An epidemiological modelling study to estimate the composition of HIV-positive populations including migrants from endemic settings: an application in the UK

Fumiyo Nakagawa on behalf of the Writing Group on HIV Epidemiologic Estimates in Countries With Migrant Populations From High Prevalence Areas*

*listed in acknowledgements

Corresponding author: Fumiyo Nakagawa, Research Department of Infection and Population Health, University College London (UCL), Royal Free Hospital, Rowland Hill Street, London NW3 2PF, UK. fnakagawa@ucl.ac.uk

Conflicts of interest: none declared.
Abstract

Objective: Migrants account for a significant number of people living with HIV in Europe, and it is important to fully consider this population in national estimates. Using a novel approach with the UK as an example, we present key public health measures of the HIV epidemic, taking into account both in-country infections and infections likely to have been acquired abroad.

Design: Mathematical model calibrated to extensive data sources.

Methods: An individual-based stochastic simulation model is used to calibrate to routinely collected surveillance data in the UK. Data on number of new HIV diagnoses, number of deaths, CD4 count at diagnosis, as well as time of arrival into the UK for migrants and the annual number of people receiving care were used.

Results: An estimated 106,400 (90% plausibility range: 88,700-124,600) people were living with HIV in the UK in 2013. 23% of these people, 24,600 (15,000-36,200) were estimated to be undiagnosed; this number has remained stable over the last decade. An estimated 32% of the total undiagnosed population have CD4 count <350 cells/mm$^3$ in 2013. 25% and 23% of black African male and female heterosexuals living with HIV were undiagnosed respectively.

Conclusions: We have shown a working example to characterize the HIV population in a European context which incorporates migrants from countries with generalised epidemics. Despite all aspects of HIV care being free and widely available to anyone in need in the UK, there is still a substantial number of people who are not yet diagnosed and thus not in care.

Key words: HIV; Mathematical modelling; epidemiology; UK; undiagnosed infections
Introduction

It is imperative for countries to have a good understanding of the populations most at risk of and affected by HIV infection, so as to inform and evaluate the public health response to the epidemic, including the targeting of prevention efforts and to adequately fund treatment and health care services. Monitoring and surveillance is vital to achieve this, and in combination with modelling techniques, data generated can be used to estimate the size of the undiagnosed and diagnosed population as well as HIV incidence[1-5]. One major limitation of much modelling work however, has been assumption that HIV was acquired within the country. This is one of the reasons why estimates of undiagnosed infections and incidence have focused predominantly on men-who-have-sex-with-men (MSM)[6-8]. In the UK, where all HIV services are provided free at the point of need, much research estimating the extent to which HIV affects the population has also focussed predominantly on MSM[4, 5]. The number of new diagnoses in MSM continues to rise[6], and is believed to be a result of a rise in HIV incidence over time as well as improved testing[5, 6, 9]. In contrast, the number of diagnoses of HIV acquired through heterosexual contact has gradually declined, mainly due to fewer diagnoses among people born in countries with generalised HIV epidemics, specifically sub-Saharan Africa[10]. However, this group still constitutes the second largest proportion of people living with HIV (PLHIV) in the UK. Furthermore, there is some evidence that the proportion of migrants with HIV who have acquired HIV within the UK has increased[11-13].

Most estimates of key measures to describe HIV-positive populations in the UK are derived using, but not limited to, case-based surveillance of new diagnoses, CD4 counts at time of diagnosis and the number of deaths in diagnosed people[1]. We have previously reported on a method of estimation, demonstrated using data on the HIV-positive MSM population in the UK[14]. Here, we build on this previous work and extend it to map out the national epidemic
whilst taking into account migration of people from sub-Saharan Africa, which is rarely considered explicitly in other back-calculation and modelling studies[6-8]. The main aim of this work is to demonstrate a novel approach using a mathematical model with the UK as an example, to present key public health outputs on the national level, including the size of the undiagnosed population and HIV incidence for PLHIV.

Methods

The model

The HIV Synthesis progression model is an individual-based stochastic simulation model of HIV progression and the effect of antiretroviral therapy (ART). Model details are available elsewhere[15-17]. In brief, the model aims to simulate the course of HIV infection by following them over time in three monthly intervals. Model assumptions and variables are based on data from observational cohort studies and clinical trials. Variables modelled include the individual’s CD4 count, viral load, use of specific antiretroviral drugs, treatment interruptions, presence of resistance mutations and adherence levels. The risk of loss to follow-up, AIDS events, AIDS deaths and non-AIDS deaths are also modelled.

The model is used to simulate details for all individuals who have lived with HIV in the UK at any time from the beginning of the epidemic (assumed to be 1980) until 2015. For the purposes of calibrating to data from the UK, individuals are categorised into risk groups as MSM, black African heterosexuals, and all other individuals. This latter category includes all other heterosexual individuals, people who inject drugs (PWID) and other less common routes of HIV acquisition. Risk groups are not categorised further as the main aim of this study is to produce estimates for the PLHIV population as a whole.

The model simulates longitudinal data on people from the time of infection. For risk groups other than black Africans, for simplicity, individuals are assumed to have lived in the UK at
the time of infection. We also assume that some black Africans living with HIV in the UK were likewise infected after migration to the UK. In addition, in order to capture the effects of migration to the UK of people already living with HIV from sub Saharan Africa, we simulate a “pool” of HIV-positive people born and living in sub-Saharan Africa (Figure 1). From this population, we assume a rate of migration. This rate of arrival into the UK from the pool of people who acquired HIV in sub-Saharan Africa is assumed to have the same trend as observed migration trends seen for the general population[18]. Additionally, we assume that the rate of arrival would be reduced to 0 (i.e. an individual will not migrate to the UK) in the presence of any pre-AIDS symptoms or AIDS-defining conditions. In order to simulate the population pool of people living with HIV in sub-Saharan Africa we use the HIV incidence curve from sub-Saharan Africa based on UNAIDS estimates generated using Spectrum/EPP[2, 3, 19].

Whilst HIV testing in the UK was possible from 1984 onwards[20], the possibility of diagnosis for people whilst in sub-Saharan Africa is assumed only to be available from 1996 onwards (when testing in antenatal clinics started), with the probability increasing with time. ART is assumed to be available for people in sub-Saharan Africa from 2000 (with scale-up from 2003 onwards) as per WHO guidelines at the time.

Calibration to data

The model is calibrated using Approximate Bayesian Computation (ABC) methods[21], which involves running the simulation model multiple times, each time sampling values for a set of input parameters from their respective prior distributions. ABC methods are suitable for complex simulation models where the likelihood of the outputs cannot be exactly calculated. It also allows us to explore a wide parameter space and to consider multiple sets of parameter values which are consistent with the observed data, instead of converging to a single set of parameter values.
For any single simulation, we define the ‘calibration-score’ to be the weighted average of the deviances of the modelled outputs from the observed data. Therefore, the lower the calibration-score, the better the model fits to the data. The calibration process is continued until at least 50 accepted sets of parameter values are within the pre-defined tolerance threshold. We chose to simulate at least 50 parameter sets due to the trade-off between precision and run time. Further simulations are then performed using the accepted parameter sets, to reduce stochastic variability (i.e. variability in model outputs between simulation runs when using the same parameter set). We use the 50 best fitting simulations generated in this step for our presented results; the median, 5th and 95th percentiles of the distribution of these model outputs are considered the point estimate and plausibility range (PR) limits, respectively. The main parameter values to be estimated are those associated with the incidence (in the UK) and rate of diagnosis. All risk groups are jointly modelled and calibrated. Further details of the calibration method are described in [14] and in the Supplementary material, http://links.lww.com/QAD/B16. The model and calibration method are programmed in SAS software, Version 9.3 (SAS Institute Inc., Cary, NC).

Data from the HIV and AIDS reporting system (HARS), which is the national HIV surveillance system developed and managed by Public Health England (PHE), were used to calibrate the model. We used case-based reports on the number of new HIV diagnoses, number of deaths and CD4 count at diagnosis. Data collected on those diagnosed and accessing care, such as the number seen in care and number receiving ART, and time of arrival for sub-Saharan African migrants were also used. The observed data were available until 2013. Weights for each type of data used to calculate the calibration-score are chosen a priori to reflect perceived relative quality and completeness. Although HIV surveillance in the UK is based on voluntary reports from clinicians, virologists, public health practitioners amongst others, it is thought to be one of the most comprehensive in the world due to the
diverse range and methods by which information is collected and its direct link to the commissioning of HIV services.

**Results**

A range of model outputs from the simulations using the 50 accepted parameter sets alongside the surveillance data are presented in the Supplementary Materials, http://links.lww.com/QAD/B16. The following results are based on these parameter sets which are calibrated to the observed data in the UK.

**Principal findings**

The incidence of HIV, defined as the number of new infections acquired in the UK per year, is estimated to have remained at high levels in recent years (Figure 2). After the initial peak in incidence in 1980-81 with an average of 5,400 new infections per year, it is estimated to have steadily declined to an average of 1,600 infections in the early 1990s. The incidence is estimated to have risen to over 4,500 new infections per year since the 2000s. Between 2010 and 2013, an estimated 4,700 (90% PR: 2,000-9,800) new infections occurred annually. This wide PR arises because there is a time lag between individuals acquiring HIV to then being diagnosed and reported into the surveillance system, meaning there is less data to inform the incidence for later years.

By 2013, 106,400 (90% PR: 88,700-124,600) people were estimated to be living with HIV in the UK (Figure 3). As expected, there has been a sharper rise in the number of PLHIV in the post-combination therapy era due to the high effectiveness of ART in reducing death rates and non-zero incidence.

The estimated total number of people living with undiagnosed HIV in 2013 was 24,600 (90% PR: 15,000-36,200) (Figure 3). This figure has remained stable during the last decade. Of the 24,600 undiagnosed population living in the UK, 32% (8,000; 90% PR: 4,500-12,100) are
estimated to have CD4 count <350 cells/mm³ and are therefore at higher risk of progression to AIDS or clinical manifestation. We also estimate that 95% of the undiagnosed population have a viral load >500 copies/ml.

The estimated HIV cascade of care for the UK in 2013 according to our model is shown in Figure 4. Using our estimate of the total number of people living with HIV (100%), 77% are diagnosed, 76% are retained in care, 68% are receiving ART and 62% have viral load <50 copies/ml.

*Estimates by risk group*

Of the 4,700 new infections which took place in the UK each year between 2010 and 2014, we estimate that 2,500 (90% PR: 900-5,800) were attributed to MSM. We also estimate that 1,200 (90% PR: 800-2,300) new infections were acquired annually in the UK in the same period among black African heterosexual individuals.

The rate of diagnosis, defined as the probability of being diagnosed in the UK per year (given the person is HIV-positive and asymptomatic), is estimated to have increased over time to 18.9% and 15.3% for MSM and heterosexual individuals (regardless of ethnicity or sex) for 2008 onwards (it is assumed constant over eight year periods).

As expected, PLHIV population size has increased in all risk groups (Figure 5), though the rise seems most substantial among MSM. In 2013, MSM comprised the largest share of the HIV-positive population in the UK, approximated at 45,000 (90% PR: 30,100-53,800). The next largest groups are black African females and black African males, with an estimated 21,900 (90% PR: 17,700-27,900) and 15,400 (90% PR: 13,700-19,100) individuals respectively. Reduced migration from sub-Saharan Africa is believed to be the reason for a reduced rise in the total number of black African heterosexual individuals living with HIV and for the decline in the number living with undiagnosed HIV.
The estimated size of the undiagnosed proportion in 2013 ranges from 19% (90% PR: 9%-28%) among MSM, to 26% (90% PR: 23%-30%) and 25% (90% PR: 22%-26%) in black African male and female heterosexuals and 23% (90% PR: 17%-34%) and 20% (90% PR: 15%-33%) in all other males and females respectively. The proportion of undiagnosed people with a CD4 count <350 cells/mm$^3$ varies greatly between risk groups, with an estimated 27% among MSM but a substantially higher 46% for black African heterosexuals.

**Discussion**

Our modelling study shows that the total number of PLHIV in the UK in 2013 surpassed 100,000, with around a fifth to a quarter of this population being unaware of their infection. The model we have used has been calibrated to a wide range of UK surveillance data, including data on black African heterosexuals, in order to provide a more detailed picture of the national epidemic than has been done before.

There is a distinct possibility that the incidence of new infections taking place in recent years is even greater than the high levels seen at the beginning of the epidemic, although there is large uncertainty associated with it. We estimated that there were nearly 5,000 new cases of HIV occurring in the UK in 2013. This is despite the fact that nearly two-thirds of the PLHIV population are thought to be virally suppressed and at low risk of transmitting the virus. Although national and international guidelines now recommend that all PLHIV should initiate ART[22, 23], studies have shown that HIV transmission in the UK is largely driven by the undiagnosed population and that testing rates need to be much higher in order to limit further increases in incidence[5, 24]. Higher testing rates are especially needed among black African individuals, for which we estimate that nearly half of those undiagnosed have CD4 count <350 cells/mm$^3$ and are in immediate need for treatment.
Our results are in line with previous estimates published by PHE, using the Multi-Parameter Evidence Synthesis (MPES) method[25, 26]. Like our model, the MPES method is directly informed by surveillance data, but is combined with other data from prevalence surveys and estimates of risk group sizes to calculate the prevalence of HIV and the size of the undiagnosed population. The MPES estimate for the total number of PLHIV in 2013 was 107,800 (95% credible interval: 101,600-115,800)[10]. The PR generated using our method is considerably wider however, which is likely a result of taking into account the uncertainty associated with parameters which cannot be directly informed by data, as well as the beliefs held with regards to certain items of data (as quantified using different weights in the calibration-score). Our method is able to complement the information generated by the MPES method and provides another approach to estimate the undiagnosed population. The ECDC modelling tool[7, 27] is also a method based on an underlying mathematical model of HIV which can estimate incidence and the total size of the diagnosed and undiagnosed population, however it does not currently take into account migration.

To the best of our knowledge, our study is the first modelling study done to characterize the PLHIV population in a European context which attempts to address and incorporate migration of individuals from endemic settings. Migrants from sub-Saharan Africa account for a significant proportion of cases of HIV in Europe[28, 29], and it is important to fully consider this population in estimates and to try to disentangle the extent to which transmission is ongoing in the European country[30]. In settings where a large number of infections are thought to have occurred abroad, it is important to know how many PLHIV actually acquired their infection in the country of interest, in order to target prevention interventions most effectively. It was used to be considered that most migrants acquired HIV in sub-Saharan Africa, however more recent evidence suggests that in fact many are now acquiring HIV post-migration in European countries[12, 31]. Here we estimated that 1,200
(90% PR: 800-2,300) new infections occurred annually in the UK among black African heterosexuals in 2010-2014. In addition, we estimate that there has been a decline in the number of undiagnosed black African heterosexuals, and that many are likely to be diagnosed late, which in turn means that testing is imperative among this group.

Our estimates are based on our model being calibrated to the date of arrival into the UK among HIV-positive individuals from sub-Saharan Africa. These data are not collected as part of routine case surveillance in all countries at present however this field is now recommended by the European Centre for Disease Prevention and Control[32]. We also acknowledge that few countries have a national cohort of all people living with diagnosed HIV. These data are a critical part of the model calibration.

One limitation with this modelling study was that we did not explicitly model heterosexual individuals who are not black African, PWID, mother-to-child-transmission or other less common routes, but instead grouped them together. This also means that we were not able to make separate estimates of the size of these risk groups. In 2013, PHE figures stated that 4.3% of cases of new HIV diagnoses were acquired through injecting drug use or other routes of transmission (not including sexual transmission) and there are also a small proportion of cases without an exposure category[10]. For our risk group-specific estimates (but not our overall estimates) we also inherently assumed that the available data on route of HIV transmission is correct for each individual. Phylogenetic analyses in the UK have shown that some self-reported heterosexual men diagnosed with HIV could have acquired HIV through sex with other men, choosing not to disclose sex with men at HIV diagnosis and preferring to be identified as heterosexual[33, 34].

Our main aim was to estimate the total size of the PLHIV population and therefore the model is more closely calibrated to un-stratified data. This was done by calculating the calibration-score using higher weights for data relating to the total population compared to risk group-
specific data. Specifically, we used weights which were twice as much because it was important that the model outcomes were more closely calibrated to data on the total UK population. This means our estimates presented for the three defined groups are not as well informed as the overall presented estimates for the UK. For example, the point estimate of 2,500 (90% PR: 900-5,800) new infections among MSM per year for 2010-2014 is lower than the estimate of 3,200 (90% PR: 1,700-5,400) from our previous work where the model was specifically calibrated to data on MSM[14]. Further research is required to model the migrant population more accurately, in order to better estimate incidence and number undiagnosed. Similarly, narrower PRs may well also follow by modelling sub-populations separately by reducing the realm to which the model needs to fit the observed data.

In addition, our method of estimating the UK PLHIV population is not easy to implement, requiring knowledge of mathematical models and being computer intensive. It should also be noted that our PRs are wide as we take into account many sources of error and variability. However, they are perhaps overestimated somewhat due to residual stochastic effects relating to the fact that we model a random sample of the infected population and scale up. Although generally, more data should lead to less uncertainty, there is a limit to how much the PRs can be reduced, given that in this case, what we are trying to estimate is intrinsically uncertain.

Achieving the UNAIDS 90-90-90 target requires more than 73% of all PLHIV with an undetectable viral load[35]. Even in the UK, where HIV services are comprehensive and universally free, we estimate that only 62% of the PLHIV population have suppressed viral load. In a previous modelling study, we showed that 90% of all HIV-positive MSM would need to be suppressed to achieve substantial reductions in incidence[9]. In most settings, including the UK, the largest breakpoint in the cascade is diagnosis, reiterating the importance of testing[36]. It is recognised that late diagnosis, particularly among heterosexuals remains a problem in the UK[10]. There is therefore a need for scaling up
access to HIV testing, and frequency of testing in those at high risk, to reduce late diagnoses and also undiagnosed HIV infection, which is the likely reason for ongoing levels of high incidence.

It is possible with access to good surveillance data to calculate the proportion of the diagnosed population who are on ART and the proportion of the treated population who are virally suppressed. However, it is not possible to know the size of the undiagnosed population without some estimation method such as mathematical models or prevalence surveys. The method presented and used here allows calculation of such information which are vital to know the extent to which targets such as those put in place by UNAIDS are likely to be achieved. It is only feasible to generate reliable estimates of the number of people living with HIV if good surveillance data are available however. Our modelling work is based on extensive national-level surveillance data, and though it would not be impossible to do without all of it, a similarly strong set of data is required for it to be applied in another setting.

In conclusion, we have presented a method to generate national-level estimates about the size and characteristics of HIV-positive populations incorporating migrants from sub-Saharan Africa. These estimates should help inform and evaluate current and future public health responses to the epidemic by providing more information about the affected population than can be understood and achieved solely from surveillance.

Acknowledgements

The authors acknowledge the Public Health England (PHE) HIV surveillance team and the use of the UCL Legion High Performance Computing Facility (Legion@UCL), and associated support services, in the completion of this work.
The research leading to these results has received funding from the European Union Seventh Framework Programme (FP7/2007-2013) under EuroCoord grant agreement n° 260694.

**Writing Group on HIV Epidemiologic Estimates in Countries With Migrant Populations From High Prevalence Areas:**

Fumiyo NAKAGAWA, Valerie DELPECH, Alioum AHMADOU, Jan ALBERT, Daniela BEZEMER, Mario CORTINA-BORJA, Alison BROWN, Valentina CAMBIANO, Dominique COSTAGLIOLA, Andrea DE LUCA, Roger KOUYOS, Rebecca LODWICK, Jens LUNDGREN, Nikos PANTAZIS, Colette SMITH, Virginie SUPERVIE, Pat TOOKEY, Giota TOULOUMI, Zheng YIN, Ard VAN SIGHEM, Andrew N PHILLIPS on behalf of the SSOPHIE project working group in EuroCoord*

* (see below)

SSOPHIE project working group in EuroCoord: Fumiyo Nakagawa, Alioum Ahmadou, Andrew Amato-Gauci, Alison Brown, Jan Albert, Daniela Bezemer, Valentina Cambiano, Colin Campbell, Jordi Casabona, Daniel Commenges, Mario Cortina Borja, Dominique Costagliola, Daniela de Angelis, Valerie Delpech, Andrea de Luca, Martin Donoghoe, Deborah Ford, Geoff Garnett, Peter Ghys, Tim Hallett, Jesper Kjaer, Roger Kouyos, Rebecca Lodwick, Jens Lundgren, Nikos Pantazis, Santi Perez-Hoyos, Patrizio Pezzotti, Anastasia Pharris, Chantal Quinten, Mika Salminen, Colette Smit, Colette Smith, Jonathan Sterne, Virginie Supervie, Rodolphe Thiebaut, Claire Thorne, Pat Tookey, Giota Touloumi, Ard van Sighem, Andrew Phillips.

EuroCoord Executive Board: Fiona Burns, University College London, UK; Geneviève Chêne, University of Bordeaux, France; Dominique Costagliola (Scientific Coordinator), Institut National de la Santé et de la Recherche Médicale, France; Carlo Giaquinto, Fondazione PENTA, Italy; Jesper Grarup, Region Hovedstaden, Denmark; Ole Kirk, Region
Hovedstaden, Denmark; Laurence Meyer, Institut National de la Santé et de la Recherche Médicale, France; Heather Bailey, University College London, UK; Alain Volny Anne, European AIDS Treatment Group, France; Alex Panteleev, St. Petersburg City AIDS Centre, Russian Federation; Andrew Phillips, University College London, UK, Kholoud Porter, University College London, UK; Claire Thorne, University College London, UK.

EuroCoord Council of Partners: Jean-Pierre Aboulker, Institut National de la Santé et de la Recherche Médicale, France; Jan Albert, Karolinska Institute, Sweden; Silvia Asandi, Romanian Angel Appeal Foundation, Romania; Geneviève Chêne, University of Bordeaux, France; Dominique Costagliola (chair), INSERM, France; Antonella d’Arminio Monforte, ICoNA Foundation, Italy; Stéphane De Wit, St. Pierre University Hospital, Belgium; Peter Reiss, Stichting HIV Monitoring, Netherlands; Julia Del Amo, Instituto de Salud Carlos III, Spain; José Gatell, Fundació Privada Clínica per a la Recerca Biomèdica, Spain; Carlo Giaquinto, Fondazione PENTA, Italy; Osamah Hamouda, Robert Koch Institut, Germany; Igor Karpov, University of Minsk, Belarus; Bruno Ledergerber, University of Zurich, Switzerland; Jens Lundgren, Region Hovedstaden, Denmark; Ruslan Malyuta, Perinatal Prevention of AIDS Initiative, Ukraine; Claus Møller, Cadpeople A/S, Denmark; Kholoud Porter, University College London, United Kingdom; Maria Prins, Academic Medical Centre, Netherlands; Aza Rakhmanova, St. Petersburg City AIDS Centre, Russian Federation; Jürgen Rockstroh, University of Bonn, Germany; Magda Rosinska, National Institute of Public Health, National Institute of Hygiene, Poland; Manjinder Sandhu, Genome Research Limited; Claire Thorne, University College London, UK; Giota Touloumi, National and Kapodistrian University of Athens, Greece; Alain Volny Anne, European AIDS Treatment Group, France.

EuroCoord External Advisory Board: David Cooper, University of New South Wales, Australia; Nikos Dedes, Positive Voice, Greece; Kevin Fenton, Public Health England, USA;
David Pizzuti, Gilead Sciences, USA; Marco Vitoria, World Health Organisation, Switzerland.

EuroCoord Secretariat: Silvia Faggion, Fondazione PENTA, Italy; Lorraine Fradette, University College London, UK; Richard Frost, University College London, UK; Andrea Cartier, University College London, UK; Dorthe Raben, Region Hovedstaden, Denmark; Christine Schwimmer, University of Bordeaux, France; Martin Scott, UCL European Research & Innovation Office, UK.
Reference List


Figure 1: The modelled population in HIV Synthesis Progression model, illustrating the composition of people living with HIV in the UK, including those who have migrated from sub-Saharan Africa.
Figure 2: Estimated incidence (number of new HIV infections in a year) in the UK
Figure 3: Estimates of the total number of people living with HIV and total number of people living with undiagnosed HIV in the UK, by calendar year. Columns and error bars: modelled median and 90% plausibility range.
Figure 4: Estimated treatment cascade and population characteristics of all individuals living with HIV in the UK in 2013. Columns and error bars: modelled median and 90% plausibility range.
Figure 5: Estimates of the total number of people living with HIV and living with undiagnosed HIV by calendar year among MSM, black African heterosexual men and women, and all other men and women. Columns and error bars: modelled median and 90% plausibility range.