

Appendix: Model technical details

Microsimulation model

The microsimulation began by generating baseline characteristics for all individuals, with distributions based on data observed in individuals infected with HIV in British Columbia (BC). These characteristics included demographics as well as random effects for each of the cost models and frailty terms for the resistance model. The random effects for the cost models represented an individual's tendency to utilize particular health services, above and beyond what would be described by the covariates included in the cost models (i.e. CD4 cell count, adherence, treatment status, demographics). Similarly, the frailty terms for the resistance model described an individual's tendency to develop medication resistance beyond what would be expected based on their adherence to HAART and other variables included in the resistance model. An additional variable that was generated at the outset of the microsimulation was the expected time following infection at which an individual was expected to present for CD4 monitoring. This was assumed to follow a uniform distribution on 0–10 years, which was chosen based on the observed initial CD4 cell counts in treated individuals in BC. It was further assumed that after an individual presented for CD4 monitoring, they would initiate treatment with HAART once they were considered clinically eligible, with eligibility based on a CD4 cell count below 350 cells/mm³.

The microsimulation then proceeded by moving forward in one-month time cycles. At each time step, CD4 cell count and viral load were updated. Treatment uptake in the upcoming month for individuals not yet receiving HAART was determined based on their CD4 cell count and whether or not they had presented for monitoring.

Resistance mutation development was randomly determined assuming probabilities based on the Weibull frailty model. It was assumed that if an individual initiating a non-nucleoside reverse transcriptase inhibitor (NNRTI)-based regimen developed an NNRTI resistance mutation, they would switch to a protease inhibitor (PI)-based regimen, and remain on PI-based regimens for the remainder of their life, due to the increased likelihood of cross-class resistance to NNRTIs. Similarly, individuals who started on a PI-based regimen were assumed to switch to an NNRTI-based regimen if they developed PI resistance, and to switch back to a PI-based regimen if they subsequently became resistant to NNRTIs. The distribution of times associated with these regimen switches were estimated based on empirical data and accounted for in the frailty model.

Monthly costs associated with the various health services categories (hospitalizations, outpatient visits, laboratory tests, emergency room visits and medications) were then generated. These costs were generated based on a series of random effects models that were fit using comprehensive health services data. For a population-based cohort of 1,895 individuals infected with HIV, data were available describing all charges associated with physician visits and days spent in hospital. In addition, for all individuals receiving treatment for HIV in BC, all prescription records and laboratory tests were available. These data were converted to longitudinal data series with a person-month unit of analysis and this longitudinal dataset was used to fit a series of two-stage random effects statistical models which were then used to predict costs within the microsimulation. The first stage random effects models were used to determine the probability of non-zero utilization for each of the categories, and random numbers were generated to determine which individuals would utilize each of the categories during the upcoming month. Conditional on non-zero utilization to a particular health services category, the second stage random effects model was used to estimate the expected log-cost associated with that category for the upcoming month. The expected log-costs calculated using the second stage models were treated as point estimates associated with normally-distributed random variables. Actual log-costs were then randomly sampled from normal distributions with mean and standard deviation equal to the point estimate and the residual standard error of the second stage model, respectively. Finally, these log-costs were transformed back to the original cost scale using the Duan smearing factor.

Survival throughout the upcoming interval was randomly determined based on a Cox proportional hazards model. For individuals who died during the upcoming one-month interval, the amount of time they contributed to the interval prior to death was sampled from a uniform distribution, and the final month-of-life indicator was included in the cost models to reflect the increase in expected costs associated with this interval.

In the sensitivity analysis, all coefficients of the statistical models (costs, resistance, survival) were randomly generated based on the covariance distributions defined by the model-fitting process. Further information regarding the microsimulation or individual statistical models can be obtained by contacting the authors.

Transmission model structure

The transmission model structure is shown in Appendix Figure 1. Simulated time moved forward in one-month intervals, during which individual disease histories and resource utilization profiles were updated within the

microsimulation, and new infections were estimated using a series of difference equations. Newly- infected individuals were then added to the cohort being followed within the microsimulation, and their disease trajectories and resource utilization were simulated prospectively. As individuals died within the microsimulation process they were removed from the cohort. The pool of susceptible individuals varied over time as individuals migrated in or left due to death or infection. This process was repeated at monthly intervals throughout the entire simulated time period. Discrete-time difference equations were used in place of continuous-time differential equations because the model was structured to “stop” at each one-month time interval and calculate new infections based on the distribution of clinical variables determined by the microsimulation. The difference equations described a process in which susceptible individuals became infected at a rate determined by the respective numbers of susceptible and infected individuals, the distribution of viral load amongst infected individuals, the baseline level of risk behaviour observed in the population, and any decrease in risk behaviour due to increased viral load. Mathematically, the equations are expressed:

$$S_{t+30} = S_t - \left(S_t \beta_0 \sum_{j=0}^4 I_t^{(j)} \gamma^{(j)} (1 - \delta^{(j)}) / N \right) + \lambda$$

$$I_{t+30} = I_t + \left(S_t \beta_0 \sum_{j=0}^4 I_t^{(j)} \gamma^{(j)} (1 - \delta^{(j)}) / N \right) - \mu_t$$

In this notation, t is measured in days, S_t refers to the number of susceptible individuals at time t , $I_t^{(j)}$ refers to the number of individuals in viral load category j at time t , β_0 is the baseline population force of infectivity, $\gamma^{(j)}$ is the increase in infectivity associated with viral load category j , $\delta^{(j)}$ is a dampening of the increased infectivity associated with viral load category j due to a decrease in risk behaviour, N is the total number of individuals in the population (susceptible and infected), λ is the net migration into the susceptible population accounting for migration in and out and mortality (note that λ could be positive or negative, depending on migration patterns), and μ_t is the mortality rate of infected individuals at time t , which varies according to the disease stage distribution at time t . The viral load categories considered were: primary infection during the 60 days following seroconversion ($I [0]$); viral load below 3 log copies per mL ($I [1]$), viral load in the interval [3,4) log copies per mL ($I [2]$), viral load in the interval [4,5) log copies per mL ($I [3]$), and viral load greater than or equal to 5 log copies per mL ($I [4]$). Category $I [1]$ was taken as the baseline category, so $\gamma [1]$ and $\delta [1]$ were both defined to be 1, meaning that individuals in this viral load category represent the baseline level of infectivity and risk behaviour. For increasing viral load categories, it was assumed that an increase of 1 log copy/mL was associated with a 2.45-fold increase in infectivity [13] and a 10% reduction in risk

behaviour, based on the assumption that viral load could be used as an acceptable proxy for morbidity. This decrease in risk behaviour was varied in the sensitivity analysis. It was further assumed that during the period of primary infection, viral load reached a level of 6 log copies/mL with no reduction in risk behaviour.

In order to reduce model complexity, we assumed a single baseline infectivity parameter (β_0), and did not differentiate between different routes of transmission (e.g. men who have sex with men, injection drug use, sex trade work). Thus, β_0 can be viewed as implicitly incorporating the relative sizes of different risk groups, and the level of risk behaviour and risk of transmission associated with each risk group. It is unlikely that β_0 is actually static, as the relative size of risk groups and the behaviour of individuals change over time. However, we made the simplifying assumption that the β_0 parameter would remain constant over the simulated time period. We do not expect that this parameter uncertainty in β_0 had a major impact on model results. Within this study, it was not our objective to exactly replicate transmission patterns observed in BC or predict the specific number of infections to be expected in upcoming years. Rather, we wished to evaluate the difference in outcome between two treatment strategies in the context of a mature, concentrated epidemic consistent with that observed in BC. Because our primary outcomes were incremental in nature and all comparisons used the same parameter estimates, it was unlikely that any incorrect assumptions regarding specific parameter values would have biased overall results in either direction.

Net Benefit

The primary outcome of the model was the incremental net benefit (INB), which is defined as:

$$INB = \lambda \Delta QALY - \Delta Cost$$

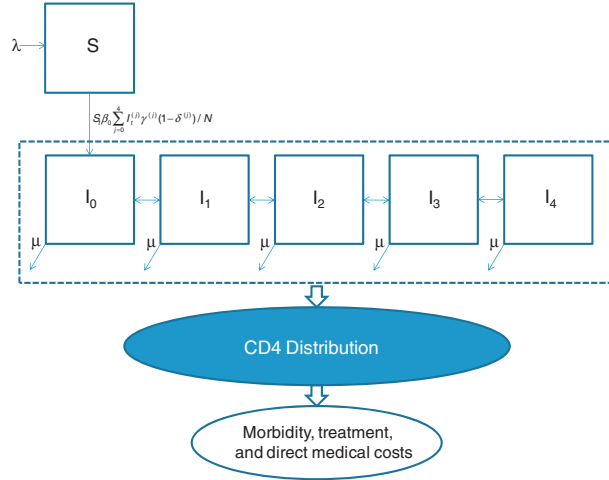
In the above formula, $\Delta QALY$ refers to the difference across scenarios in quality-adjusted lifeyears, $\Delta Cost$ refers to the difference across scenarios in total costs, and λ refers to the societal willingness-to-pay for an additional quality-adjusted life year, assumed here to be \$50,000.

Another commonly-used metric in health economics is the incremental cost-effectiveness ratio (ICER), defined as:

$$ICER = \frac{\Delta Cost}{\Delta QALY}$$

Using the ICER metric, a scenario can be considered cost-effective if the related ICER is below a pre-specified willingness-to-pay threshold λ . [43] Noting the relationship between the INB and ICER formulas, for a given value of λ , a cost-effective ICER is equivalent to an $INB > 0$.

Appendix Figure 1: Schematic of model structure.



Appendix Table 1: Parameters associated with transmission model.

Parameter	Description	Point estimate	Distribution used in probabilistic sensitivity analysis/ Corresponding 95% Confidence Interval (CI)	Source(s)
S_0	Susceptible population at baseline	15,000	Normal ($\sigma=2,000$) 95% CI: (11,080–18,920)	• Assumption
β_0	Baseline force of infectivity (associated with viral load of 3 log copies/mL)	0.02	Beta ($\alpha=15.66, \beta=767.34$) 95% CI: (0.011–0.031)	• Empirical calibration [26]
$\gamma^{(j)}$	Increase in infectivity associated with viral load category j relative to baseline	2.45 per log increase	Normal ($\sigma=0.15$) 95% CI: (2.16–2.74)	• Beta distribution parameters estimated using method of moments assuming standard deviation of 0.005.
$\delta^{(j)}$	Decrease in risk behaviour associated with viral load category j relative to baseline	0.10 per log increase (excluding primary infection)	Beta ($\alpha=0.8, \beta=7.2$) 95% CI: (0.001–0.367)	• Quinn <i>et al.</i> [13]
λ	Net annual migration into susceptible population (accounting for mortality)	500	Normal($\sigma=100$) 95% CI: (304–696)	• Assumption
μ_t	Mortality rate at time t .	Dependent on disease stage and age distribution and treatment history of simulated population.	N/A	• Beta distribution parameters estimated using method of moments assuming standard deviation of 0.10.
				• Empirical calibration
				• Untreated: Babiker <i>et al.</i> [25]
				• Treated: Cox proportional hazards analysis of British Columbia data.
				• Non-HIV related: British Columbia age-specific mortality. [44, 45]
Viral load set point	Baseline viral load (assumed to increase by 0.1 log-copies per year)	4 log-copies	Normal($\sigma=0.5$) 95% CI: (3.02–4.98)	• Fraser <i>et al.</i> [24] British Columbia data