Extending current approaches to HIV incidence estimation

Back-calculation on the basis of data on the number of new HIV and AIDS diagnoses is a key tool for estimating levels of HIV incidence, diagnosis rates and the prevalence of undiagnosed infection throughout the HIV epidemic. These estimates provide an indicator of the success of interventions designed to mitigate the spread of HIV and also to predict the future burden on healthcare services due to the virus. Accordingly, for a number of years, Bayesian back-calculation methods have been developed at the MRC-BSU, with outputs feeding directly into Public Health England's annual reporting on the state of the HIV epidemic in the United Kingdom (Kirwan et al., 2016).

In brief, data giving numbers diagnosed has been augmented by additional information on the distribution of the CD4 counts of the diagnosed to inform progression into, through and out of a multi-stage model describing the prevalence of undiagnosed infection (and its severity) in the population (Sweeting et al., 2005). In later work, it was demonstrated how to derive posterior distributions of the level of undiagnosed HIV infection (Birrell et al., 2012) (amongst other quantities) and, in 2013, this method was a key component of a high-impact study of the effectiveness of the increased HIV testing rates and antiretroviral distribution (Birrell et al., 2013). Recent work has stratified the analysis to obtain age-specific inferences on incidence, diagnosis rates and undiagnosed prevalence (Brizzi, 2015; Brizzi et al 2017, in preparation).

The weakness of a back-calculation approach is in the uncertainty of the resulting estimates in the most recent years. The time between HIV infection and diagnosis is, on average, around 3 years, so the diagnosis data can only be weakly informative about rates of infection over the last 1-3 years. This is problematic: the success, or otherwise, of public health interventions to disrupt transmission of HIV, such as the widespread administration of pre-exposure prophylaxis (PrEP) cannot easily be assessed until a number of further years have passed.

To address this limitation, the back-calculation model can be further extended to incorporate additional surveillance data from the testing of newly diagnosed individuals using recent infection testing algorithms (RITA). Data from RITA classify new diagnoses into recently infected (on average in the last 6 months) and not recently infected. This will result in more robust estimates of incidence and prevalence of undiagnosed infection over the latest years. Similar data have been shown to achieve this in a different healthcare setting (Yan et al., 2011). Data on RITA test results are available from PHE for the last few years and this availability presents an unmissable opportunity to fully exploit these data, illustrating the value of the RITA programme, and to contribute to current policy decision and evaluation.

Work is currently ongoing to adapt the back-calculation model to incorporate this information in the age-dependent setting, but it is in age-independent estimation where this lack of robustness is most striking and is most in need of additional research. Spanning a similar period, data are also available on the most recent negative HIV test result for some of the newly diagnosed. How these data are most appropriately used is not immediately clear, there remains a concern that for diagnoses where a last negative test is unavailable then this represents informative missingness. Does this mean that they only tested because of the presence of seroconversion illness or is it more likely that a non-regular tester will have a long-held infection? The answer to this question will determine whether or not such data necessitates population-level modelling or modelling at the individual level to account for personal propensities to test.

The aims of this project are:
1. To extend the current age-dependent back-calculation model to incorporate data from RITA testing;
2. To investigate the informativeness of information on the last negative test.
3. Following on from 2, to extend the backcalculation model to include this information as appropriate;
4. To investigate the feasibility of incorporating migration in the backcalculation model.


