

Analysis Plan for Individual-based Model Projections of the HPTN 071 (PopART) Intervention

William Probert, Rafael Sauter, Anne Cori, Michael Pickles, Sian Floyd,
David MacLeod, Ethan A Wilson, Deborah Donnell, Helen Ayles, Nulda
Beyers, Peter Bock, Sarah Fidler, Richard Hayes, Christophe Fraser

Version 1.0

December 14, 2018

This document presents the analysis plan for the mathematical modelling component of the HPTN 071 (PopART) trial. The modelling team of the HPTN 071 (PopART) trial are blinded to all but baseline data from the PC (PC0), in which the primary endpoint will be measured. This provides an opportunity to test the predictive ability of the mathematical model developed as part of the trial and to understand factors contributing to why model predictions were or were not concordant with the reported primary endpoint (cumulative reduction in incidence between arms A or B, and C between PC12 and PC36). This report therefore outlines, and provides rationale for, proposed analyses using the mathematical model before and after the trial unblinding in December 2018.

Contents

1	Background	3
2	Trial unblinding	4
3	Analysis	5
3.1	Model projections of trial impact	6
3.2	Model projections of secondary endpoints	7
3.3	Comparison of projected and observed trial impact	8
3.4	Primary endpoint	8
3.5	Secondary endpoints	9
4	PopART-IBM	9
4.1	Overview	9
4.2	Input parameters	10
5	Inference framework	12

Abbreviations:

- ABC: approximate Bayesian computation
- ART: antiretroviral treatment
- CF: counter-factual scenario
- CHiPs: Community HIV-care providers
- DHS: Demographic and Health Surveys
- HPTN: HIV Prevention Trials Network
- HSRC: Human Sciences Research Council
- IBM: individual-based model
- PC: population cohort
- PC0: population cohort at baseline
- PopART: Population Effects of Antiretroviral Therapy to Reduce HIV Transmission
- RR: rate ratio
- SPVL: set-point viral load
- TasP: treatment as prevention
- TMC: traditional male circumcision
- VMMC: voluntary medical male circumcision
- VS: virally suppressed
- UTT: universal testing and treatment

1 Background

The PopART individual-based model (IBM) is a computer simulation model of the HIV epidemic in the communities of the HPTN 071 (PopART) trial. It models a simulated population of approximately the same size as each community in the HPTN 071 (PopART) trial, modelling every individual within that simulated population. The model has several components that model demography, the introduction of HIV, heterosexual partnership formation and dissolution, HIV progression, and both the PopART intervention and a background care cascade.

A mathematical model was developed before the HPTN071 trial began that was used for sample size calculations, to aid development of the trial protocol, to develop targets for process variables (such as coverage), and quantify the effects of factors such as sexual partnerships with partners living outside the trial communities, to predict the impact of the proposed packages of interventions relative to the standard of care arm, and to provide scenarios for power calculations. During the trial, a stochastic individual-based model was developed and used to address four main objectives:

1. **To help interpret the results of the trial:** Process data showing the extent of uptake of the intervention compared with similar data from the control arm will be used to obtain model projections of expected impact under these conditions. By examining projected impact under the conditions prevailing in Zambia and South Africa, and in different trial communities, we will be able to examine whether the level of impact and variations in impact are in accordance with expectations and to explore possible explanations for any discordance between model projections and empirical data.
2. **To project longer-term impact:** Modelling shows that the full impacts of universal test and treat (UTT) as well as male circumcision are not seen for several years. Impact measured during 3 years of intervention may therefore underestimate the long-term impact of the program. Models fitted to the impact seen during the first 3 years will be used to project the likely impact over longer time periods.
3. **To explore likely impact in different settings:** If the trial demonstrates impact, it is likely that similar interventions will be implemented in a wide range of settings. The model will be used to explore how impact would be expected to vary depending on epidemiological, demographic and other characteristics of populations, and thus to project likely impact in a range of settings.
4. **To explore the likely impact of alternative intervention packages:** Our study design will provide empirical data on the impact of the specific package of preventive interventions incorporated in the PopART programme. However, the model can be used to explore the effect of adding or removing components. For example, we can project the impact of an intervention in which male circumcision is not promoted, or where testing and linkage to care is prioritised for certain subpopulations or population strata.

2 Trial unblinding

The unblinding of the trial provides an opportunity to test the predictive power of mathematical modelling. The PopART modelling team is blinded to all but the baseline data from the Population Cohort (PC), but has access to CHiPs intervention data in all arm A and B communities.

During the course of the trial the community-specific IBM projections were updated with PC baseline data in all communities of the HPTN 071 (PopART) trial and with data collected by CHiPs, who delivered the trial intervention, in all arm A and B communities. The results of the IBM projections were reported to the Data and Safety Monitoring Board (DSMB) in 2016 and 2017. Until the unblinding of the trial, projections of trial impact are only presented as comparisons of arm A or B communities with a self-counterfactual scenario. The counter-factual scenario (CF) was generated by switching off the trial intervention in the IBM.

At the end of the pre-unblinding stage of the HPTN 071 trial the data collection for all PC rounds is complete but trial results are not unblinded (the modelling team is still blinded to all but the PC0 data). This situation allows three important questions to be addressed:

1. What is the predictive ability of the PopART IBM with respect to the primary outcome?
2. What explains the mismatches (if any) between the simulation and the trial outcome?
3. How does knowledge from the trial change predictions of long-term impact of treatment as prevention (TasP), and prioritisation of prevention modalities for the future?

To address these questions, three rounds of IBM projections, the first before unblinding, are to be carried out to predict the primary endpoint of the trial:

Round 1 Pre-unblinding projections: before the 17th of December all input parameters of the IBM will be revisited and an update of the CHiPs data will be used. Model projections will be recalibrated. The primary endpoint of the trial, which is the relative reduction in HIV incidence between PC12 and PC36, will be computed for all seven triplets. The trial impact will be assessed based on community-specific IBM projections which are calibrated to arm A communities compared to a counter-factual scenario (CF). The CF uses the same calibrated parameter combinations but the trial intervention in the model is switched off. Cumulative incidence, overall and stratified by age and gender, will be also be computed for each of the arm A and B communities in the HPTN 071 (PopART) trial.

Post-unblinding projections: after the 17th of December 1) the realised impact of the trial will be reported and 2) the modelling team of the HPTN 071 (PopART) trial will gain access to all the PC data. The release of the realised impact of the trial allows comparison with the model projections. Access to the PC data allows the model to be calibrated to PC data, therefore allowing projections of the epidemic in arm C communities, and allowing model-generated comparisons of arm A (or arm B) to arm C communities instead of to counter-factual scenarios. The post-unblinding projections will allow assessment of the validity of the parameterisation and processes used in the IBM. From

December 2018 until March 2019 the input parameters of the IBM will be revised, including the new information which will be made available through the PC data.

Round 2 Post-unblinding projections, updating uptake from PC: Using parameter values fitted to the pre-unblinding data, community-specific parameters associated with ART uptake, VMMC uptake, and uptake of HIV testing will be updated in the background care cascade, informed from the new PC data (i.e. PC12, PC12N, PC24, PC24N, PC36). That is, PC data as well as the CHiPs data will be used to inform the care cascade. The model will not be recalibrated and all other parameters will be kept as their pre-unblinding estimates. Projections of primary and secondary endpoints in all A and B communities and their CF will be generated, as per the pre-unblinding projections. Cumulative incidence, overall and stratified by age and gender, will be also be computed for each of the arm A and B communities in the HPTN 071 (PopART) trial.

Round 3 Post-unblinding projections, updating uptake from PC, and recalibrating to PC: The new PC data (i.e. PC12, PC12N, PC24, PC24N, PC36), including data on HIV incidence, will be used to re-calibrate model projections. Projections of primary and secondary endpoints in all A and B communities and their CF will be generated, as per the pre-unblinding projections. Community-specific arm C projections will be generated for the first time, in addition to the CF projections, as PC data (beyond PC0) for arm C communities will become available. Cumulative incidence, overall and stratified by age and gender, will be also be computed for each of the arm A, B, and C communities in the HPTN 071 (PopART) trial.

Prior to unblinding, a report will be generated that documents pre-unblinding projections of the trial impact and this report will be lodged with the Data and Safety Monitoring Board (DSMB). Projections of trial impact and secondary endpoints (discussed below) will be included in this report. Included with the report will be detailed model input and output files, relevant code for running the IBM and reproducing the model simulations, and code scripts for processing and generating the pre-unblinding results and report. This procedure guarantees that the modelling team does not incorporate any post-unblinding information, coming from the PC data or other sources, in the pre-unblinding projections, that the modelling team does not manipulate pre-unblinding projections ex post, and that the pre-unblinding projections are fully reproducible.

3 Analysis

The model projections of the trial impact and secondary endpoints will be reported in the pre- and post-unblinding report in the same format. The comparison of projected pre- and post-unblinding and observed trial impact will be carried out after the unblinding of the PC data.

All measures will be reported as median, 2.5% and 97.5% quantiles of the simulations from the IBM which were accepted in the final step of the ABC-algorithm (see inference framework in Section 5). The interval from the 2.5% to 97.5% quantile can be interpreted as 95% credible interval, given

the model, data and applied methods, which reflects the uncertainty with respect to the model parameterisation.

The DSMB reports from 2016 and 2017 used several scenarios that looked at different epidemiological parameters. For the pre-unblinding report, the 'high acute, high contamination' scenario as well as the 'low acute, low contamination' scenario will be reported. The 'high acute, high contamination' scenario is the most conservative scenario but it is most in line with current literature on acute HIV infection (e.g. Bellan et al. (2015) PLOS Medicine), and CHiPs data on what is reported regarding partners outside the community (contamination).

3.1 Model projections of trial impact

Trial impact from the mathematical model will be reported as the relative reduction in cumulative incidence between intervention and control arms, and as relative cumulative incidence. The relative cumulative incidence is a rate ratio (RR) and is in line with what is reported in the primary endpoint in the Statistical Analysis Plan for the HPTN071 trial (version 2.0, 21 November 2017). The relative reduction in cumulative incidence and the RR will be reported as one number for each A and B community compared to counter-factual scenarios in the same communities (the expectation according to the IBM had the PopART intervention not taken place). For the post-unblinding projections the RR and the relative reduction in cumulative incidence will also be reported for each A and B community to the arm C communities. These measures will be reported according to the following time frames: PC0-PC36, PC0-PC12, PC12-PC24, PC24-PC36, PC12-PC36, 2020-2030. These measures will also be reported using three different calculations of cumulative incidence for the community in question:

1. Cumulative incidence in the whole simulated population.
2. Cumulative incidence for the whole simulated population, stratified by age and gender.
3. Cumulative incidence, weighted so as to be in line with a 'PC-like' sample including gender and age distribution.

The re-weighting according to a 'PC-like' population will be done at the country and/or community level if there is substantial evidence for a difference in the gender and/or age distribution at the respective geographical level. The cumulative incidence (Cum.I.) between PC round t_0 and t_1 (reported in 100 person years) for each stratum and population will be computed as

$$\text{Cum.I.}_c^d(t_0, t_1) = \frac{I_c^d}{N_c^d}$$

where I is the number of new, incident infections between t_0 and t_1 . N is the number of observed or simulated person years between t_0 and t_1 . The superscript d indicates from where the data are derived: either the projected pre-unblinding simulation results ($d = \text{pre.}$), the post-unblinding projections ($d = \text{post.}$), or observed PC data ($d = \text{obs.}$). The subscript c indicates whether the cumulative incidence is computed for community A ($c = A$), B ($c = B$), C ($c = C$), or in the simulated

CF scenario ($c = CF$).

The RR is computed as

$$RR_{c,c'}^d(t_0, t_1) = \frac{\text{Cum.I.}_c^d}{\text{Cum.I.}_{c'}^d}$$

where c is either a community in arm A or B, and c' is a community in arm C or a CF scenario. All comparisons will be between communities in the same triplet (that is, c and c' are in the same triplet) according to the HPTN 071 (PopART) trial study protocol.

The projected cumulative incidence for each arm A and B community (and arm C in the post-unblinding projections) will also be reported overall, and stratified by age and gender.

3.2 Model projections of secondary endpoints

Secondary endpoints, as computed from simulations, will also be reported. The following will be reported stratified by age and gender for each community, and at the midpoint of PC12, PC24, and PC36 in two tables, one for the whole population and one for the 'PC-like' population:

- Proportion of HIV positive individual in the population.
- Proportion of HIV positive individuals that are aware of status.
- Proportion of those aware of status that are on ART.
- Proportion of those on ART that are viral suppressed.

Overall levels of viral suppression, stratified by arm, will also be reported. After unblinding, simulations will be generated based on two counterfactual scenarios in order to disentangle the impact of different components of the PopART intervention package on HIV incidence, and their combined effect (if any):

1. Scenario 1, no VMMC: the coverage of VMMC, as offered in the simulated HPTN 071 (PopART) trial, will be set to zero in the IBM simulations.
2. Scenario 2, no ART: the coverage of ART, as offered in the simulated HPTN 071 (PopART) trial, will be set to zero in the IBM simulation.

Several additional analyses will be performed after the unblinding of the trial. Firstly, future projections of three scenarios will be generated: 1) discontinuation of the CHiPS intervention and reversion to national standard of care, 2) roll-out of a PopART-like intervention on a national scale, 3) roll-out of different versions of a PopART-like intervention that target subgroups in the population. Secondly, the impact of migration on trial impact, and on the estimation of model parameters, will be evaluated using data on in- and out-migration from CHiPs round 3. Thirdly, the potential impact of drug resistance on nation-wide roll out of PopART-like interventions will be evaluated.

Levels of drug resistance will be informed using data from 'HPTN 071-2 Phylogenetics in HPTN 071', the phylogenetics ancillary study of the HPTN 071 (PopART) trial, and the main PC resistance survey results if these are available at the time the work progresses.

3.3 Comparison of projected and observed trial impact

The following analysis will be carried out after the PC data are unblinded and once the post-unblinding IBM projections are completed. This analysis will be based on RR using the reweighted, 'PC-like' population.

The HPTN 071 (PopART) trial is a three arm, cluster-randomised trial. Communities were matched in seven triplets. As mentioned above, RR are computed for each triplet, comparing arm A or B communities with arm C or CF of the same triplet.

3.4 Primary endpoint

The following three null-hypotheses will be treated as primary endpoints:

1. The trial impact based on pre-unblinding projections are not different from the observed trial impact.

$$RR_{A,CF}^{pre}(t_0, t_1) = RR_{A,C}^{obs}(t_0, t_1)$$

2. The trial impact based on post-unblinding CF projections are not different from observed trial impact.

$$RR_{A,CF}^{post}(t_0, t_1) = RR_{A,C}^{obs}(t_0, t_1)$$

3. The trial impact based on post-unblinding arm C projections are not different from observed trial impact.

$$RR_{A,C}^{post}(t_0, t_1) = RR_{A,C}^{obs}(t_0, t_1)$$

The three hypotheses will be examined with $t_0 = \text{PC12}$ and $t_1 = \text{PC36}$. The primary endpoint will be the comparison between the overall RR, computed as the geometric mean across triplets (R^d with $d = \{\text{pre}, \text{post}\}$), and the geometric mean based on the observed RR (R^{obs}) computed in the same way. We will compare the posterior predictive distribution of the geometric mean with the distribution of the observed geometric mean. The estimate and uncertainty for the observed geometric mean R^{obs} will be computed as specified in the Statistical Analysis Plan for the HPTN 071 trial. We will also perform similar comparisons for RR in each triplet.

3.5 Secondary endpoints

This analysis may be complemented by the RR based on the whole, simulated population instead of the 'PC-like' reweighted population.

In order to investigate differences in the primary endpoint in more detail the cumulative incidence for each community (denoted above as $\text{Cum.I}_c^d(t_0, t_1)$), for observed or simulated and stratified by gender and age-groups will be analysed as a secondary, complementary analysis. Uncertainty with respect to the IBM projections will be included and discussed in this secondary analysis.

Model projections for secondary endpoints, such as proportion of individuals on ART, aware of HIV status, or virally suppressed will be compared between arm A or B, and C or CF communities, as well as stratified by gender and age-groups.

Post-unblinding recalibration will include fitting to prevalence, proportion awareness of status, proportion on treatment, and VMMC coverage using data from arm C, as measured in the Population Cohort (as listed below). Post-unblinding recalibration will also be performed by fitting to incidence data from the Population Cohort.

4 PopART-IBM

4.1 Overview

The PopART IBM model has several components. Data from the UN Population Division (UNPD) are used to model mortality and fertility within Zambia and South Africa from 1900 and until 2050 using UNPD projections. Introduction of HIV into the simulated population is assumed to occur between 1970-80, and HIV transmission is modelled in heterosexual partnerships. Partnership formation and dissolution is informed using data from the baseline PC data. Three risk groups, allowing different numbers of concurrent partners and duration of partnerships, are modelled. HIV disease progression is assumed to follow data from the AIDS Therapy Evaluation in the Netherlands (ATHENA) study.

The IBM models both the PopART intervention and HIV testing and treatment through a background care cascade. Within arm A communities, CHiPs teams are assumed to visit individuals within each community with a coverage, stratified by age and sex, that matches data from the field. HIV testing, with assumed 100% sensitivity and specificity, is carried out with each simulated CHiPs visit, and individuals with an HIV+ test result are offered ART immediately within simulated arm A communities. Time until ART initiation following an HIV+ test result is modelled following data from the trial. Individuals starting ART can either become virally suppressed, virally unsuppressed, or drop out of care at a later date, and risk of HIV transmission to partners is dependent upon an individual's profile within the care cascade.

The simulated background care cascade differs from the PopART intervention in ART eligibility, and drop out rates. Repeat CD4 testing (for those not eligible for ART immediately) is simulated in

the background care cascade.

Voluntary medical male circumcision (VMMC) is offered to any uncircumcised, HIV negative, male following a negative HIV test and VMMC uptake is assumed to be different between the PopART intervention and the background care cascade. VMMC is assumed to offer a 60% reduction in susceptibility and traditional male circumcision (TMC) is assumed to offer no protection. Circumcision coverage, both VMMC and TMC, differs by country in accordance with PC data.

4.2 Input parameters

The PopART IBM has over 350 parameter. Like many models of HIV in generalised epidemics, the PopART IBM is over-parameterised, given the data available. The majority of these are informed from the HIV literature, UNPD demographic projections, CHiPs process data, and PC0 survey data. Several parameters are allowed to vary within the model to represent processes that may naturally be stochastic, such as individual differences in the progression of CD4 decline after HIV infection. There are 13 input parameters that are estimated using an inference framework (outlined below) which are presented in the following table.

Table 1: List of parameters varied in calibration, and corresponding ranges used.

Parameter	Values	Notes
Initialisation parameters		
<i>Proportion by activity group when entering population</i>		
Low-activity men	30-60%	Based on variability between communities of the HPTN 071 (PopART) trial at baseline
Low-activity women	30-60%	
Medium-activity men (of those not in low activity group)	50-99%	
Medium-activity women (of those not in low activity group)	50-99%	
Factor multiplying initial % of population seeded HIV+ at the start of the HIV epidemic	1-100	Sampled on log scale
HIV-related parameters		
Average annual hazard of (uncircumcised) man getting HIV from a HIV+ partner who has maximal SPVL	0.05-0.3 yr ⁻¹	Hazard is adjusted down via a Hill function for lower SPVL. Literature estimate 0.313 yr ⁻¹ for individuals with SPVL 1 million copies/ml ?
Relative infectivity of male-to-female transmission	1.0-3.0	Lower limit no difference, upper limit is mean of low and high-income country estimates from Boily LID 2009.
Probability collect background CD4 test results	0.75-0.95	Lower limit ?
Multiplier for rate of background HIV testing from 2006 onwards	1.0-4.5	Assumption
Mean time to start ART through background HIV testing (of those who decide to start ART)	0.4-0.7 yrs	Range from analysis of CHiPs data by country.
Lifetime probability of becoming virally suppressed after period of early ART	0.55-0.9 yrs	Assumption, including levels in with ? and ?.
Partnership-related parameters		
Risk assortativity	0.05-0.95	Large range to reflect uncertainty. 1.0 means that people report the correct number of partners, > 1 means people are underreporting. Lower range taken to ensure 1 is sampled well and to allow over-reporting. Partnership durations assumed to be exponentially distributed with mean 15.4/6.1/3.8 (Zambia) and 9.8/7.0/4.2 (South Africa) years for low/medium/high activity partnerships respectively.
Multiplier to account for mis-reporting of number of sexual partners	0.5-4.0	
Partnership duration mean multiplier	1.0-2.0	

5 Inference framework

The PopART IBM is parameterised using a Bayesian framework called Approximate Bayesian Computation (ABC). ABC is a suite of algorithms that approximate a conditional probability density function as described in Beaumont et al. (2009).

In some situations a likelihood function is not available or is intractable, but a simulation from a corresponding distribution, here by using the PopART-IBM, is possible. The simulated projections therefore replace the intractable likelihood. A distance measure compares simulated and observed summary statistics. A simulated projection, for a given IBM input parameter combination, is accepted if the distance measure is below a predefined threshold. The resulting distribution of the model projections is proportional to the posterior distribution given the data and the IBM. As summary statistics the sex and age-group stratified HIV prevalence from historical data as well as from the CHiPs data is used. These are complemented with the proportion of individuals aware of their status and on ART. ABC can be done with rejection sampling, MCMC and sequential Monte Carlo (SMC) (McKinley et al., 2018). We use the algorithm suggested by Lenormand et al. (2015) entitled as the adaptive PMC-ABC algorithm (see appendix in Lenormand et al. (2015)). ABC inference for the PopART simulations is based on 2,000 simulation in each step, with an acceptance rate of 50% and the PMC-ABC algorithm will stop if the proportion of simulations with decreased distance measure is below 9%. The model projections will be reported as quantiles of the posterior distribution, which consists of the 1,000 accepted simulations produced in the final ABC-step.

6 Software and Code

The C code for the PopART-IBM is housed in a private repository on (github.com). The C code does not house any data that is not publicly available (such as parameter values sourced from the literature). R and Python code used for pre-processing parameter values and the inference framework is housed in a separate private repository on github.com. Additional scripts of helper functions and housekeeping scripts are coded in Python, R, and as shell scripts and are also stored on repositories on github.com.

The ABC inference will be carried out using the implementation in the R-library EasyABC (Jabot et al., 2015) version 1.5, which is discussed in Jabot et al. (2013).

In each report, the pre- and post-unblinding reports, a commit number will be added for each repository. The commit number and date allows users to access the software and script version used to generate each report. Should there be changes made to one of the scripts hosted on a repository at a later point, this will be reproducible, as previous versions will be stored and only incremental changes will be added to the repositories.

References

- BEAUMONT, M., CORNUET, J.-M., MARIN, J.-M. and ROBERT, C. P. (2009). Adaptive approximate bayesian computation. *Biometrika* **96** 983–990.
URL <http://www.jstor.org/stable/27798882>
- JABOT, F., FAURE, T. and DUMOULIN, N. (2013). Easyabc: performing efficient approximate bayesian computation sampling schemes using r. *Methods in Ecology and Evolution* **4** 684–687.
URL <https://besjournals.onlinelibrary.wiley.com/doi/abs/10.1111/2041-210X.12050>
- JABOT, F., FAURE, T., DUMOULIN, N. and ALBERT, C. (2015). *EasyABC: Efficient Approximate Bayesian Computation Sampling Schemes*. R package version 1.5.
URL <https://CRAN.R-project.org/package=EasyABC>
- LENORMAND, M., JABOT, F. and DEFFUANT, G. (2015). Adaptive approximate bayesian computation for complex models. *arXiv:1111.1308 [math.ST]* .
- MCKINLEY, T. J., VERNON, I., ANDRIANAKIS, I., MCCREESH, N., OAKLEY, J. E., NSUBUGA, R. N., GOLDSTEIN, M. and WHITE, R. G. (2018). Approximate bayesian computation and simulation-based inference for complex stochastic epidemic models. *Statist. Sci.* **33** 4–18.
URL <https://doi.org/10.1214/17-STS618>