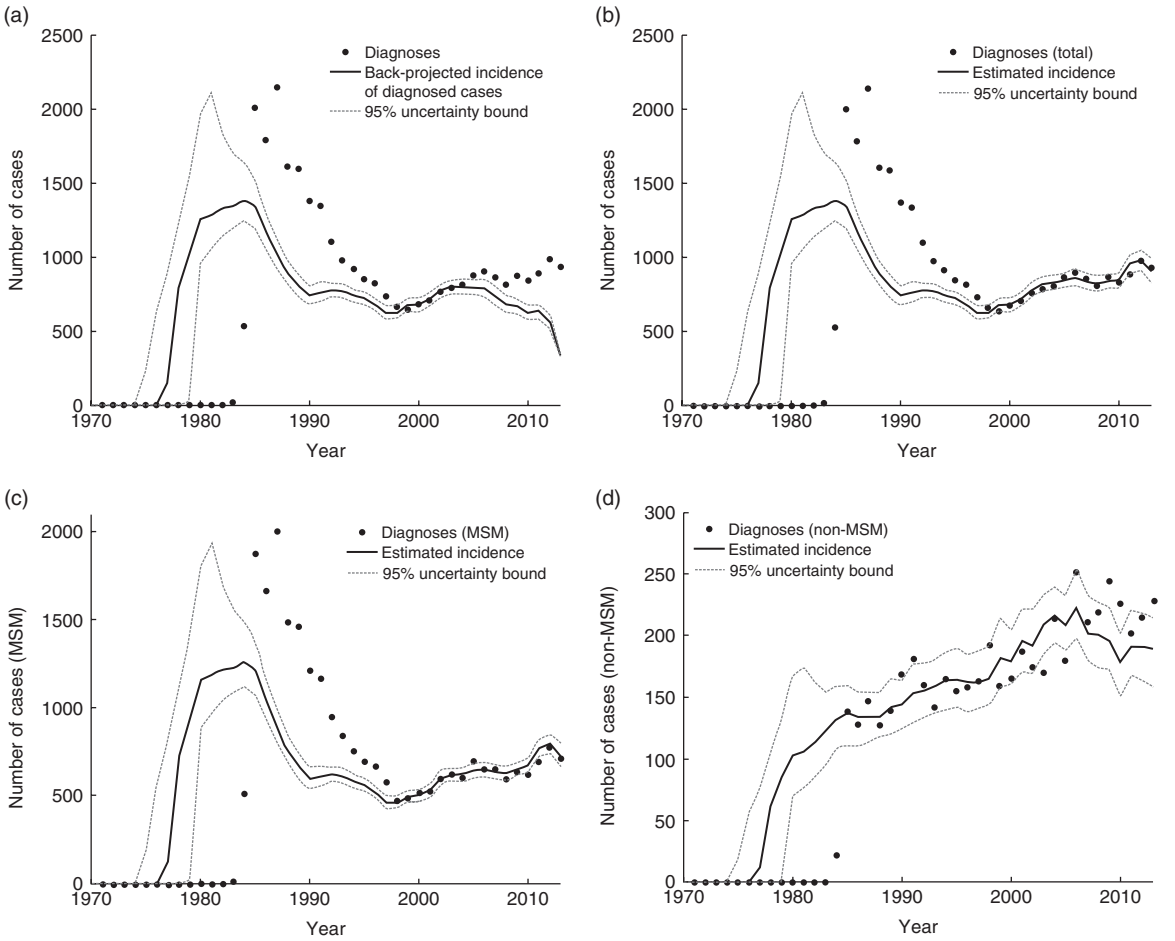

Fig. 3. (a) Back-projected infections of diagnosed individuals in Australia. (b) Sum of back-projected and estimated undiagnosed infections. (c) Sum of back-projected and estimated undiagnosed infections in Australia for MSM. (d) Sum of back-projected and estimated undiagnosed infections in Australia for people who are not MSM.

that while the representation of the CD4$^+$ declines in Fig. 1b are relatively smooth, these are summary statistics of the population. The CD4$^+$ cell count for an individual in the simulation can vary greatly with time and would appear much less smooth.

The CD4$^+$ cell count at diagnosis depends on both the time since infection and the annual probability of testing (and thus being diagnosed). This means the median time for the CD4$^+$ cell count of an untreated person to reach 200 cells/μl (7.0 years; Fig. 1b) is different from the time between infection and diagnosis for a person with a CD4$^+$ cell count of 200 cells/μl (5.0 years; Fig. 2), because the CD4$^+$ cell count at diagnosis depends on both the time since infection and the annual probability of diagnosis (i.e. testing). For example, a person who starts with a CD4$^+$ cell count of 1100 cells/μl will take longer, on average, to reach a CD4$^+$ cell count of 200 cells/μl than a person who starts with a CD4$^+$ cell count of 700 cells/μl. The individual with 1100 cells/μl has more opportunities for testing prior to reaching 200 cells/μl and hence is less likely to reach 200 cells/μl. In Australia, the median estimated time between infection and diagnosis in 2013 was 2.0 (IQR 0.9–3.6) years for diagnoses with CD4$^+$ cell counts of 500 cells/μl at diagnosis, 2.8 (IQR: 1.4–4.7) years for 350 cells/μl, and 4.5 (IQR: 2.9–6.6) years for 200 cells/μl (Fig. 2). The distribution of time until diagnosis in Fig. 2 is specific to Australia in 2013.

The estimated number of cases diagnosed in 2013 in which their infection also occurred in 2013 was 350 (95% uncertainty bound 331–377) (Fig. 3a). In Fig. 3b, the combined estimates of currently diagnosed and undiagnosed infections are presented. The peak of infections [1384 (95% uncertainty bound 1251–1647)] is estimated to have occurred in 1984, at a level that is lower than the peak in diagnoses (2145) in 1987. This is due to a backlog of undiagnosed infections prior to the discovery of HIV and the availability of HIV-diagnostic testing. Following this peak, the estimated number of infections reached a low of 627 (95% uncertainty bound 586–677) in 1997.

In recent years, the trend and magnitude in diagnoses are estimated to reflect incidence, but with a slight lag (Fig. 3b). There were an estimated 912 (95% uncertainty bound 835–1002) infections in 2013, slightly less than the 938 estimated diagnosed cases. The high uncertainty in the incidence extrapolation is due to the small number of back-projected cases available in recent years. Future diagnosis data will be necessary to reduce the uncertainty regarding the number of infections in 2013. Since the HIV epidemic in Australia is predominantly driven by MSM, we separately conducted incidence back-estimates, disaggregating the total data into cases attributable to male homosexual exposure and other cases. Figure 3c and d shows the estimated incidence of HIV infection for MSM and non-MSM in Australia, respectively.

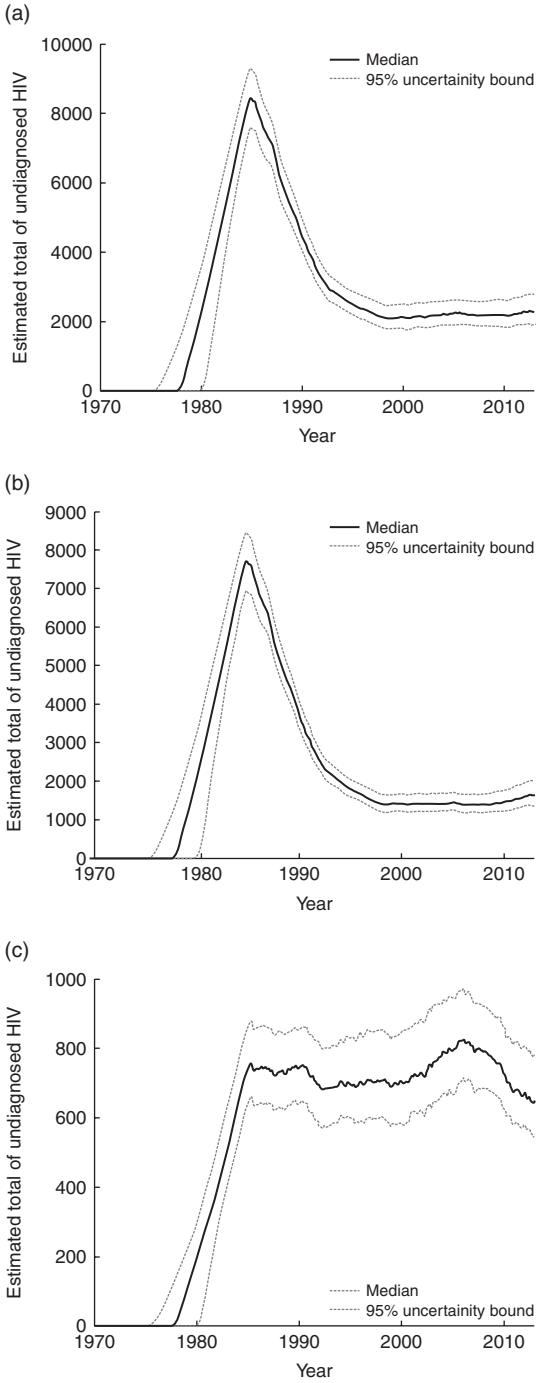

Fig. 4. Estimated number of people living with undiagnosed HIV in Australia by year for (a) the entire population, (b) MSM, and (c) people who are not MSM.

By combining the known time of diagnosis and the estimated range in times of infections, the total number of people with undiagnosed HIV at any point in time is estimated (see Fig. 4). The model estimates that the number of people living with undiagnosed HIV peaked in 1985 at 8456 (95% uncertainty bound 7616–9319), before decreasing to a low of 2084 (95% uncertainty bound 1795–2454) people in 1998. At the end of 2013,

an estimated 2281 (95% uncertainty bound 1900–2454) people were living with undiagnosed HIV. We estimate that there are 24 359 people living with currently diagnosed HIV, using a method described previously [52]; this corresponds to 9.4% (95% uncertainty bound 7.8–10.1%) of PLHIV being undiagnosed overall, consisting of 7.9% (6.4–9.5%) for MSM and 10.7% (8.9–12.8%) for non-MSM.

The annual probability of testing for undiagnosed PLHIV in Australia is estimated to have stayed relatively constant over time (Fig. 5a). The model estimates an annual testing probability for those who have acquired HIV of 0.26 (95% uncertainty bound 0.23–0.30) in 1999 – the year coinciding with the lowest number of diagnoses. This model deduces that testing probabilities have increased slightly over time, with the annual testing probability in 2013 at 0.34 (95% uncertainty bound 0.28–0.39). The estimated time between infection and diagnosis has reduced substantially from a median of 4.1 (IQR 2.1–5.7) years for people diagnosed in 1985 to a median of 0.9 (IQR 0.1–3.7) years for people diagnosed in 2013 (Fig. 5b). This reflects the reported increase in $CD4^+$ cell count at diagnosis over time (see Appendix Fig. A2, http://links.lww.com/QAD/A690). However, the mean time between infection and diagnosis was 2.4 (95% uncertainty bound 2.0–3.0) years in 2013; the right skew of this distribution indicates some people are undiagnosed for longer periods. The mean time between infection and diagnosis was estimated to differ by population group in Australia: 2.0 (1.7–2.5) years in 2013 for MSM and 3.5 (2.9–4.4) years for non-MSM.

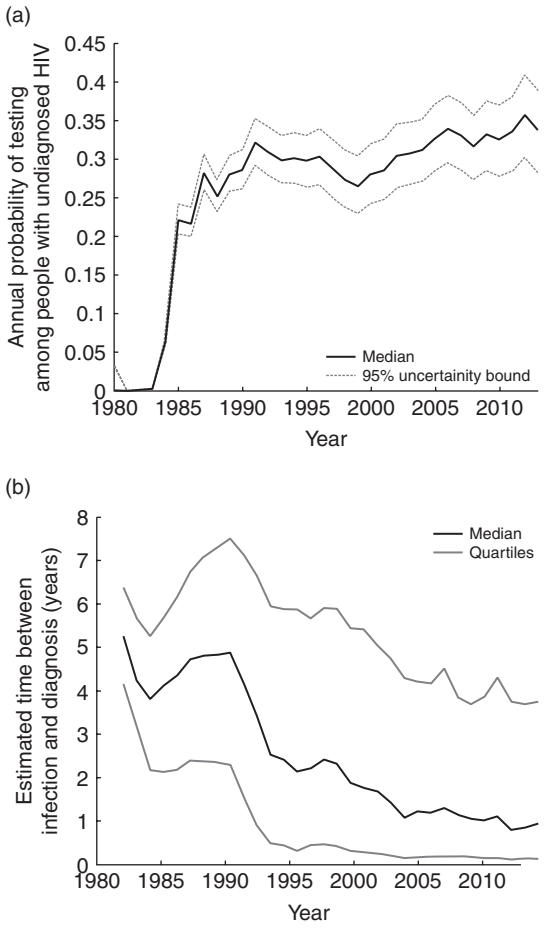

Fig. 5. (a) Estimated testing rate of undiagnosed HIV-infected people in Australia over time. (b) Median time between infection and testing by year of diagnosis.

Discussion

Our $CD4^+$-based back-projection methodology translates surveillance registries recording information on HIV-diagnosed cases into estimates of numbers of new infections by year, size of the undiagnosed population, annual probability of testing, and distributions in the time between infection and diagnosis. An advantage of the methodology is that it gives a distribution of expected times between infection and diagnosis for each person's $CD4^+$, not simply an aggregate estimate of incidence.

Using $CD4^+$ cell counts to estimate time between infection and diagnosis produces relatively large uncertainty bounds as our method captures the large variability in $CD4^+$ cell count between and within individuals. The wide error bounds for 2013 incidence are indicative of this intrinsic uncertainty. Given further data for the following years, the error bounds for 2013 will reduce substantially as the data from 2014 and later years will inform the size of the undiagnosed population in 2013. This methodology demonstrates the substantial difference in the trends in incidence between MSM and those who are not MSM, with a substantial growth in incidence in MSM followed by a decrease in the 1980s. We estimate that there has been a steady increase in the incidence of HIV from 1985 onwards among non-MSM in Australia.

The choice of median $CD4^+$ cell count following the primary infection (peak of $CD4^+$ rebound level) and the method's structural assumption of a square root decline influences the outcome of time between infection and diagnosis and hence all results in the model (as demonstrated in section 12 of the Appendix, http://links.lww.com/QAD/A690). There is also a large degree of uncertainty in these key parameters but our model incorporates the range of values obtained from multiple studies in order to produce results reflective of this uncertainty (see Appendix sections 3, 4, and 5, http://links.lww.com/QAD/A690). However, our approach ignores environmental or other factors, which can affect $CD4^+$ cell count. For example, $CD4^+$ cell count may be affected by smoking [53] and $CD4^+$ decline may be affected by previous intravenous drug use [54]. Differences in $CD4^+$ cell counts and declines by age [55] and between sexes [56] are not taken into account in this

model. The population of PLHIV in Australia is overwhelmingly men, which is in alignment with the majority of studies used to source parameters for this study. Therefore, the results reported in this study should not be heavily affected by this limitation, but an analysis of the incidence in women only may require an adjustment of parameters used in the model.

The lack of testing prior to 1983 causes the interaction of testing rate and CD4$^+$ cell count during the optimization phase to be unconstrained, increasing the uncertainty in results. The uncertainty bounds of the incidence curve in the early 1980s and 1970s only represent the uncertainty in the numerical calculations, and do not incorporate the uncertainty associated with the limitations of the model for this period (Fig. 3). As CD4$^+$ cell count was not reported routinely during the early stages of the epidemic, many records required an imputed CD4$^+$ cell count (93% of cases prior to 1990). This may skew our results for earlier years as CD4$^+$ cell counts were likely collected more often from those who showed signs of immunosuppression. Additionally, imputation may decrease the uncertainty in the data for earlier years and hence present uncertainty bounds that are smaller than expected, given the limited data.

Early back-projection models relied on progression to AIDS in blood recipients [3] that may not be representative of the majority of the population with HIV, given the high degree of variability of CD4$^+$ cell counts and dependence of progression to AIDS based on both lifestyle [57] and age [58]. Models that rely heavily on AIDS diagnosis as a determinant of time between infection and diagnosis [5,8,9] have become less relevant as AIDS becomes rarer, especially in high-income countries, due to earlier testing and initiation of effective ART.

Estimates for HIV incidence in Australia have previously been made using a modified back-projection method that combines HIV diagnosis, AIDS diagnosis, and recent infection information [59]. This method was used to estimate incidence by age cohort [60], subpopulation category (MSM, heterosexual, people who inject drugs) [61], and Australian state and territory [62], and relies on AIDS diagnoses and recent diagnosis data to inform the calculation of an incidence estimate. We believe that our new method offers an improvement on this method as we consider CD4$^+$ at diagnosis in addition to recent infection.

A simple approach using CD4$^+$ cell count around diagnosis can be useful for assessing trends in early versus late diagnosis. Previously, simplified CD4$^+$-based methods have been used by us in policy forums and others (e.g. [12]) to estimate the average time from infection to diagnosis. For example, by assuming that the median CD4$^+$ cell count for someone without HIV is 891 cells/μl, approximately 25–30% of cells are lost in the first year of infection and ~60 cells/μl are lost per year thereafter (until a CD4$^+$ cell count of 200 cells/μl) then an average time from infection to diagnosis can be estimated based on the average CD4$^+$ cell count of newly diagnosed cases. In Australia, the median CD4$^+$ cell count at diagnosis in 2009-2013 was ~450 cells/μl, leading to an estimated average time from infection to diagnosis of ~4.5 years. The method we present in this paper yields substantially lower estimates, namely, 2.4 years as a mean and 0.9 years as a median. There are two main reasons for this difference. Firstly, the rate at which people at risk of HIV are tested in a population influences the population average time from infection to particular CD4$^+$ levels. This creates a selection bias; for example, in a population in which testing rates are typically high, if someone is diagnosed with a low CD4$^+$ cell count it is more likely that this person has a lower baseline CD4$^+$ cell count prior to infection than them arriving at that CD4$^+$ cell count after an extended period of infection. Secondly, the rate of CD4$^+$ decline averages around 60 cells/μl per year during the overall course of natural infection but is greater initially and slower later in infection. Our current method incorporated the statistical distribution in starting CD4$^+$ cell counts and decline rates across the population. It also inferred a time-dependent testing rate to match a simulated distribution of CD4$^+$ cell counts at diagnosis with the observed distribution of CD4$^+$ cell counts at diagnosis. Additionally, our current method incorporates separate information about diagnosed cases which have evidence of recent infection. These adjustments account for the majority of the difference in magnitude of estimates in time from infection to diagnosis. Both approaches provide consistent time trends in the duration of infection prior to diagnosis. We believe that our new method provides a more accurate estimate of the average time between infection and diagnosis. However, this is subject to future validation. The range of estimates across methods suggests that currently there remains a lack of precision for estimating the magnitude of this indicator.

The number of diagnoses is not adjusted for underreporting in this study because we believe that available records are representative of the total diagnoses in Australia. HIV is a notifiable infection in Australia, which means new diagnoses must be reported to the relevant state authority. Additionally, subsequent adjustment of total reported diagnoses in later years have been low (<2% increase). Although we believe adjustment for underreporting and delayed notification is unnecessary in Australia, such an adjustment may be necessary in other locations.

Our method may provide an improvement for estimating HIV incidence and the time between infection and diagnosis from routinely collected diagnosis data, or at least provide supplemental information on existing

approaches. This is useful for understanding the current and future direction of HIV epidemics and the likely impact of HIV interventions and public health policy.

Acknowledgements

The study was funded by the following sources: the Australian Government Department of Health and Ageing; Australian Research Council (DP1093026; FT0991990); National Health and Medical Research Council.

Conflicts of interest

The views expressed in this publication do not necessarily represent the position of the Australian Government. The Kirby Institute is affiliated with the Faculty of Medicine, University of New South Wales.

References

1. Korenromp EL, Williams BG, Schmid GP, Dye C. **Clinical prognostic value of RNA viral load and cd4 cell counts during untreated HIV-1 infection: a quantitative review.** *PLoS One* 2009; **4**:e5950.
2. Malone JL, Simms TE, Gray GC, Wagner KF, Burge JR, Burke DS. **Sources of variability in repeated T-helper lymphocyte counts from human immunodeficiency virus type 1-infected patients: total lymphocyte count fluctuations and diurnal cycle are important.** *J Acquir Immune Defic Syndr* 1990; **3**:144–151.
3. Brookmeyer R, Gail MH. **A method for obtaining short-term projections and lower bounds on the size of the AIDS epidemic.** *J Am Stat Assoc* 1988; **83**:301–308.
4. Brookmeyer R, Gail MH. **Minimum size of the acquired immunodeficiency syndrome (AIDS) epidemic in the United States.** *Lancet* 1986; **328**:1320–1322.
5. Cui J, Becker NG. **Estimating HIV incidence using dates of both HIV and AIDS diagnoses.** *Stat Med* 2000; **19**:1165–1177.
6. Becker NG, Lewis JJC, Li Z, McDonald A. **Age-specific back-projection of HIV diagnosis data.** *Stat Med* 2003; **22**:2177–2190.
7. Hall H, Song R, Rhodes P, Prejean J, An Q, *et al*. **Estimation of HIV incidence in the United States.** *JAMA* 2008; **300**:520–529.
8. Aalen OO, Farewell VT, de Angelis D, Day NE, Nöel Gill O. **A Markov model for HIV disease progression including the effect of HIV diagnosis and treatment: application to AIDS prediction in England and Wales.** *Stat Med* 1997; **16**:2191–2210.
9. Sommen C, Alioum A, Commenges D. **A multistate approach for estimating the incidence of human immunodeficiency virus by using HIV and AIDS French surveillance data.** *Stat Med* 2009; **28**:1554–1568.
10. 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.
11. Sweeting MJ, De Angelis D, Aalen OO. **Bayesian back-calculation using a multistate model with application to HIV.** *Stat Med* 2005; **24**:3991–4007.
12. Birrell PJ, Gill ON, Delpech VC, Brown AE, Desai S, Chadborn TR. **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.
13. Lodwick R, Sabin C, Phillips A. **A method to estimate the number of people in a country or region with HIV who are undiagnosed and in need of ART.** *J Int AIDS Soc* 2010; **13**:165–168.
14. The Kirby Institute. *HIV, viral hepatitis and sexually transmissible infections in Australia Annual Surveillance Report 2013*. Sydney: The Kirby Institute, the University of New South Wales; 2013.
15. Larsen SO. **Estimation of the number of people in a register from the number of birthdates.** *Stat Med* 1994; **13**:177–183.
16. Law MG, McDonald AM, Kaldor JM. **Estimation of cumulative HIV incidence in Australia, based on national case reporting.** *Aust N Z J Public Health* 1996; **20**:215–217.
17. Tindall B, Cooper DA, Donovan B, Barnes T, Philpot CR, Gold J, Penny R. **The Sydney AIDS Project: development of acquired immunodeficiency syndrome in a group of HIV seropositive homosexual men.** *Aust N Z J Med* 1988; **18**:8–15.
18. Bodsworth NJ, Slade M. **Immunological assessment in the treatment of HIV infection.** *Pathology* 1992; **24**:230–1230.
19. Bryant JA, Wylie BR, Yuan FF, Ribeiro A, Thomson AR, Cooley MA, Fletcher A. **Effect of blood donation on the establishment of normal ranges of lymphocyte subsets.** *Transfusion* 1996; **36**:559–566.
20. Vuillier F, Lapresle C, Dighiero G. **Comparative analysis of CD4-4B4 and CD4-2H4 lymphocyte subpopulations in HIV negative homosexual, HIV seropositive and healthy subjects.** *Clin Exp Immunol* 1988; **71**:8–12.
21. Tollerud DJ, Clark JW, Brown LM, Neuland CY, Pankiw-Trost LK, Blattner WA, Hoover RN. **The influence of age, race, and gender on peripheral blood mononuclear-cell subsets in healthy nonsmokers.** *J Clin Immunol* 1989; **9**:214–222.
22. Giorgi JV, Cheng H-L, Margolick JB, Bauer KD, Ferbas J, Waxdal M. **Quality control in the flow cytometric measurement of T-lymphocyte subsets: The Multicenter AIDS Cohort Study experience.** *Clin Immunol Immunopathol* 1990; **55**:173–186.
23. Hannet I, Erkeller-Yuksel F, Lydyard P, Deneys V, DeBruyère M. **Developmental and maturational changes in human blood lymphocyte subpopulations.** *Immunol Today* 1992; **13**:215–218.
24. Bofill M, Janossy G, Lee CA, Macdonald-Burns D, Phillips AN, Sabin C, *et al*. **Laboratory control values for CD4 and CD8 T lymphocytes. Implications for HIV-1 diagnosis.** *Clin Exp Immunol* 1992; **88**:243–252.
25. Hulstaert F, Hannet I, Deneys V, Munhyeshuli V, Reichert T, De Bruyere M, Strauss K. **Age-related changes in human blood lymphocyte subpopulations: II. Varying kinetics of percentage and absolute count measurements.** *Clin Immunol Immunopathol* 1994; **70**:152–158.
26. Howard RR, Fasano CS, Frey L, Miller CH. **Reference intervals of CD3, CD4, CD8, CD4/CD8, and absolute CD4 values in Asian and non-Asian populations.** *Cytometry* 1996; **26**:231–232.
27. Comans-Bitter WM, de Groot R, van den Beemd R, Neijens HJ, Hop WCJ, Groeneveld K, *et al*. **Immunophenotyping of blood lymphocytes in childhood: reference values for lymphocyte subpopulations.** *J Pediatr* 1997; **130**:388–393.
28. Santagostino A, Garbaccio G, Pistorio A, Bolis V, Camisasca G, Pagliaro P, Girotto M. **An Italian national multicenter study for the definition of reference ranges for normal values of peripheral blood lymphocyte subsets in healthy adults.** *Haematologica* 1999; **84**:499–504.
29. Tsegaye A, Messele T, Tilahun T, Hailu E, Sahlu T, Doorly R. **Immunohematological reference ranges for adult Ethiopians.** *Clin Diagn Lab Immunol* 1999; **6**:410–414.
30. Kassu A, Tsegaye A, Petros B, Wolday D, Hailu E, Tilahun T, *et al*. **Distribution of lymphocyte subsets in healthy human immunodeficiency virus-negative adult Ethiopians from two geographic locales.** *Clin Diagn Lab Immunol* 2001; **8**:1171–1176.
31. Bisset LR, Lung TL, Kaelin M, Ludwig E, Dubs RW. **Reference values for peripheral blood lymphocyte phenotypes applicable to the healthy adult population in Switzerland.** *Eur J Haematol* 2004; **72**:203–212.
32. Jentsch-Ullrich K, Koenigsmann M, Mohren M, Franke A. **Lymphocyte subsets' reference ranges in an age- and gender-balanced population of 100 healthy adults: a monocentric German study.** *Clin Immunol* 2005; **116**:192–197.
33. Yaman A, Çetiner S, Kibar F, Taşova Y, Şeydaoğlu G, Dündar IH. **Reference ranges of lymphocyte subsets of healthy adults in Turkey.** *Med Princ Pract* 2005; **14**:189–193.
34. Kaufmann GR, Cunningham P, Zaunders J, Law M, Vizzard J, Carr A, *et al*. **Impact of early HIV-1 RNA and T-lymphocyte dynamics during primary HIV-1 infection on the subsequent course of HIV-1 RNA levels and CD4+ T-lymphocyte counts in the first year of HIV-1 infection.** *J Acquir Immune Defic Syndr* 1999; **22**:437–444.

35. Lee CA, Phillips A, Elford J, Miller EJ, Bofill M, Griffiths PD, Kernoff PB. **The natural history of human immunodeficiency virus infection in a haemophilic cohort.** *Br J Haematol* 1989; **73**:228–234.
36. Veugelers PJ, Strathdee SA, Kaldor JM, Shafer KAP, Moss AR, Schechter MT, *et al*. **Associations of age, immunosuppression, and AIDS among homosexual men in the Tricontinental Seroconverter Study.** *J Acquir Immune Defic Syndr* 1997; **14**:435–441.
37. Prins M, Robertson JR, Brettle RP, Aguado IH, Broers B, Boufassa F, *et al*. **Do gender differences in CD4 cell counts matter?** *AIDS* 1999; **13**:2361–2364.
38. Deeks SG, Kitchen CM, Liu L, Guo H, Gascon R, Narvaez AB, *et al*. **Immune activation set point during early HIV infection predicts subsequent CD4+ T-cell changes independent of viral load.** *Blood* 2004; **104**:942–947.
39. Mellors JW, Margolick JB, Phair JP, Rinaldo CR, Detels R, Jacobson LP, Munoz A. **Prognostic value of HIV-1 RNA, CD4 cell count, and CD4 cell count slope for progression to AIDS and death in untreated HIV-1 infection.** *JAMA* 2007; **297**:2349–2350.
40. Drylewicz J, Matheron S, Lazaro E, Damond F, Bonnet F, Simon F, *et al*. **Comparison of viro-immunological marker changes between HIV-1 and HIV-2-infected patients in France.** *AIDS* 2008; **22**:457–468.
41. Muller V, von Wyl V, Yerly S, Boni J, Klimkait T, Burgisser P, *et al*. **African descent is associated with slower CD4 cell count decline in treatment-naive patients of the Swiss HIV Cohort Study.** *AIDS* 2009; **23**:1269–1276.
42. Wolbers M, Babiker A, Sabin C, Young J, Dorrucci M, Chêne G, *et al*. **Pretreatment CD4 cell slope and progression to AIDS or death in HIV-infected patients initiating antiretroviral therapy: the CASCADE collaboration: a collaboration of 23 cohort studies.** *PLoS Med* 2010; **7**:e1000239.
43. Seng R, Lewden C, Thiebaut R, Boufassa F, Coulibaly A, Malateste K, *et al*. **Comparison of early CD4 T-cell count in HIV-1 seroconverters in Cote d'Ivoire and France: the ANRS PRIMO-CI and SEROCO cohorts.** *J Acquir Immune Defic Syndr* 2010; **53**:260–265.
44. CASCADE Collaboration. **Differences in CD4 cell counts at seroconversion and decline among 5739 HIV-1-infected individuals with well estimated dates of seroconversion.** *J Acquir Immune Defic Syndr* 2003; **34**:76–83.
45. Lodi S, Phillips A, Touloumi G, Pantazis N, Bucher HC, Babiker A. **CD4 decline in seroconverter and seroprevalent individuals in the precombination of antiretroviral therapy era.** *AIDS* 2010; **24**:2697–2704.
46. Lodi S, Phillips A, Touloumi G, Geskus R, Meyer L, Thiébaut R, *et al*. **Time from human immunodeficiency virus seroconversion to reaching CD4+ cell count thresholds <200, <350, and <500 cells/mm3: assessment of need following changes in treatment guidelines.** *Clin Infect Dis* 2011; **53**:817–825.
47. Keller M, Lu Y, Lalonde RG, Klein MB. **Impact of HIV-1 viral subtype on CD4+ T-cell decline and clinical outcomes in antiretroviral naive patients receiving universal healthcare.** *AIDS* 2009; **23**:731–737.
48. Malone JL, Oldfield EC, Wagner KF, Simms TE, Daly R, O'Brian J, Burke DS. **Abnormalities of morning serum cortisol levels and circadian rhythms of CD4+ lymphocyte counts in human immunodeficiency virus type 1-infected adult patients.** *J Infect Dis* 1992; **165**:185–186.
49. Hughes MD, Stein DS, Gundacker HM, Valentine FT, Phair JP, Volberding PA. **Within-subject variation in CD4 lymphocyte count in asymptomatic human immunodeficiency virus infection: implications for patient monitoring.** *J Infect Dis* 1994; **169**:28–36.
50. Lang W, Perkins H, Anderson RE, Royce R, Jewell N, Winkelstein WJ. **Patterns of T lymphocyte changes with human immunodeficiency virus infection: from seroconversion to the development of AIDS.** *J Acquir Immune Defic Syndr* 1989; **2**:63–69.
51. Jansson J. HIV back projection git repository. 2014. https://github.com/JamesJansson/HIVBackProjection.
52. Jansson J, Wilson DP. **Projected demographic profile of people living with HIV in Australia: planning for an older generation.** *PLoS One* 2012; **7**:e38334.
53. Maini MK, Gilson RJ, Chavda N, Gill S, Fakoya A, Ross EJ, *et al*. **Reference ranges and sources of variability of CD4 counts in HIV-seronegative women and men.** *Genitourin Med* 1996; **72**:27–31.
54. Meijerink H, Wisaksana R, Iskandar S, den Heijer M, van der Ven AJAM, Alisjahbana B, van Crevel R. Injecting drug use is associated with a more rapid CD4 cell decline among treatment naive HIV-positive patients in Indonesia. *J Int AIDS Soc* 2014; **17**.
55. Collaboration C. **Differences in CD4 cell counts at seroconversion and decline among 5739 HIV-1-infected individuals with well estimated dates of seroconversion.** *J Acquir Immune Defic Syndr* 2003; **34**:76–83.
56. Farzadegan H, Hoover DR, Astemborski J, Lyles CM, Margolick JB, Markham RB, *et al*. **Sex differences in HIV-1 viral load and progression to AIDS.** *Lancet* 1998; **352**:1510–1514.
57. Ronald PJM, Robertson JR, Elton RA. **Continued drug use and other cofactors for progression to AIDS among injecting drug users.** *AIDS* 1994; **8**:339–344.
58. Phillips AN, Lee CA, Elford J, Webster A, Janossy G, Timms A, *et al*. **More rapid progression to AIDS in older HIV-infected people: the role of CD4+ T-cell counts.** *J Acquir Immune Defic Syndr* 1991; **4**:970–975.
59. Yan P, Zhang F, Wand H. **Using HIV diagnostic data to estimate HIV incidence: method and simulation.** *Stat Commun Infect Dis* 2011:3.
60. Wand H, Wilson D, Yan P, Gonnermann A, McDonald A, Kaldor J, Law M. **Characterizing trends in HIV infection among men who have sex with men in Australia by birth cohorts: results from a modified back-projection method.** *J Int AIDS Soc* 2009; **12**:19.
61. Wand H, Yan P, Wilson D, McDonald A, Middleton M, Kaldor J, Law M. **Increasing HIV transmission through male homosexual and heterosexual contact in Australia: results from an extended back-projection approach.** *HIV Med* 2010; **11**:395–403.
62. Mallitt K-A, Wilson DP, McDonald A, Wand H. **HIV incidence trends vary between jurisdictions in Australia: an extended back-projection analysis of men who have sex with men.** *Sex Health* 2012; **9**:138–143.