

# Methods to estimate the number of people living with undiagnosed HIV

## ? Questions

- How do we know how many people are living with undiagnosed HIV infection?
- What population groups make up people living with undiagnosed HIV infection?

## 🔑 Key Take-Home Messages

- At the end of 2016, the Public Health Agency of Canada estimated that 14% of 63,110 individuals living with HIV in Canada were undiagnosed (1).
- Different types of data can be used across a variety of methodologies to estimate the undiagnosed proportion of people living with HIV (2, 3).
- The proportion of individuals living with HIV in Canada who were undiagnosed in 2016 was estimated using a complex statistical method known as “back-calculating”, which requires reliable data on HIV diagnoses and deaths from previous years (4).
- Of the 21% of people living with HIV in Canada in 2014 who were undiagnosed, approximately 18% were attributed to the men who have sex with men exposure category, 20% to injection drug use, and 28% to heterosexual contact (5).

## ! The Issue and Why It's Important

In Canada, the estimated prevalence rate of HIV at the end of 2016 was 173 per 100,000 population (1). This prevalence estimate refers to the number of individuals living with HIV who are diagnosed *and* undiagnosed at a given time (1, 6). In addition to providing HIV prevalence data, the Public Health Agency of Canada (PHAC) also provides estimates for the undiagnosed proportion, which refers to people living with HIV who have not been diagnosed and are therefore *unaware* of their HIV-positive status (5). People living with HIV may be undiagnosed due to the nature of the virus, as symptoms may not appear at the time of infection (2, 7). Undiagnosed HIV infection is problematic as individuals who are unaware that they are HIV-positive cannot benefit from highly effective treatment and may unknowingly transmit the virus (1, 8). At the end of 2016, PHAC estimated that 14% (one in seven) of individuals living with

## References

1. Public Health Agency of Canada. Summary: Estimates of HIV incidence, prevalence and Canada's progression on meeting the 90-90-90 HIV targets, 2016. 2018. Available from: <https://www.canada.ca/en/public-health/services/publications/diseases-conditions/summary-estimates-hiv-incidence-prevalence-canadas-progress-90-90-90.html#m> Accessed December 12, 2018.
2. 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(8):1017-23.
3. Medland NA, McMahon JH, Chow EP, Elliott JH, Hoy JF, Fairley CK. The HIV care cascade: A systematic review of data sources, methodology and comparability. *Journal of the International AIDS Society*. 2015;18(1):20634.
4. Lodwick RK, Nakagawa F, van Sighem A, Sabin CA, Phillips AN. Use of surveillance data on HIV diagnoses with HIV-related symptoms to estimate the number of people living with undiagnosed HIV in need of antiretroviral therapy. *PLoS ONE*. 2015;10(3):e0121992.

HIV in Canada were undiagnosed (1).

Figures reflecting the prevalence of HIV and the undiagnosed proportion are imprecise; however, there are several benefits to producing updated estimates (2). Prevalence data are crucial to building the first two stages of the HIV continuum of care (9), a tool used to model the number of individuals: living with HIV, diagnosed with HIV, linked to care, retained in care, being on antiretroviral treatment, and being virally suppressed (10). Prevalence estimates also allow for Canada and other nations to measure progress on the 90-90-90 global HIV targets established by the United Nations (9), where the first “90” is a target referring to the percent of people living with HIV who will know their status by 2020 (11). At the end of 2016, PHAC estimated that 86% of all people in Canada living with HIV were aware of their status (1). Additionally, estimating the burden of HIV is important as this number informs resource allocation such as treatment needs (2, 4, 9) and intervention approaches, such as screening strategies (9) and testing practices (2).

Evidently, estimating the prevalence of HIV and the proportion of individuals who remain undiagnosed is beneficial. However, this number may be the most challenging stage of the care continuum to estimate (12). Indeed, one might ask how the undiagnosed proportion can be estimated when the individuals living with HIV *themselves* are unaware of their HIV status. This review attempts to answer this question by providing a general overview of methodologies used by epidemiologists, statisticians, and other health professionals to calculate estimates of people living with HIV who are undiagnosed. Various methods are described and estimates of the undiagnosed proportion in Canada are provided.

## What We Found

### Methodologies for estimating the number of people living with HIV

There are several methods used to estimate the number of people living with HIV in a given jurisdiction (2). One article reviewing the data sources, methodology, and comparability of HIV care continuums found that cascades that reported undiagnosed HIV-infection (n=7) used different sets of data and employed methods that were not comparable (3).

The Working Group on Estimation of HIV Prevalence in Europe describes three main approaches for estimating the number of undiagnosed individuals living with HIV: methods that are based on calculating cumulative incidence of HIV, methods based on CD4 cell count, and methods based on prevalence surveys (2). The first approach (i.e. methods based on reporting of HIV/AIDS diagnoses involving calculation of cumulative incidence of HIV) is described in more detail as this is the approach PHAC used to estimate figures

5. Public Health Agency of Canada. Summary: Estimates of HIV incidence, prevalence and proportion undiagnosed in Canada, 2014. 2016. Available from: <https://www.canada.ca/en/public-health/services/publications/diseases-conditions/summary-estimates-hiv-incidence-prevalence-proportion-undiagnosed-canada-2014.html> Accessed December 11, 2018.

6. Johnson AS, Song R, Hall HI. Estimated HIV incidence, prevalence, and undiagnosed infections in US states and Washington, DC, 2010–2014. *Journal of Acquired Immune Deficiency Syndromes*. 2017;76(2):116–22.

7. Centers for Disease Control and Prevention. About HIV/AIDS: What are the stages of HIV? 2018. Available from: <https://www.cdc.gov/hiv/basics/whatishiv.html> Accessed December 12, 2018.

8. Pharris A, Quinten C, Noori T, Amato-Gauci AJ, van Sighem A, Surveillance EHA, et al. Estimating HIV incidence and number of undiagnosed individuals living with HIV in the European Union/European Economic Area, 2015. *Euro Surveillance: Bulletin European sur les Maladies Transmissibles = European Communicable Disease Bulletin*. 2016;21(48):1–4.

9. Nunez O, Hernando V, Diaz A. Estimating the number of people living with HIV and the undiagnosed fraction in Spain in 2013. *AIDS*. 2018;32(17):2573–81.

10. Gardner EM, McLees MP, Steiner JF, Del Rio C, Burman WJ. The spectrum of engagement in HIV care and its relevance to test-and-treat strategies for prevention of HIV infection. *Clinical Infectious Diseases*. 2011;52(6):793–800.

for Canada in 2016 (1, 13). Note that there does not appear to be a “gold standard” for calculating estimates of the undiagnosed proportion; each method has benefits and drawbacks (2, 14), and various assumptions are made to arrive at each estimate (2).

### **Methods based on cumulative incidence of HIV**

This approach is based on a method known as back-calculation, a statistically complex method that requires reliable data on past HIV diagnoses and deaths (4). The method is based on the concept that new diagnoses are from infections that happened in the past (15). While infection and diagnosis do not occur simultaneously, some individuals are diagnosed sooner than others. Back-calculation therefore assumes that *new* diagnoses of HIV or AIDS can be used to determine *historical* incidence. The back-calculation method has several mathematical variations (2). As mentioned previously, one review of HIV care cascade methodologies found that cascades including estimates of undiagnosed HIV (n=7) each utilized a different variation of the back-calculation method (3).

Back-calculation was initially used at the beginning of the HIV epidemic; AIDS incidence data was used to “back-calculate” the number of individuals previously infected (16, 17). During this time, effective treatment was not available; the virus took its course, and over time, progressed to AIDS (18). By using the number of reported AIDS cases from each year *and* the assumed length of time from HIV infection to AIDS (i.e. the incubation period), estimating the number of individuals who acquired the virus in each previous year was possible (19). Furthermore, if accurate data on the aggregate number of HIV-related deaths was available, estimating the prevalence of HIV by using the number diagnosed with HIV made it possible to estimate the number of undiagnosed individuals (2).

However, with the advent of antiretroviral therapy, life expectancy among people living with HIV has increased (20). Today, it is estimated that individuals living with HIV on antiretroviral therapy in the U.S. and Canada have a life expectancy similar to that of the general population (21). As a result, back-calculation based on new AIDS cases could no longer produce reliable estimates, as the incubation period became more challenging to calculate (2, 19). Therefore, using this method with current AIDS diagnostic data would likely not produce accurate estimates (19).

In the early 2000s, authors began to incorporate the number of HIV diagnoses into back-calculation models to improve precision (22, 23). Today, back-calculating models utilize HIV diagnoses in addition to other types of data (e.g. biological assays) to estimate the undiagnosed proportion, and there are several variations (2). These include the Cambridge method, the Atlanta method, the Ottawa/Sydney method, the Paris method, and the Bordeaux method (2, 24). A more recent, simplified back-calculation method from the University of Washington relies on a testing history model (12).

11. UNAIDS. 90-90-90: An ambitious treatment target to help end the AIDS epidemic. 2017. Available from: <http://www.unaids.org/en/resources/documents/2017/90-90-90> Accessed December 11, 2018.

12. Fellows IE, Morris M, Birnbaum JK, Dombrowski JC, Buskin S, Bennett A, et al. A new method for estimating the number of undiagnosed HIV infected based on HIV testing history, with an application to men who have sex with men in Seattle/King County, WA. *PLoS ONE*. 2015;10(7):e0129551.

13. Yan P, Zhang F, Wand H. Using HIV diagnostic data to estimate HIV incidence: Method and simulation. *Statistical Communications in Infectious Diseases*. 2011;3(1).

14. Mallitt KA, Wilson DP, McDonald A, Wand H. Is back-projection methodology still relevant for estimating HIV incidence from national surveillance data? *The Open AIDS Journal*. 2012;6(Suppl 1: M7):108–11.

15. Morris M, Fellows I, Golden M. Estimating the undiagnosed fractions: A new ‘testing history’ approach. 2013. Available from: <https://depts.washington.edu/cfar/sites/default/files/uploads/core-program/user70/Estimating%20the%20Undiagnosed%20Fraction.pdf> Accessed December 18, 2018.

16. Brookmeyer R, Gail MH. A method for obtaining short-term projections and lower bounds on the size of the AIDS epidemic. *Journal of the American Statistical Association*. 1988;83(402):301–8.

17. Brookmeyer R, Gail M. Minimum size of the acquired immunodeficiency syndrome (AIDS) epidemic in the United States. *The Lancet*. 1986;328(8519):1320–2.

The Ottawa/Sydney method (also referred to in the literature as extended back-calculation or modified back-projection) is a mathematical modelling technique that utilizes national HIV/AIDS surveillance data to produce estimates (1). These kinds of models simulate real-life circumstances by using equations to predict future trends (25). PHAC used statistical modelling to estimate HIV incidence, prevalence, and the 90-90-90 measures for Canada for 2016 (1).

This modified back-calculation method requires two kinds of data: HIV and AIDS diagnostic data, and the proportion of recent infections among newly diagnosed individuals (1). As surveillance data can only record the date of diagnosis and not the date of infection, a mathematical model is used to estimate the trend of past infections over time until the present (1). Using this trend of past HIV infections, the model projects *forward*, calculating the expected number of HIV diagnoses (1).

Once the trend of the number of infections over time has been estimated, HIV incidence estimates for all years are added together to calculate the cumulative HIV incidence (1). Prevalence for the most recent year is calculated by subtracting the estimated total mortality of people living with HIV from the cumulative incidence (1). Finally, the number of undiagnosed infections among people living with HIV is calculated by subtracting the number of diagnosed individuals living with HIV from the prevalence (1). To estimate the first measure in the 90-90-90 targets, the number of diagnosed people living with HIV is divided by the prevalence (1).

An in-depth explanation of the mathematical formulations, data assumptions, and algorithms this technique employs can be found by reviewing a 2009 publication by Yan et al., *Using HIV diagnostic data to estimate HIV incidence: Method and simulation* (13).

### **Methods based on CD4 cell count**

CD4 cells are white blood cells that are targeted by HIV, and subsequently destroyed (26). Therefore, the CD4 count (measured in cells per cubic millilitre of blood, cells/mm<sup>3</sup>) is a useful indicator for people living with HIV; it reveals the health of the immune system and provides information about disease progression (27). If someone is living with HIV and does not take HIV treatment, their CD4 count will decrease over time, and their risk for developing illness will increase (27).

The underlying concept in the CD4 count approach is that individuals living with undiagnosed HIV will eventually develop HIV-related symptoms, present for care, and be diagnosed with HIV (4). If the CD4 count is recorded at diagnosis, it is possible to approximate when they were infected with HIV based on CD4 depletion data collected from cohort studies.

18. U.S. Department of Health and Human Services. The stages of HIV infection. 2018. Available from: <https://aidsinfo.nih.gov/understanding-hiv-aids/fact-sheets/19/46/the-stages-of-hiv-infection> Accessed December 12, 2018.

19. Wand H, Yan P, Wilson D, McDonald A, Middleton M, Kaldor J, et al. Increasing HIV transmission through male homosexual and heterosexual contact in Australia: Results from an extended back-projection approach. *HIV Medicine*. 2010;11(6):395–403.

20. Antiretroviral Therapy Cohort Collaboration. Life expectancy of individuals on combination antiretroviral therapy in high-income countries: A collaborative analysis of 14 cohort studies. *The Lancet*. 2008;372(9635):293–9.

21. Samji H, Cescon A, Hogg RS, Modur SP, Althoff KN, Buchacz K, et al. Closing the gap: Increases in life expectancy among treated HIV-positive individuals in the United States and Canada. *PLoS ONE*. 2013;8(12):e81355.

22. Cui J, Becker NG. Estimating HIV incidence using dates of both HIV and AIDS diagnoses. *Statistics in Medicine*. 2000;19(9):1165–77.

23. Becker NG, Lewis JJ, Li Z, McDonald A. Age-specific back-projection of HIV diagnosis data. *Statistics in Medicine*. 2003;22(13):2177–90.

24. Hague C. Estimate of the number of persons living with HIV in Massachusetts. 2016. Available from: <https://targethiv.org/sites/default/files/supporting-files/6767Hague.pdf> Accessed December 12, 2018.

The publication from the Working Group on Estimation of HIV Prevalence in Europe outlines two methods that use CD4 cell counts to estimate the undiagnosed proportion – London method 1 and London method 2 (2). However, since the publication of that review in 2011, it appears that other novel methods using CD4 cell counts have emerged that also use CD4 depletion data to calculate the undiagnosed proportion (28, 29). The commonalities among all these methods is the use of CD4 count data obtained from cohorts of people living with HIV, such as CASCADE, a cohort that includes more than 21,000 people living with HIV drawn from over 300 clinics from Europe, Canada, Australia, and Africa (30).

One study that employs a variation of the London method 1 estimates the undiagnosed proportion (4). The authors elaborate on the logic and assumptions of this method in detail. The method requires surveillance data on numbers of new HIV diagnoses with HIV-related symptoms, and the CD4 count at diagnosis. The CD4 count-specific rate at which HIV-related symptoms develop are estimated from cohort data, and 95% confidence intervals can be constructed using a simple simulation method (4). The method is straightforward to implement within a short period once a surveillance system of all new HIV diagnoses, which collects data on HIV-related symptoms at diagnosis, is in place. It is most suitable for estimating the number of undiagnosed people with CD4 count <200 cells/mm<sup>3</sup> due to the low rate of developing HIV-related symptoms at higher CD4 counts (4).

### **Methods based on prevalence surveys**

There are two methods used in this approach – the direct method, and the multiparameter evidence synthesis method (MPES) (2). First used in 1994 (31), the direct method estimates prevalence of HIV by risk category (e.g. injection drug use) and the estimated population size of the risk category (2, 32). Estimates across all risk categories are then added together to produce the number of people living with HIV (2). Currently, UNAIDS uses software called Spectrum (33), a collection of models used to project future trends in HIV and AIDS based on the direct method (2). The second approach, MPES, is similar, but it synthesizes diverse types of data and performs statistical triangulation in order to ensure that estimates are consistent (34).

### **Estimates by exposure category: Canada**

As previously stated, PHAC estimates that 14% of 63,110 individuals living with HIV in Canada were undiagnosed at the end of 2016 (1). However, the publication containing these statistics does not describe the exposure categories that contributed to the proportion of undiagnosed individuals.

A similar document using data from the end of 2014 describes the undiagnosed proportion in three exposure categories: men who

25. Simon Fraser University. Chapter 1: What is mathematical modeling? Available from: <https://www.sfu.ca/~vdabbagh/Chap1-modeling.pdf> Accessed December 13, 2018.
26. U.S. Department of Veteran Affairs. CD4 count (or T-cell count). 2018. Available from: <https://www.hiv.va.gov/patient/diagnosis/labs-CD4-count.asp> Accessed January 3, 2019.
27. Hughson G. NAM: AIDSMAP. Fact-sheet: CD4 cell counts. 2017. Available from: <http://www.aidsmap.com/CD4-cell-counts/page/1044596/> Accessed December 20, 2018.
28. Hall HI, Song R, Szwarcwald CL, Green T. Brief report: Time from infection with the human immunodeficiency virus to diagnosis, United States. *Journal of Acquired Immune Deficiency Syndromes*. 2015;69(2):248–51.
29. Song R, Hall HI, Green TA, Szwarcwald CL, Pantazis N. Using CD4 data to estimate HIV incidence, prevalence, and percent of undiagnosed infections in the United States. *Journal of Acquired Immune Deficiency Syndromes*. 2017;74(1):3–9.
30. EuroCoord. CASCADE: Concerted Action on SeroConversion to AIDS and Death in Europe. Available from: [http://www.eurocoord.net/partners/founding\\_networks/cascade/](http://www.eurocoord.net/partners/founding_networks/cascade/) Accessed January 3, 2019.
31. McGarrigle CA, Cliffe S, Copas AJ, Mercer CH, DeAngelis D, Fenton KA, et al. Estimating adult HIV prevalence in the UK in 2003: The direct method of estimation. *Sexually Transmitted Infections*. 2006;82(Suppl 3):iii78–86.

have sex with men, injection drug use, and heterosexual contact (5). At the end of 2014, 21% of people living with HIV in Canada were undiagnosed; of this, approximately 18% of infections among undiagnosed individuals were attributed to the men who have sex with men exposure category, 20% to injection drug use, and 28% to heterosexual contact (5).

## Factors That May Impact Local Applicability

This review is a general presentation of the different methods used when calculating the proportion of undiagnosed among people living with HIV. It is not intended to be a comprehensive overview of assumptions, models, and equations. Additionally, it is important to consider if the quality of the available data is appropriate for a chosen method, as some methods require robust data sets. Therefore, the methods presented may not be suitable for jurisdictions that do not collect pertinent HIV surveillance data.

## What We Did

We searched Medline (including Epub Ahead of Print, In-Process & Other Non-Indexed Citations) using a combination of text terms (HIV and estimate\* and undiagnosed). All searches were conducted on November 13, 2018 and results limited to English articles without publication date restrictions. Reference lists of identified articles were also searched. The search yielded 314 references from which 34 were included.

32. Petruckevitch A, Nicoll A, Johnson A, Bennett D. Direct estimates of prevalent HIV infection in adults in England and Wales for 1991 and 1993: An improved method. *Sexually Transmitted Infections*. 1997;73(5):348–54.

33. Avenir Health. Spectrum models. Available from: <https://www.avenirhealth.org/software-spectrummodels.php> Accessed December 17, 2018.

34. Presanis AM, Gill ON, Chadborn TR, Hill C, Hope V, Logan L, et al. Insights into the rise in HIV infections, 2001 to 2008: A Bayesian synthesis of prevalence evidence. *AIDS*. 2010;24(18):2849–58.

### Rapid Response: Evidence into Action

The OHTN Rapid Response Service offers quick access to research evidence to help inform decision making, service delivery and advocacy. In response to a question from the field, the Rapid Response Team reviews the scientific and grey literature, consults with experts, and prepares a brief fact sheet summarizing the current evidence and its implications for policy and practice.

#### Suggested Citation

Rapid Response Service. Methods to estimate the number of people living with undiagnosed HIV. Toronto, ON: Ontario HIV Treatment Network; January 2019.

#### Prepared by

Danielle Giliauskas

#### Program Leads / Editors

David Gogolishvili



THE ONTARIO HIV  
TREATMENT NETWORK

#### Contact

[rapidresponse@ohntn.on.ca](mailto:rapidresponse@ohntn.on.ca)

The Ontario HIV Treatment Network  
1300 Yonge Street, Suite 600  
Toronto ON M4T 1X3

#### For more information visit

[www.ohntn.on.ca/rapid-response-service](http://www.ohntn.on.ca/rapid-response-service)

[www.ohntn.on.ca](http://www.ohntn.on.ca)