

AN EXACT APPROACH TO CALIBRATING INFECTIOUS DISEASE MODELS TO SURVEILLANCE DATA: THE CASE OF HIV AND HSV-2

DAVID J. GERBERRY

Department of Mathematics
Xavier University
3800 Victory Parkway
Cincinnati, OH 45207, USA

ABSTRACT. When mathematical models of infectious diseases are used to inform health policy, an important first step is often to calibrate a model to disease surveillance data for a specific setting (or multiple settings). It is increasingly common to also perform sensitivity analyses to demonstrate the robustness, or lack thereof, of the modeling results. Doing so requires the modeler to find multiple parameter sets for which the model produces behavior that is consistent with the surveillance data. While frequently overlooked, the calibration process is nontrivial at best and can be inefficient, poorly communicated and a major hurdle to the overall reproducibility of modeling results.

In this work, we describe a general approach to calibrating infectious disease models to surveillance data. The technique is able to match surveillance data to high accuracy in a very efficient manner as it is based on the Newton-Raphson method for solving nonlinear systems. To demonstrate its robustness, we use the calibration technique on multiple models for the interacting dynamics of HIV and HSV-2.

1. Introduction. In order to take a mathematical model of infectious disease from a theoretical construct to a practical tool for informing health policy in a specific setting, an important initial step is model calibration. The terms “model calibration”, “model fitting” and “model parameterization” all describe the process of estimating values for the parameters used within a mathematical model so that its output is relevant to the situation of interest. Despite its importance, calibration is commonly overlooked as a part of the modeling process with its details often relegated to appendices or supporting materials. Insufficient communication of the calibration process presents a significant obstacle to the overall reproducibility of modeling work.

In this work we present an exact approach to calibrating infectious disease models to surveillance data that aims to streamline the process while increasing efficiency and reducing reliance on computational methods. At its core, the approach reserves specified model parameters to be used as fitting variables where the number of fitting variables is determined by the number of calibrating conditions to be satisfied. Estimates for the remaining model parameters are established from empirical studies

2010 *Mathematics Subject Classification.* Primary: 92D30; Secondary: 92D25, 37N25, 37N30, 92C60.

Key words and phrases. Mathematical modeling, calibration, model-fitting, HIV, HSV-2, surveillance data.

and/or previous modeling studies. Once a set of values has been specified for non-fitting parameters, values of the fitting variables are found so that the mathematical model exactly matches the given calibration conditions.

The structure of the paper is as follows: in the next section we describe three general approaches to model calibration. In Section 3, we outline our general approach to exact model calibration and illustrate it with a simple toy model. Focusing our attention to the calibration of models for the interaction of HIV and HSV-2 (genital herpes) for illustration, we then demonstrate our calibration approach for multiple mathematical models of varying complexity: a 4-compartment *SI*-type model for the interaction of HIV and HSV-2 in Section 4, the *Granich et al.* model [11] for HIV in Section 5, an HIV/HSV-2 coinfection model with behavioral response in Section 6 and some extensions to the calibration approach in Section 7. Finally, in Section 8 we provide a summary of the approach and discuss generalizations, strengths and weaknesses.

2. Background. When parameterizing a model, the subset of parameters that will be allowed to vary in order to achieve a desirable fit to data is identified. We refer to these as calibration or fitting parameters. In an ideal situation, reliable estimates and/or ranges of values exist for the remaining model parameters. Model calibration or fitting is then the process of specifying values for the calibration parameters. While many techniques for doing so exist, we emphasize three general approaches.

While easy to overlook as a method, manual calibration is quite possibly the most ubiquitous approach to model fitting. We use the term “manual calibration” to loosely refer to scenarios in which a modeler simply plays with parameter values until a satisfactory fit is achieved. Despite its lack of rigor, the hands-on approach of adjusting parameter values and observing the effects offers the researcher an opportunity to gain valuable insights and familiarity with a model. Of course, obvious downsides for the manual approach include the facts that the process can be inefficient, produce less than desirable calibration results and can be difficult to reproduce.

A more formal, yet still intuitive, approach involves obtaining random samples of parameter sets and finding those that adequately fit the data. In this approach, one begins by specifying probability distributions for the parameters of interest. Multiple random parameter sets are then sampled from these distributions (often using a Latin Hypercube Sampling technique) with the model simulated for each parameter set. Parameter sets that do not produce model behavior that adequately fits the given data are then filtered out leaving only those that are relevant. The remaining parameter sets then represent the result of the calibration process. Such a calibration is conducive to uncertainty and sensitivity analysis as variation in multiple model parameters is included in the parameter sets. However, a drawback of the approach is its ultimate reliance on sophisticated statistical techniques (e.g. partial rank correlation coefficients, logistic regression, etc.) to demonstrate relationships between model parameters and behaviors of interest. The required familiarity with these techniques (including their appropriateness in the given situation and their metrics) can impede clear understanding of the modeling results for some readers. Lastly, the fact that parameters are randomly sampled means that, although the qualitative behavior should be the same, the modeling results are not completely reproducible.

The last and most formal approach to calibration is to use a numerical fitting algorithm. In this approach, optimal parameter values are found that minimize a

specified metric of error (e.g. absolute or squared error between model simulation and data). Numerical fitting algorithms can be deterministic (e.g. Nelder-Meade simplex method and Levenburg-Marquardt method) or involve a stochastic search component (e.g. Markov chain Monte Carlo search and genetic algorithms) and are often quite computationally demanding. The benefit of such rigorous methods is that they reproducibly produce the very best fit possible for a given model. The price for their performance is that the details are complicated and hidden within subroutines of many proprietary software packages. Consequently, the process of model calibration is hidden within a “black box” for many practitioners.

3. Illustration of the exact approach to model calibration. The goal of this work is to introduce an exact approach to model calibration. Our approach is not intended to replace those described above but rather as an alternative that may be preferable in certain situations or could be used in conjunction with other methods in others.

For example, it is often the case that one wishes to calibrate an epidemiological model to disease surveillance data (i.e. standardized estimates of disease prevalence, incidence and other important indicators) such as those contained in WHO Global TB database and the UNAIDS/WHO Global HIV/AIDS Online Database. More specifically, assume that we are examining an infectious disease in a particular setting for which disease prevalence is at or near endemic equilibrium.

As a simple illustrative example of the overall approach, consider calibrating a basic SI infectious disease model, given by

$$\begin{aligned} S' &= \Lambda - \beta SI - \mu S, \\ I' &= \beta SI - \mu I, \end{aligned}$$

to disease prevalence at endemic equilibrium. We must first identify the calibration parameter. In general, we note that disease transmission coefficients are particularly well-suited for this role as reliable estimates for transmission coefficients rarely exist and because transmission coefficients incorporate a multitude of factors that can vary significantly among different settings and populations (e.g. biological and environment factors, human behavior, effect of interventions, etc.). Using β as our calibration parameter, we note that the model yields a disease prevalence of $P^* = 1 - \mu/\beta$ at endemic equilibrium. Hence, when values for the other model parameters are established, the transmission coefficient β can be used to exactly calibrate to a specified disease prevalence, P^* , using the fact that $\beta = \mu/(1 - P^*)$.

Of course, calibration in this simple example is greatly facilitated by an analytic expression for disease prevalence at equilibrium that will not be available for more detailed models. In following sections, we show how the construction of a calibration system of equations that is then solved by Newton’s Method can overcome this hurdle for models of varying complexity.

4. Four-compartment SI -type HIV/HSV-2 coinfection model. In this section, we consider the simplest framework for the interacting dynamics of HIV and HSV-2. The model includes only four classes of individuals: susceptible to infection by both HIV and HSV-2, infected with HSV-2 only, infected with HIV only, and co-infected with both HIV and HSV-2. To model the epidemics of the two infections, we incorporate the multiples ways that HIV and HSV-2 interact.

Many epidemiological studies [9, 21, 14, 5, 15] have shown that infection with HSV-2 increases an individual's susceptibility to infection with HIV. The mechanisms behind this increased susceptibility are that genital ulcers resulting from HSV-2 infection provide passage to HIV viruses through epithelial cell layers to the more vulnerable cell layers beneath and that HSV-2 infection promotes the recruitment of HIV target cells to the genital region. In our model, HIV-negative individuals that are infected with HSV-2 are more likely to become infected with HIV than those not infected with HSV-2. Higher HIV infectivity has also been observed for individuals that are coinfecting with HIV and HSV-2 compared to those infected with only HIV [12, 4] due to increased HIV viral load and/or increased access through mucosal membranes via herpetic lesions. Consequently, we model HIV-positive individuals that are also HSV-2 infected as more infectious than those that are only infected with HIV. Lastly, it has also been shown that infection with HIV increases HSV-2 infectivity of coinfecting individuals by increasing the duration of HSV-2 shedding episodes and increases the frequency of such episodes [1].

4.1. The model. The interacting dynamics of HIV and HSV-2 infection are modeled using the following system:

$$\begin{aligned} S' &= \Lambda - \frac{\beta S(A + r_2 C)}{N} - \frac{\sigma S(H + r_3 C)}{N} - \mu S, \\ H' &= \frac{\sigma S(H + r_3 C)}{N} - \frac{r_1 \beta H(A + r_2 C)}{N} - \mu H, \\ A' &= \frac{\beta S(A + r_2 C)}{N} - \frac{\sigma A(H + r_3 C)}{N} - (\mu + \mu_A)A, \\ C' &= \frac{r_1 \beta H(A + r_2 C)}{N} + \frac{\sigma A(H + r_3 C)}{N} - (\mu + \mu_A)C, \end{aligned} \tag{1}$$

where S is the number of fully susceptible individuals (i.e. infected with neither HIV nor HSV-2); H is the number of individuals infected with HSV-2 only; A is the number of individuals infected with HIV only; C is the number of individuals infected with both HIV and HSV-2; and $N = S + H + A + C$ is the total size of the population. We assume a background mortality rate of μ and that disease mortality is only induced by HIV infection and is modeled at a rate of μ_A . The model incorporates the most important aspects of the interaction between HIV and HSV-2 infection. Namely, being infected with HSV-2 renders an individual more susceptible to HIV infection (by a factor of r_1); co-infected individuals are more infectious and likely to spread both HIV and HSV-2 (r_2 and r_3 , respectively). Consequently, we have that $r_i \geq 1$ for $i = 1, 2, 3$. For detailed descriptions of how the empirical findings on the interaction of HIV and HSV-2 infection from [9, 21, 14, 5, 15, 9, 21, 14, 5, 15, 12, 4] are transformed into transmission cofactors for mathematical modeling, see [10, 1]. The flow diagram for the model is depicted in Figure 4.

4.2. Calibration to HIV and HSV-2 prevalence at equilibrium. As in Section 3, we will consider the situation where we wish to parameterize the current model in (1) for a setting with specified HIV and HSV-2 prevalence levels (which we denote as \hat{A} and \hat{H} , respectively) assuming both diseases are at equilibrium. To achieve the two prevalence conditions, we use the transmission coefficient for HIV, β , and that for HSV-2, σ , as fitting parameters. We assume that values for the remaining model parameters have been established (e.g. fixed from the literature).

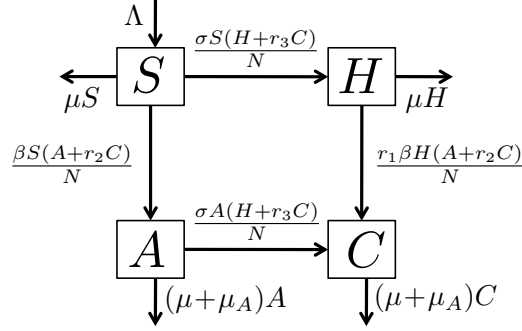


FIGURE 1. Flow diagram of the four-compartment SI -type HIV/HSV-2 coinfection model. $N = S + H + A + C$.

It is worth noting that despite the relative simplicity and small dimension of the model in (1), finding an explicit expression for the its endemic equilibrium is intractable via direct methods due to the highly nonlinear nature of the system; even when using modern computer algebra systems (CAS).

Our calibration approach is facilitated by rescaling the state variables of the model so that $s = \frac{S}{N}$, $h = \frac{H}{N}$, $a = \frac{A}{N}$ and $c = \frac{C}{N}$. Hence, state variables of the rescaled model represent proportions of the total population rather than numbers of individuals. Noting that $s = 1 - h - a - c$, the full model (1) reduces to

$$\begin{aligned} h' &= \sigma(1 - h - a - c)(h + r_3c) - r_1\beta h(a + r_2c) - \mu h, \\ a' &= \beta(1 - h - a - c)(a + r_2c) - \sigma a(h + r_3c) - (\mu + \mu_A)a, \\ c' &= r_1\beta h(a + r_2c) + \sigma a(h + r_3c) - (\mu + \mu_A)c. \end{aligned} \quad (2)$$

If the endemic equilibrium is denoted by (h^*, a^*, c^*) and we calibrate to the prevalence data

$$\begin{aligned} \hat{A} &: \text{overall HIV prevalence at endemic equilibrium,} \\ \hat{H} &: \text{overall HSV-2 prevalence at endemic equilibrium,} \end{aligned} \quad (3)$$

it follows that $\hat{A} = a^* + c^*$ and $\hat{H} = h^* + c^*$. At endemic equilibrium, it must be the case that $h' = a' = c' = 0$. Including our calibration conditions, we arrive at our calibration system of equations

$$0 = \sigma^*(1 - h^* - a^* - c^*)(h^* + r_3c^*) - r_1\beta^*h^*(a^* + r_2c^*) - \mu h^*, \quad (4)$$

$$0 = \beta^*(1 - h^* - a^* - c^*)(a^* + r_2c^*) - \sigma^*a^*(h^* + r_3c^*) - (\mu + \mu_A)a^*, \quad (5)$$

$$0 = r_1\beta^*h^*(a^* + r_2c^*) + \sigma^*a^*(h^* + r_3c^*) - (\mu + \mu_A)c^*, \quad (6)$$

$$0 = a^* + c^* - \hat{A}, \quad (7)$$

$$0 = h^* + c^* - \hat{H}, \quad (8)$$

where the first three equations impose the equilibrium conditions and the last two impose the calibration conditions on disease prevalence.

4.2.1. Newton's Method. Introduced in many first semester Calculus courses, Newton's Method is perhaps the most famous numerical method because of its efficiency and elegance. In the one-dimensional case, the method finds successively better approximations to the solution of $f(x) = 0$ using the iteration $x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$.

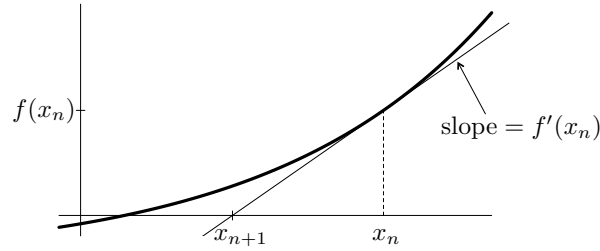


FIGURE 2. Visual illustration of Newton's Method in 1-dimension for finding the solution of $f(x) = 0$. As illustrated, the iteration $x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$ results from finding the root of the tangent line at the current iteration $(x_n, f(x_n))$.

As seen in Figure 2, Newton's Method essentially approximates the root of a function using the root of its first order approximation (i.e. its tangent line in the 1-dimensional case). In n -dimensions where we wish to solve $f(\vec{x}) = \vec{0}$, Newton's Method can be written as $\vec{x}_{n+1} = \vec{x}_n - [Df(\vec{x}_n)]^{-1}f(\vec{x}_n)$, where $Df(\vec{x})$ is the Jacobian matrix of f evaluated at \vec{x} . As matrix inverses are computationally unstable, the typical iteration is $\vec{x}_{n+1} = \vec{x}_n + \Delta\vec{x}$ where $\Delta\vec{x}$ is the solution to $Df(\vec{x}_n)\Delta\vec{x} = -f(\vec{x}_n)$. As long as the initial approximation \vec{x}_0 is close enough to the actual root and the Jacobian is nonsingular, Newton's Method converges quickly to the desired root.

Unfortunately, there is no theoretical result dictating how close is "close enough" to guarantee convergence and one must often experiment with initial approximations until a suitable one is found [6]. To use Newton's Method to solve of our calibration system (8), it turns out that finding an initial approximation close enough to the actual solution is not terribly demanding. In fact, using the simplifying assumption that HIV and HSV-2 infection are independent produces an adequate initial approximation in every scenario tested. Specifically, we let

$$c_0 = \hat{A}\hat{H}, \quad a_0 = \hat{A} - \hat{A}\hat{H}, \quad h_0 = \hat{H} - \hat{A}\hat{H}, \quad (9)$$

and solve (4) and (6) for σ and β to obtain initial estimates

$$\sigma_0 = \frac{(\mu + \mu_A)c_0 + \mu h_0}{(c_0 r_3 + h_0)(1 - c_0 - h_0)} = \frac{\hat{A}\mu_A + \mu}{(1 - \hat{H})(1 - (1 - r_3)\hat{A})}, \quad (10)$$

$$\beta_0 = \frac{(\mu + \mu_A)(1 - a_0 - h_0 - c_0)c_0 - a_0 h_0 \mu}{h_0 r_1 (1 - c_0 - h_0)(c_0 r_2 + a_0)} = \frac{\mu_A}{(1 - (1 - r_2)\hat{H})r_1}. \quad (11)$$

Given its prominence among numerical methods, a form of Newton's Method should be available in all computing platforms. In most software, Newton's Method is at the core of the default root-finding function. Among options for the required precision of the solution, maximum iterations, etc., most implementations can compute an automatic numerical approximation to Jacobian matrix. Alternatively, the user can choose to supply the exact Jacobian function to increase efficiency and accuracy. In every scenario tested, automatic approximations to the Jacobian have performed admirably.

4.3. Demonstration of calibration to HIV and HSV-2 prevalence. In this section, we demonstrate the performance of our calibration technique. We begin with two specific examples that illustrate the convergence of the numerical iteration for our calibration parameters in Section 4.3.1. We then demonstrate the robustness of the approach by calibrating the model for 50 independent, randomly sampled parameter sets and calibration conditions in Section 4.3.2. In Table 1, we summarize model parameters and calibration conditions as well as establish baseline values and/or ranges to be used in both demonstrations.

TABLE 1. Model parameters and calibration conditions for the four-compartment *SI*-type HIV/HSV-2 coinfection model. [†] Values used in the high HIV burden example; South Africa. [‡] Values used in the low HIV burden example; United Kingdom.

	Symbol	Baseline	Range	References
Model Parameters				
Average time in sexually active population (years)	$1/\mu$	35	30 – 40	
Recruitment rate into the sexually active population (/yr)	Λ	$\frac{1}{35}$	$\frac{1}{40} - \frac{1}{30}$	
Diseased-induced mortality due to HIV infection (/yr)	μ_A	$\frac{1}{20}^\dagger, \frac{1}{100}^\ddagger$	$\frac{1}{100} - \frac{1}{20}$	
Cofactor for increased HIV susceptibility due to HSV-2 infection	r_1	1.80	1.4 – 2.4	[9, 21, 13, 1]
Cofactor for increased HIV infectivity due to HSV-2 infection	r_2	1.45	1.2 – 1.7	[1, 12, 17, 4]
Cofactor for increased HSV-2 infectivity due to HIV infection	r_3	1.30	1.2 – 1.5	[1]
Calibration Conditions				
HIV prevalence at endemic equilibrium (%)	\hat{A}	17 [†] , 0.3 [‡]	0.3 – 17	[18]
HSV-2 prevalence at endemic equilibrium (%)	\hat{H}	50 [†] , 4 [‡]	4 – 50	[2, 16, 22, 20]
Calibration Parameters				
Transmission coefficient for HIV	β	determined by calibration procedure		
Transmission coefficient for HSV-2	σ	determined by calibration procedure		

4.3.1. *Specific examples.* To demonstrate the performance of our calibration approach for a high HIV prevalence setting, we consider the case of South Africa. Data from the World Bank [18] shows that, like in many sub-Saharan African countries, HIV prevalence has stabilized over last several years. Figure 3 shows HIV prevalence among South Africans aged 15-49 from 1990-2013. Given that HSV-2

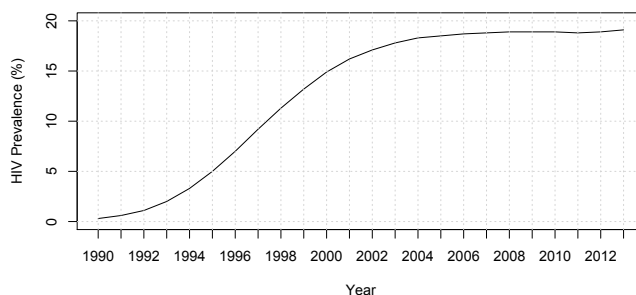


FIGURE 3. HIV prevalence in South Africa (% of population ages 15-49 infected with HIV) from 1990-2013. Data from the World Bank, World Development Indicators [18].

was prevalent in human populations long before HIV, it is reasonable to assume

that HSV-2 prevalence would have also been at or near endemic equilibrium. While more uncertainty exists in estimates of HSV-2 prevalence than HIV prevalence, data suggests an HSV-2 prevalence of roughly 50% [2, 16, 22].

For the South Africa example, we use the baseline parameter values from Table 1 which includes three values specific to South Africa (i.e. $\mu_A = \frac{1}{20}$ to incorporate the impact of HIV on life expectancy in South Africa, $\hat{A} = 0.17$ and $\hat{H} = 0.50$). Our initial approximation to the solution of the calibration system of equations (4)–(8) is found using (9)–(11) to get

$$c_0 = 0.085, \quad a_0 = 0.085, \quad h_0 = 0.415, \quad \sigma_0 = 0.0705, \quad \beta_0 = 0.0367.$$

Table 2 shows the iterations of Newton’s Method to complete the calibration. We note that the method converges after only 3 iterations.

We use data from the United Kingdom as an example of low HIV prevalence setting. Clearly, calibrating the unstructured model from (1) to HIV prevalence in the general population is not appropriate for making policy decisions regarding HIV in the UK given the concentrated nature of infection among particular risk groups. Given such low HIV burden, it is also difficult to argue that prevalence is at or near equilibrium. Here, we use this extreme example simply to illustrate the robust nature of our calibration approach. For this example, we assume an HIV prevalence of 0.3% [18] and an HSV-2 prevalence of 4.0% [20].

For the United Kingdom example, we use the baseline parameter values from Table 1 which includes three parameters values specific to the United Kingdom (i.e. $\mu_A = \frac{1}{100}$ to reflect that life expectancy of HIV-infected individuals is near that of uninfected individuals in the United Kingdom, $\hat{A} = 0.003$ and $\hat{H} = 0.04$). Our initial approximation to the solution of the calibration system of equations (4)–(8) is found using (9)–(11) to get

$$c_0 = 0.00012, \quad a_0 = 0.00288, \quad h_0 = 0.03988, \quad \sigma_0 = 0.029766, \quad \beta_0 = 0.035352.$$

As shown in Table 2, convergence to the solution of the calibration system is achieved in only 2 iterations of Newton’s Method.

4.3.2. Robustness of calibration approach. In this section, we calibrate the four-compartment *SI*-type HIV/HSV-2 coinfection model for a range of randomly sampled model parameters and calibration conditions. Specifically, we obtain 50 independent samples of model parameters and calibration conditions from uniform distributions over the ranges given in Table 1. The calibration system (4)–(8) was then solved using Newton’s Method with initial approximation given by (9)–(11). For each sample, the calibrated value of β and σ was noted along with the resulting HIV and HSV-2 prevalence at endemic equilibrium and the necessary number of iterations of Newton’s Method. The results are shown in Table 3. For all 50 simulations, the calibration approach successfully matched the calibration conditions (i.e. HIV and HSV-2 prevalence) while requiring no more than 4 iterations of Newton’s Method.

5. Granich et al. HIV model. In this section, we consider an HIV-only model based on that of Granich et al. [11]. This highly influential modeling work, which appeared in *The Lancet*, sparked considerable interest among the HIV community in the strategy of “Treatment as Prevention” which is essentially the idea that providing early antiretroviral treatment (ART) to HIV-infected individuals reduces their HIV viral load (and consequently their infectiousness) thus preventing future

TABLE 2. Calibration of the four-compartment *SI*-type HIV/HSV-2 coinfection model for a high prevalence and low prevalence setting; South Africa and the United Kingdom, respectively. Note that $i = 0$ the “0th iteration” refers to our initial approximation to the solution of the calibration system (4)–(8). HIV and HSV-2 prevalence at the endemic equilibrium that results from using the i th approximations for β and σ are denoted \hat{A}_i and \hat{H}_i , respectively. Precision is measured by average absolute value of the calibration system (4)–(8) evaluated at the current approximation (i.e. $\frac{\|f(h_i, a_i, c_i, \sigma_i, \beta_i)\|_1}{5}$). A stopping criteria of $\max\{|\beta_i - \beta_{i-1}|, |\sigma_i - \sigma_{i-1}|\} < 10^{-9}$ is used for Newton’s Method.

i	β_i	σ_i	\hat{A}_i	\hat{H}_i	precision _{i}
South Africa (HIV prevalence 15%, HSV-2 prevalence 50%)					
0	0.0367143341	0.0705450591	0.0000000447	0.5949903301	3.10×10^{-3}
1	0.0528598059	0.0770336930	0.1775025497	0.4988482844	2.77×10^{-4}
2	0.0522429752	0.0764216142	0.1700166525	0.5000125340	1.44×10^{-6}
3	0.0522423466	0.0764181115	0.1700000000	0.5000000000	3.38×10^{-11}
United Kingdom (HIV prevalence 0.3%, HSV-2 prevalence 4%)					
0	0.0353517304	0.0297663650	0.0000221420	0.0401412229	5.78×10^{-6}
1	0.0359324850	0.0297727618	0.0033368789	0.0399611984	1.86×10^{-8}
2	0.0359185167	0.0297727520	0.0030000000	0.0400000000	1.10×10^{-13}

TABLE 3. Robustness of calibrating the four-compartment *SI*-type HIV/HSV-2 coinfection model to HIV and HSV-2 prevalence at the endemic equilibrium. For each sample $j = 1, 2, \dots, 50$, the resulting values of the calibration parameters (β^*, σ^*) and the resulting disease prevalences at equilibrium (\hat{A}^*, \hat{H}^*) are given. The number of iterations of Newton’s Method required to achieve a stopping criterion of $\max\{|\beta_i - \beta_{i-1}|, |\sigma_i - \sigma_{i-1}|\} < 10^{-9}$ is denoted by m .

j	Sampled parameter values and calibration conditions										Calibration results				
	λ	μ	μ_A	r_1	r_2	r_3	r_4	\hat{A}	\hat{H}	β^*	σ^*	\hat{A}^*	\hat{H}^*	m	
1	0.026	0.028	0.037	1.98	1.65	1.24	0.088	0.276	0.043	0.045	0.088	0.276	4		
2	0.033	0.027	0.015	2.39	1.23	1.22	0.146	0.164	0.039	0.035	0.146	0.164	4		
3	0.028	0.027	0.022	1.55	1.25	1.23	0.007	0.340	0.037	0.042	0.007	0.340	3		
4	0.025	0.022	0.042	2.39	1.52	1.50	0.025	0.275	0.043	0.046	0.025	0.275	3		
5	0.025	0.028	0.016	2.24	1.33	1.36	0.145	0.291	0.034	0.042	0.145	0.291	4		
6	0.032	0.027	0.021	2.01	1.51	1.28	0.123	0.294	0.034	0.042	0.123	0.294	4		
7	0.031	0.033	0.033	1.89	1.57	1.41	0.065	0.290	0.043	0.049	0.065	0.290	4		
8	0.031	0.030	0.020	1.67	1.67	1.48	0.110	0.339	0.034	0.046	0.110	0.339	4		
9	0.027	0.028	0.030	2.19	1.29	1.45	0.016	0.409	0.033	0.048	0.016	0.409	3		
10	0.029	0.028	0.029	1.84	1.67	1.47	0.170	0.134	0.054	0.038	0.170	0.134	4		
11	0.025	0.030	0.035	1.66	1.40	1.47	0.147	0.477	0.046	0.066	0.147	0.477	4		
12	0.032	0.028	0.022	1.56	1.24	1.37	0.143	0.176	0.050	0.037	0.143	0.176	4		
13	0.026	0.028	0.040	1.59	1.54	1.38	0.026	0.419	0.041	0.050	0.026	0.419	3		
14	0.027	0.028	0.048	2.14	1.41	1.33	0.025	0.152	0.059	0.036	0.025	0.152	3		
15	0.031	0.027	0.044	1.69	1.47	1.43	0.121	0.312	0.055	0.048	0.121	0.312	4		
16	0.029	0.033	0.045	1.43	1.42	1.44	0.162	0.354	0.067	0.061	0.162	0.354	4		
17	0.028	0.029	0.033	1.54	1.31	1.48	0.118	0.234	0.056	0.042	0.118	0.234	4		
18	0.028	0.029	0.012	1.59	1.58	1.32	0.027	0.377	0.026	0.047	0.027	0.377	3		
19	0.026	0.029	0.029	1.87	1.24	1.37	0.074	0.315	0.044	0.045	0.074	0.315	4		
20	0.033	0.031	0.021	2.39	1.41	1.49	0.145	0.127	0.047	0.037	0.145	0.127	4		
21	0.030	0.031	0.017	1.60	1.38	1.39	0.024	0.054	0.046	0.033	0.024	0.054	3		
22	0.025	0.032	0.010	1.46	1.46	1.30	0.098	0.412	0.031	0.055	0.098	0.412	4		
23	0.029	0.028	0.046	1.64	1.59	1.30	0.150	0.256	0.060	0.049	0.150	0.256	4		
24	0.031	0.026	0.044	2.16	1.57	1.50	0.040	0.360	0.038	0.044	0.040	0.360	4		
25	0.028	0.030	0.022	2.05	1.54	1.24	0.132	0.093	0.051	0.038	0.132	0.093	4		
26	0.033	0.027	0.019	1.64	1.54	1.35	0.140	0.384	0.033	0.048	0.140	0.384	4		
27	0.033	0.028	0.025	2.12	1.23	1.43	0.125	0.488	0.034	0.058	0.125	0.488	4		
28	0.026	0.030	0.024	2.06	1.29	1.46	0.050	0.497	0.033	0.062	0.050	0.497	3		
29	0.028	0.032	0.021	2.25	1.65	1.26	0.117	0.419	0.027	0.059	0.117	0.419	4		
30	0.032	0.027	0.014	2.15	1.45	1.40	0.007	0.336	0.023	0.041	0.007	0.336	3		
31	0.029	0.028	0.021	2.30	1.32	1.24	0.007	0.416	0.026	0.048	0.007	0.416	3		
32	0.031	0.032	0.028	1.98	1.46	1.34	0.143	0.166	0.054	0.044	0.143	0.166	4		
33	0.025	0.030	0.012	2.31	1.67	1.36	0.160	0.044	0.045	0.032	0.160	0.044	4		
34	0.027	0.025	0.018	1.66	1.25	1.24	0.122	0.099	0.045	0.031	0.122	0.099	3		
35	0.029	0.032	0.024	2.00	1.36	1.38	0.141	0.197	0.049	0.044	0.141	0.197	4		
36	0.029	0.031	0.012	2.01	1.20	1.45	0.006	0.367	0.027	0.049	0.006	0.367	3		
37	0.026	0.028	0.047	2.07	1.60	1.37	0.155	0.070	0.076	0.041	0.155	0.070	4		
38	0.032	0.027	0.027	1.84	1.60	1.42	0.161	0.397	0.036	0.051	0.161	0.397	4		
39	0.033	0.032	0.043	1.88	1.32	1.38	0.021	0.412	0.047	0.055	0.021	0.412	3		
40	0.026	0.033	0.037	1.83	1.66	1.34	0.022	0.390	0.038	0.055	0.022	0.390	3		
41	0.033	0.032	0.035	1.86	1.24	1.26	0.073	0.049	0.068	0.038	0.073	0.049	3		
42	0.027	0.026	0.031	2.14	1.37	1.33	0.113	0.069	0.057	0.033	0.113	0.069	3		
43	0.033	0.027	0.022	1.95	1.41	1.28	0.126	0.217	0.041	0.039	0.126	0.217	4		
44	0.030	0.026	0.014	2.34	1.39	1.22	0.119	0.317	0.027	0.041	0.119	0.317	4		
45	0.029	0.028	0.021	2.06	1.38	1.22	0.048	0.480	0.026	0.055	0.048	0.480	4		
46	0.026	0.027	0.031	1.77	1.67	1.25	0.143	0.439	0.036	0.057	0.143	0.439	4		
47	0.029	0.032	0.031	2.09	1.23	1.40	0.030	0.337	0.042	0.050	0.030	0.337	3		
48	0.028	0.029	0.025	1.67	1.22	1.35	0.160	0.303	0.049	0.047	0.160	0.303	4		
49	0.033	0.033	0.040	1.83	1.62	1.47	0.041	0.375	0.042	0.055	0.041	0.375	4		
50	0.029	0.030	0.045	2.23	1.22	1.39	0.044	0.101	0.067	0.037	0.044	0.101	3		

infection. Subsequent clinical trials motivated in large part by the modeling study have confirmed the prevention benefit of early initiation of ART, showing up to a 96% reduction in HIV transmission [7, 3, 19, 8].

5.1. The model. We consider a modified version of the *Granich et al.* model which tracks the number of susceptible individuals, S , and infected individuals I_1, I_2, I_3, I_4

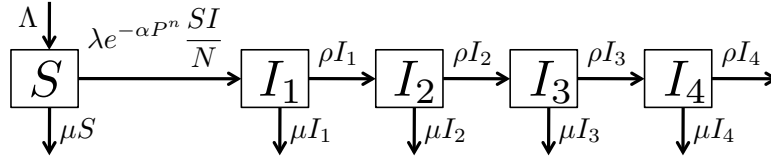


FIGURE 4. Flow diagram of the modified *Granich et al.* HIV model. $I = I_1 + I_2 + I_3 + I_4$, $N = S + I_1 + I_2 + I_3 + I_4$ and $P = \frac{I}{N}$.

as follows:

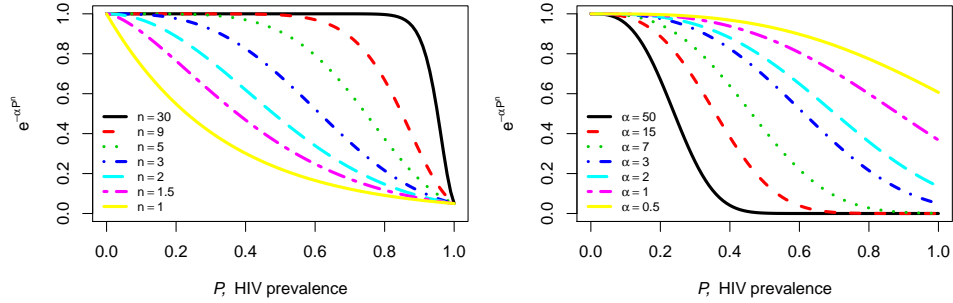
$$\begin{aligned}
 \frac{dS}{dt} &= \Lambda - \frac{\lambda e^{-\alpha P^n} SI}{N} - \mu S, \\
 \frac{dI_1}{dt} &= \frac{\lambda e^{-\alpha P^n} SI}{N} - \rho I_1 - \mu I_1, \\
 \frac{dI_2}{dt} &= \rho I_1 - \rho I_2 - \mu I_2, \\
 \frac{dI_3}{dt} &= \rho I_2 - \rho I_3 - \mu I_3, \\
 \frac{dI_4}{dt} &= \rho I_3 - \rho I_4 - \mu I_4,
 \end{aligned} \tag{12}$$

where $I = I_1 + I_2 + I_3 + I_4$ is total number of HIV-infected individuals, $N = S + I_1 + I_2 + I_3 + I_4$ is the total population size and $P = I/N$ is the prevalence of HIV. We have modified the original model of *Granich et al.* by replacing density depending recruitment with constant recruitment, Λ , and by neglecting classes of individuals on ART.

Notably, infectivity is the same in each of the four infectious classes. Hence, the infectious classes do not represent the stages of HIV infection (e.g. primary, chronic and AIDS) but rather are used to produce a gamma distribution for the survival of HIV-infected individuals that approximates the Weibull survival distribution that has been observed in data [11].

Another defining characteristic of the *Granich et al.* model is the transmission term given by $\lambda e^{-\alpha P^n}$ where P is the prevalence of HIV infection at a given time. As *Granich et al.* describe, this allows for the population's risk behavior to decrease as disease prevalence increases. Figure 5 illustrates the form of the population's response to disease prevalence. In [11], λ , α and n are referred to as the "initial value", "location" and "shape" of the transmission term, respectively. From a practical modeling perspective, this formulation allows the model to exhibit the rapid initial takeoff of the epidemic observed in several sub-Saharan African settings (e.g. $\mathcal{R}_0 \approx 8$), followed by a relatively modest disease prevalence at endemic equilibrium (e.g. $< 20\%$). In a standard formulation with fixed transmission coefficient, prevalence at endemic equilibrium is typically $1 - 1/\mathcal{R}_0$.

5.2. Calibration to HIV prevalence at equilibrium. We begin by calibrating the model in (12) to a given HIV prevalence at endemic equilibrium which we denote by \hat{A} . To satisfy this single calibration condition, we use the initial value of the transmission term, λ , as a calibration parameter. We assume that values for the remaining model parameters have been established.



(a) $e^{-\alpha P^n}$ with varying values of n and fixing $\alpha = 3$. (b) $e^{-\alpha P^n}$ with varying values of α and fixing $n = 3$.

FIGURE 5. Illustration of population's behavioral response to HIV prevalence in the *Granich et al.* model via the transmission term $e^{-\alpha P^n}$. Notably, if $n = 1$ sexual risk behavior decreases exponentially as a function of HIV prevalence. As $n \rightarrow \infty$, the decline in sexual risk behavior approaches a step function.

As in our calibration of the four compartment HIV/HSV-2 model in Section 4.2, we begin by rescaling the state variables to represent proportions of the total population by letting $s = \frac{S}{N}$, $i_1 = \frac{I_1}{N}$, $i_2 = \frac{I_2}{N}$, $i_3 = \frac{I_3}{N}$ and $i_4 = \frac{I_4}{N}$. As $s = 1 - i_1 - i_2 - i_3 - i_4$, the model (12) reduces to

$$\begin{aligned} \frac{di_1}{dt} &= \lambda e^{-\alpha i^n} (1 - i)i - \rho i_1 - \mu i_1, \\ \frac{di_2}{dt} &= \rho i_1 - \rho i_2 - \mu i_2, \\ \frac{di_3}{dt} &= \rho i_2 - \rho i_3 - \mu i_3, \\ \frac{di_4}{dt} &= \rho i_3 - \rho i_4 - \mu i_4, \end{aligned} \tag{13}$$

where $i = i_1 + i_2 + i_3 + i_4$. Combing the four conditions that ensure the disease is at equilibrium (14)–(17) with a calibration condition (18), we have the following calibration system of equations

$$0 = \lambda e^{-\alpha i^n} (1 - i)i - \rho i_1 - \mu i_1, \tag{14}$$

$$0 = \rho i_1 - \rho i_2 - \mu i_2, \tag{15}$$

$$0 = \rho i_2 - \rho i_3 - \mu i_3, \tag{16}$$

$$0 = \rho i_3 - \rho i_4 - \mu i_4, \tag{17}$$

$$0 = i_1 + i_2 + i_3 + i_4 - \hat{A}. \tag{18}$$

To complete our calibration, we use Newton's Method to find the solution $(i_1^*, i_2^*, i_3^*, i_4^*, \lambda^*)$ of the calibration system (14)–(18) using the following initial approximation:

$$i_{1,0} = \frac{\hat{A}}{4}, \quad i_{2,0} = \frac{\hat{A}}{4}, \quad i_{3,0} = \frac{\hat{A}}{4}, \quad i_{4,0} = \frac{\hat{A}}{4}, \quad \lambda_0 = \frac{\mu + \rho}{4e^{-\alpha \hat{A}^n} (1 - \hat{A})}. \tag{19}$$

Here, our initial estimate λ_0 results from solving (14) with the initial approximations $i = \hat{A}$ and of $i_1 = \hat{A}/4$.

5.2.1. *Demonstration of calibration to HIV prevalence.* In this section, we use our calibration technique to parameterize the *Granich et al.* HIV model to HIV prevalence at equilibrium. In Table 4(a), we summarize model parameters and establish baseline values and ranges. Baseline values are those from [11] that resulted from fitting the full version of the model to South Africa's HIV epidemic. To demonstrate the robustness of our approach, we calibrate the model for random sets of model parameters and calibration conditions. Specifically, we obtain 50 independent samples of parameters and calibration conditions from uniform distributions over the ranges given in Table 4(a). The calibration system given by (14)–(18) was then solved using Newton's Method using the initial conditions in (19).

For each sample, the calibrated value of λ was noted along with the resulting HIV prevalence and the necessary number of iterations of Newton's Method. The results are shown in Table 4(b). For all 50 samples, the calibration approach successfully matched the calibration condition (i.e. HIV prevalence) while requiring no more than 2 iterations of Newton's Method.

6. HIV/HSV-2 coinfection model with behavioral response. In this section, we consider an HIV/HSV-2 coinfection model that combines the interactions of HIV and HSV-2 infection from the model in Section 4.1 and the behavioral response to the HIV epidemic of the *Granich et al.* model described in Section 5.1.

6.1. **The model.** The model includes ten classes of individuals: S , susceptible to infection by both HIV and HSV-2; H , infected with HSV-2 only; A_1, A_2, A_3, A_4 , infected with HIV only; C_1, C_2, C_3, C_4 , co-infected with both HIV and HSV-2. The model equations are given by

$$\begin{aligned}
\frac{dS}{dt} &= \Lambda - \Omega S - \Psi S - \mu S, \\
\frac{dH}{dt} &= \Psi S - r_1 \Omega H - \mu H, \\
\frac{dA_1}{dt} &= \Omega S - \Psi A_1 - \rho A_1 - \mu A_1, \\
\frac{dA_2}{dt} &= \rho A_1 - \Psi A_2 - \rho A_2 - \mu A_2, \\
\frac{dA_3}{dt} &= \rho A_2 - \Psi A_3 - \rho A_3 - \mu A_3, \\
\frac{dA_4}{dt} &= \rho A_3 - \Psi A_4 - \rho A_4 - \mu A_4, \\
\frac{dC_1}{dt} &= r_1 \Omega H + \Psi A_1 - \rho C_1 - \mu C_1, \\
\frac{dC_2}{dt} &= \rho C_1 + \Psi A_2 - \rho C_2 - \mu C_2, \\
\frac{dC_3}{dt} &= \rho C_2 + \Psi A_3 - \rho C_3 - \mu C_3, \\
\frac{dC_4}{dt} &= \rho C_3 + \Psi A_4 - \rho C_4 - \mu C_4,
\end{aligned} \tag{20}$$

TABLE 4. Calibration of *Granich et al.* model to HIV prevalence.

(a) Model parameters and calibration conditions for parameterizing the *Granich et al.* model with behavioral response to HIV prevalence at endemic equilibrium. [†]Baseline value taken from [11]; range is modeling assumption of this work. [‡]Recruitment rate range is same as background mortality to give a total population of 1 in the absence of HIV, without loss of generality.

	Symbol	Baseline	Range	References
Model Parameters				
Background mortality rate (/yr)	μ	0.018	$\frac{1}{60} - \frac{1}{30}$	[11] [†]
Recruitment rate into the population (/yr)	Λ	0.018	$\frac{1}{60} - \frac{1}{30}$	[‡]
Location of transmission term	α	0.055	0.05 – 50	[11] [†]
Shape of transmission term	n	0.996	0.90 – 30	[11] [†]
Rate of HIV progression (/yr)	ρ	0.303	0.1 – 0.4	[11] [†]
Calibration Condition				
HIV prevalence at endemic equilibrium	\hat{A}	17	0.3 – 17	[18]
Calibration Parameter				
Initial value of transmission term (/yr)	λ	determined by calibration procedure		

(b) Results of calibrating the *Granich et al.* model to 50 independent sets of parameters and calibration conditions (randomly sampled from ranges above). For each sample $j = 1, 2, \dots, 50$, the resulting values of the calibration parameter (λ^*) and the resulting HIV prevalence at equilibrium (\hat{A}^*) are given. The number of iterations of Newton’s Method required to achieve a stopping criterion of $|\lambda_i - \lambda_{i-1}| < 10^{-9}$ is denoted by m .

j	Sampled parameters and calibration condition					Calibration results			j	Sampled parameters and calibration condition					Calibration results		
	Λ	μ	α	n	\hat{A}	λ^*	\hat{A}^*	m		Λ	μ	α	n	\hat{A}	λ^*	\hat{A}^*	m
1	0.018	0.021	34.3	29.1	0.115	0.104	0.115	2	26	0.020	0.027	36.7	21.5	0.112	0.085	0.112	2
2	0.028	0.024	31.7	14.6	0.019	0.065	0.019	2	27	0.020	0.020	24.4	7.2	0.088	0.049	0.088	2
3	0.022	0.025	5.8	4.9	0.126	0.066	0.126	2	28	0.021	0.026	49.6	25.3	0.154	0.123	0.154	2
4	0.021	0.019	30.2	6.1	0.142	0.044	0.142	2	29	0.030	0.026	9.8	12.5	0.100	0.072	0.100	2
5	0.028	0.017	35.9	5.7	0.006	0.097	0.006	2	30	0.021	0.019	2.4	22.5	0.058	0.106	0.058	2
6	0.026	0.028	8.7	22.0	0.015	0.049	0.015	2	31	0.022	0.026	4.6	23.2	0.098	0.101	0.098	2
7	0.031	0.018	6.0	27.3	0.061	0.087	0.061	2	32	0.026	0.022	46.7	17.7	0.108	0.090	0.108	2
8	0.030	0.019	29.9	17.2	0.014	0.077	0.014	2	33	0.029	0.024	5.8	9.6	0.120	0.125	0.120	2
9	0.024	0.026	32.4	6.3	0.073	0.078	0.073	2	34	0.023	0.030	9.5	17.1	0.027	0.087	0.027	2
10	0.022	0.022	47.6	18.5	0.125	0.098	0.125	2	35	0.025	0.024	0.7	23.5	0.111	0.117	0.111	2
11	0.022	0.030	25.0	19.3	0.160	0.124	0.160	2	36	0.021	0.030	16.0	21.0	0.023	0.101	0.023	2
12	0.023	0.030	49.1	10.6	0.114	0.059	0.114	2	37	0.027	0.018	29.4	6.3	0.165	0.123	0.165	2
13	0.032	0.017	17.4	14.4	0.061	0.041	0.061	2	38	0.025	0.033	48.2	6.1	0.071	0.096	0.071	2
14	0.025	0.030	9.0	3.0	0.140	0.063	0.140	2	39	0.027	0.030	12.0	22.4	0.057	0.071	0.057	2
15	0.020	0.022	8.3	14.9	0.093	0.126	0.093	2	40	0.025	0.033	43.6	4.3	0.017	0.059	0.017	2
16	0.024	0.028	24.0	23.0	0.060	0.077	0.060	2	41	0.031	0.031	1.7	13.0	0.151	0.075	0.151	2
17	0.029	0.032	8.0	24.5	0.023	0.108	0.023	2	42	0.031	0.031	26.1	11.1	0.041	0.077	0.041	2
18	0.025	0.025	25.3	15.1	0.119	0.109	0.119	2	43	0.022	0.024	39.9	16.9	0.100	0.100	0.100	2
19	0.018	0.031	18.8	29.1	0.126	0.084	0.126	2	44	0.029	0.022	17.7	13.2	0.103	0.061	0.103	2
20	0.031	0.026	41.9	24.7	0.075	0.107	0.075	2	45	0.023	0.024	6.7	21.4	0.158	0.117	0.158	2
21	0.027	0.018	22.6	18.3	0.131	0.102	0.131	2	46	0.028	0.023	4.2	10.5	0.082	0.051	0.082	2
22	0.027	0.019	23.9	12.6	0.106	0.094	0.106	2	47	0.023	0.031	13.1	18.4	0.022	0.084	0.022	2
23	0.031	0.026	28.6	10.0	0.086	0.095	0.086	2	48	0.031	0.022	46.3	28.3	0.058	0.118	0.058	2
24	0.019	0.027	45.3	6.9	0.067	0.075	0.067	2	49	0.020	0.025	42.7	27.7	0.077	0.112	0.077	2
25	0.025	0.033	42.0	7.9	0.040	0.102	0.040	2	50	0.024	0.030	46.1	2.2	0.082	0.085	0.082	2

where $\Omega = \lambda e^{-\alpha P^n} \frac{A+r_2C}{N}$, $\Psi = \kappa e^{-\alpha P^n} \frac{H+r_3C}{N}$, $A = A_1 + A_2 + A_3 + A_4$ is the number of people infected with HIV only, $C = C_1 + C_2 + C_3 + C_4$ is the number of coinfecting individuals, $N = S + H + A + C$ is the size of the total population and $P = \frac{A+C}{N}$ is the prevalence of HIV. A flow diagram of the model is given in Figure 6.

An important implicit assumption is made by using the same transmission formulation for HIV and HSV-2 (i.e. $\Omega = \lambda e^{-\alpha P^n} \frac{A+r_2C}{N}$ and $\Psi = \kappa e^{-\alpha P^n} \frac{H+r_3C}{N}$). Specifically, we assume that the behavioral response only occurs in response to the HIV epidemic (i.e. $e^{-\alpha P^n}$ term depends only on HIV prevalence) and that

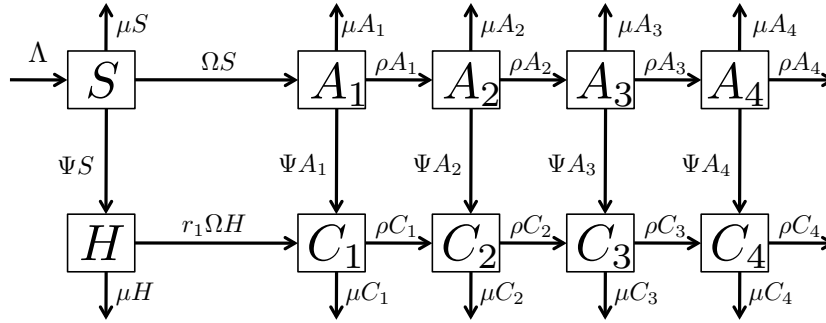


FIGURE 6. Flow diagram of the HIV/HSV-2 coinfection model with behavioral response. The rate of transmission for HIV is given by $\Omega = \lambda e^{-\alpha P^n} \frac{A+r_2 C}{N}$ where $A = A_1 + A_2 + A_3 + A_4$ is the number of people infected with HIV only, $C = C_1 + C_2 + C_3 + C_4$ is the number of coinfecting individuals, $N = S + H + A + C$ is the size of the total population and $P = \frac{A+C}{N}$ is the prevalence of HIV. The rate of transmission for HSV-2 is given by $\Psi = \kappa e^{-\alpha P^n} \frac{H+r_3 C}{N}$.

the behavioral response to the HIV epidemic (condom use, partner reduction and other safe-sex practices) would have a corresponding effect on HSV-2 transmission (i.e. $e^{-\alpha P^n}$ is included in the HSV-2 transmission terms as well).

6.2. Calibration to HIV and HSV-2 prevalence at equilibrium. In this section, we calibrate the expanded HIV/HSV-2 coinfection model with behavioral response to match both HIV and HSV-2 prevalence at the endemic equilibrium (as was done for the four-compartment coinfection model in Section 4.2) using the initial values of the HIV and HSV-2 transmission terms (λ and κ , respectively) as calibration parameters. The calibration conditions are given by

$$\begin{aligned}\widehat{A} &= \text{HIV prevalence at endemic equilibrium,} \\ \widehat{H} &= \text{HSV-2 prevalence at endemic equilibrium.}\end{aligned}$$

As in our previous calibration efforts, we begin by rescaling the state variables to represent proportions of the total population by letting $s = \frac{S}{N}$, $h = \frac{H}{N}$, $a_1 = \frac{A_1}{N}$, $a_2 = \frac{A_2}{N}$, $a_3 = \frac{A_3}{N}$, $a_4 = \frac{A_4}{N}$, $c_1 = \frac{C_1}{N}$, $c_2 = \frac{C_2}{N}$, $c_3 = \frac{C_3}{N}$ and $c_4 = \frac{C_4}{N}$. Correspondingly, we will let $a = a_1 + a_2 + a_3 + a_4$ and $c = c_1 + c_2 + c_3 + c_4$. As $s = 1 - h - a - c$, we can reduce the dimension of the in model (20) to get

$$\begin{aligned}\frac{dh}{dt} &= \Psi(1 - h - a - c) - r_1 \Omega h - \mu h, \\ \frac{da_1}{dt} &= \Omega(1 - h - a - c) - \Psi a_1 - \rho a_1 - \mu a_1, \\ \frac{da_2}{dt} &= \rho a_1 - \Psi a_2 - \rho a_2 - \mu a_2, \\ \frac{da_3}{dt} &= \rho a_2 - \Psi a_3 - \rho a_3 - \mu a_3, \\ \frac{da_4}{dt} &= \rho a_3 - \Psi a_4 - \rho a_4 - \mu a_4,\end{aligned}$$

$$\begin{aligned}
\frac{dc_1}{dt} &= r_1\Omega h + \Psi a_1 - \rho c_1 - \mu c_1, \\
\frac{dc_2}{dt} &= \rho c_1 + \Psi a_2 - \rho c_2 - \mu c_2, \\
\frac{dc_3}{dt} &= \rho c_2 + \Psi a_3 - \rho c_3 - \mu c_3, \\
\frac{dc_4}{dt} &= \rho c_3 + \Psi a_4 - \rho c_4 - \mu c_4,
\end{aligned} \tag{21}$$

where the transmission terms have become $\Omega = \lambda e^{-\alpha P^n} (a+r_2c)$ and $\Psi = \kappa e^{-\alpha P^n} (h+r_3c)$.

Combing the nine conditions that ensure the disease is at equilibrium in (22)–(30) with the HIV and HSV-2 prevalence conditions in (31) and (32), respectively, we have the following calibration system

$$0 = \Psi(1 - h - a - c) - r_1\Omega h - \mu h, \tag{22}$$

$$0 = \Omega(1 - h - a - c) - \Psi a_1 - \rho a_1 - \mu a_1, \tag{23}$$

$$0 = \rho a_1 - \Psi a_2 - \rho a_2 - \mu a_2, \tag{24}$$

$$0 = \rho a_2 - \Psi a_3 - \rho a_3 - \mu a_3, \tag{25}$$

$$0 = \rho a_3 - \Psi a_4 - \rho a_4 - \mu a_4, \tag{26}$$

$$0 = r_1\Omega h + \Psi a_1 - \rho c_1 - \mu c_1, \tag{27}$$

$$0 = \rho c_1 + \Psi a_2 - \rho c_2 - \mu c_2, \tag{28}$$

$$0 = \rho c_2 + \Psi a_3 - \rho c_3 - \mu c_3, \tag{29}$$

$$0 = \rho c_3 + \Psi a_4 - \rho c_4 - \mu c_4, \tag{30}$$

$$0 = a_1 + a_2 + a_3 + a_4 + c_1 + c_2 + c_3 + c_4 - \widehat{A}, \tag{31}$$

$$0 = h + c_1 + c_2 + c_3 + c_4 - \widehat{H}. \tag{32}$$

In order to complete the calibration, we use Newton's Method to find the solution $(h^*, a_1^*, a_2^*, a_3^*, a_4^*, c_1^*, c_2^*, c_3^*, c_4^*, \lambda^*, \kappa^*)$ of the calibration system (22)–(32). For the initial approximation to the solution of the calibration system required by Newton's Method, we again make the simplifying assumption that HIV and HSV-2 infection are independent (i.e. that the total number of coinfecting people is $c_0 = \widehat{A}\widehat{H}$, total number infected with HIV alone is $a_0 = \widehat{A} - \widehat{A}\widehat{H}$ and the total number infected with HSV-2 alone is $h_0 = \widehat{H} - \widehat{A}\widehat{H}$) and distribute HIV-infected individuals evenly among the four infected stages. To obtain initial approximations for λ and κ , we solve the 2-dimensional system given by (22) and (23) for λ and κ . The resulting initial approximation is given by:

$$h_0 = \widehat{H} - \widehat{A}\widehat{H}, \quad a_{1,0} = a_{2,0} = a_{3,0} = a_{4,0} = \frac{\widehat{A} - \widehat{A}\widehat{H}}{4}, \quad c_{1,0} = c_{2,0} = c_{3,0} = c_{4,0} = \frac{\widehat{A}\widehat{H}}{4}, \tag{33}$$

$$\lambda_0 = \frac{a_{1,0} (\mu + \rho - a_0\mu - a_0\rho - c_0\mu - c_0\rho - h_0\rho) e^{\alpha\widehat{A}^n}}{(a_0^2 - a_{1,0}h_0r_1 + 2a_0c_0 + 2a_0h_0 + c_0^2 + 2c_0h_0 + h_0^2 - 2a_0 - 2c_0 - 2h_0 + 1)(r_2c_0 + a_0)} \tag{34}$$

and

$$\kappa_0 = \frac{-h_0 (-a_{1,0}\mu r_1 - a_{1,0}r_1\rho + a_0\mu + c_0\mu + \mu h_0 - \mu) e^{\alpha\hat{A}^n}}{(a_0^2 - a_{1,0}h_0r_1 + 2a_0c_0 + 2a_0h_0 + c_0^2 + 2c_0h_0 + h_0^2 - 2a_0 - 2c_0 - 2h_0 + 1)(r_3c_0 + h_0)}. \quad (35)$$

6.3. Demonstration of calibration to HIV and HSV-2 prevalence. In this section, we show the robustness of our calibration technique to parameterize the HIV/HSV-2 coinfection model with behavioral response to HIV and HSV-2 prevalence at equilibrium. In Table 5(a), we summarize model parameter values and ranges. Demonstrating the robustness of our approach, we again calibrate the model for random sets of model parameters and calibration conditions. Specifically, we obtain 50 independent samples of parameters and calibration conditions from uniform distributions over the ranges given in Table 5(a). The calibration system given by (22)–(32) is then solved using Newton’s Method with initial conditions determined by (33)–(35).

For each sample, the calibrated value of λ and κ is noted along with the resulting HIV prevalence and the necessary number of iterations of Newton’s Method. The results are shown in Table 7(b). For all 50 samples, the calibration approach successfully matched the calibration conditions (i.e. HIV and HSV-2 prevalence) while requiring no more than 4 iterations of Newton’s Method.

7. Extensions of the calibration approach. In the previous demonstrations of our calibration approach, the mathematical model at hand was always calibrated to disease prevalence at endemic equilibrium. Fortunately, the approach can be generalized to achieve additional calibration conditions. In this section, we show how conditions that relate to the transient phase of an epidemic (the magnitude and timing of peak disease incidence, specifically) can be incorporated into our calibration.

7.1. Exact calibration of *Granich et al.* model to HIV prevalence at equilibrium, peak HIV incidence and timing of peak HIV incidence. In Section 5.2, we calibrated the *Granich et al.* model as stated in (12) to HIV prevalence at equilibrium using λ as a calibration parameter and assuming that values for α and n had been established *a priori* (recall the transmission term of $\lambda e^{-\alpha P^n}$). As α and n determine the shape of the population’s behavioral response to HIV prevalence (see Figure 5), they can be used to incorporate additional detail into our calibration. In this section, we will describe the calibration of the model in (12) to the following three conditions:

$$\begin{aligned} \hat{A} &= \text{HIV prevalence at endemic equilibrium,} \\ \hat{I} &= \text{peak HIV incidence,} \\ \hat{P} &= \text{HIV prevalence at time of peak HIV incidence.} \end{aligned}$$

To do so, we use the initial value, location and shape of the HIV transmission term (i.e. λ, α and n) as calibration parameters. If the *Granich et al.* model explicitly included time (i.e. was nonautonomous), we could calibrate the system so that peak HIV incidence occurs at a specific time t^* . Since the *Granich et al.* model is autonomous, we use HIV prevalence at the time of peak HIV incidence, \hat{P} , as a proxy for the time of peak incidence to calibrate to different types of epidemic curves (e.g. early, mid or late peaking epidemics).

TABLE 5. Calibration of HIV/HSV-2 coinfection model with behavioral response to HIV and HSV-2 prevalence.

(a) Model parameters and calibration conditions for parameterizing the HIV/HSV-2 coinfection model with behavioral response to HIV and HSV-2 prevalence at endemic equilibrium. ‡Recruitment rate range is same as background mortality to give a total population of 1 in the absence of HIV, without loss of generality.

	Symbol	Range	References
Model Parameters			
Background mortality rate (/yr)	μ	$\frac{1}{60} - \frac{1}{30}$	[11]
Recruitment rate into the population (/yr)	Λ	$\frac{1}{60} - \frac{1}{30}$	‡
Location of HIV transmission term	α	0.05 – 50	[11]
Shape of HIV transmission term	n	0.90 – 30	[11]
Rate of HIV progression (/yr)	ρ	0.1 – 0.4	[11]
Calibration Conditions			
HIV prevalence at endemic equilibrium (/yr)	\hat{A}	0.3 – 17	[18]
HSV-2 prevalence at endemic equilibrium (/yr)	\hat{H}	4 – 50	[2, 16, 22, 20]
Calibration Parameters			
Initial value of HIV transmission term (/yr)	λ	determined by calibration procedure	
Initial value of HSV-2 transmission term (/yr)	κ	determined by calibration procedure	

(b) Results of calibrating the HIV/HSV-2 coinfection model with behavioral response to 50 independent sets of parameters and calibration conditions (randomly sampled from ranges above). For each sample $j = 1, 2, \dots, 50$, the resulting values of the calibration parameters (λ^*, κ^*) and the resulting diseases prevalences at equilibrium (\hat{A}^*, \hat{H}^*) are given. The number of iterations of Newton’s Method required to achieve a stopping criterion of $\max\{|\lambda_i - \lambda_{i-1}|, |\kappa_i - \kappa_{i-1}|\} < 10^{-9}$ is denoted by m .

Sampled parameters and calibration condition										Calibration results															
j	Λ	μ	ρ	α	n	\hat{A}	\hat{H}	λ^*	κ^*	\hat{A}^*	\hat{H}^*	m	j	Λ	μ	ρ	α	n	\hat{A}	\hat{H}	λ^*	κ^*	\hat{A}^*	\hat{H}^*	m
1	0.033	0.030	0.136	25.388	7.775	0.100	0.429	0.040	0.057	0.100	0.429	4	26	0.028	0.031	0.357	19.775	5.288	0.086	0.101	0.094	0.049	0.086	0.101	4
2	0.026	0.025	0.294	25.452	10.004	0.061	0.433	0.057	0.053	0.061	0.433	4	27	0.019	0.030	0.312	8.664	27.919	0.126	0.219	0.080	0.055	0.126	0.219	4
3	0.022	0.032	0.268	9.053	18.416	0.065	0.333	0.066	0.055	0.065	0.333	4	28	0.024	0.021	0.357	4.805	29.623	0.069	0.469	0.056	0.041	0.069	0.469	4
4	0.027	0.022	0.159	30.943	15.777	0.107	0.286	0.042	0.038	0.107	0.286	4	29	0.030	0.020	0.259	32.938	28.889	0.099	0.436	0.052	0.049	0.099	0.436	4
5	0.020	0.032	0.184	35.273	19.612	0.025	0.484	0.033	0.063	0.025	0.484	4	30	0.018	0.017	0.182	4.482	2.045	0.027	0.288	0.041	0.026	0.027	0.288	4
6	0.019	0.017	0.363	46.871	29.933	0.081	0.050	0.105	0.029	0.081	0.050	4	31	0.028	0.026	0.164	26.525	25.949	0.117	0.084	0.058	0.035	0.117	0.084	4
7	0.021	0.026	0.164	35.976	10.444	0.050	0.437	0.034	0.050	0.050	0.437	4	32	0.023	0.020	0.372	39.113	5.546	0.136	0.443	0.075	0.064	0.136	0.443	4
8	0.017	0.023	0.149	45.299	27.869	0.013	0.349	0.031	0.035	0.013	0.349	4	33	0.032	0.022	0.353	1.944	19.318	0.082	0.353	0.084	0.047	0.082	0.353	4
9	0.018	0.022	0.122	2.960	2.103	0.161	0.067	0.050	0.031	0.161	0.067	4	34	0.029	0.031	0.110	2.429	4.162	0.059	0.432	0.026	0.056	0.059	0.432	4
10	0.020	0.025	0.210	47.955	15.737	0.099	0.044	0.072	0.034	0.099	0.044	4	35	0.027	0.024	0.165	47.290	21.058	0.041	0.089	0.051	0.029	0.041	0.089	4
11	0.018	0.021	0.120	22.432	23.321	0.114	0.424	0.028	0.042	0.114	0.424	4	36	0.025	0.022	0.348	17.164	26.863	0.063	0.455	0.076	0.052	0.063	0.455	4
12	0.032	0.023	0.155	27.625	4.781	0.107	0.234	0.044	0.037	0.107	0.234	4	37	0.017	0.024	0.361	9.553	7.454	0.027	0.054	0.103	0.029	0.027	0.054	3
13	0.029	0.027	0.324	36.150	6.654	0.087	0.303	0.063	0.051	0.087	0.303	4	38	0.027	0.022	0.222	46.392	21.362	0.167	0.320	0.051	0.051	0.167	0.320	4
14	0.018	0.020	0.144	29.338	26.770	0.008	0.357	0.027	0.032	0.008	0.357	4	39	0.028	0.018	0.247	21.006	7.505	0.115	0.465	0.046	0.048	0.115	0.465	4
15	0.019	0.032	0.177	42.820	16.166	0.168	0.403	0.084	0.086	0.168	0.403	4	40	0.018	0.026	0.143	34.607	19.392	0.022	0.103	0.046	0.030	0.022	0.103	4
16	0.028	0.017	0.126	15.874	24.707	0.160	0.237	0.041	0.030	0.160	0.237	4	41	0.021	0.033	0.203	46.859	24.756	0.139	0.184	0.068	0.049	0.139	0.184	4
17	0.021	0.024	0.275	14.494	1.718	0.097	0.111	0.111	0.047	0.097	0.111	4	42	0.028	0.021	0.251	22.415	11.566	0.026	0.043	0.072	0.025	0.026	0.043	4
18	0.032	0.026	0.254	14.086	11.617	0.085	0.231	0.053	0.044	0.085	0.231	4	43	0.021	0.020	0.390	14.421	26.982	0.078	0.408	0.079	0.049	0.078	0.408	4
19	0.020	0.030	0.300	18.883	1.811	0.022	0.358	0.054	0.051	0.022	0.358	4	44	0.018	0.017	0.113	4.843	4.594	0.037	0.277	0.028	0.026	0.037	0.277	4
20	0.024	0.024	0.274	5.479	9.004	0.161	0.068	0.091	0.043	0.161	0.068	4	45	0.020	0.028	0.376	2.375	17.141	0.054	0.302	0.083	0.048	0.054	0.302	4
21	0.033	0.024	0.149	26.541	10.778	0.003	0.325	0.032	0.035	0.003	0.325	4	46	0.029	0.023	0.125	2.299	25.985	0.052	0.456	0.024	0.044	0.052	0.456	4
22	0.024	0.031	0.173	0.848	2.736	0.019	0.206	0.050	0.040	0.019	0.206	4	47	0.028	0.033	0.353	34.218	7.959	0.010	0.207	0.081	0.042	0.010	0.207	4
23	0.017	0.020	0.147	35.029	14.683	0.102	0.117	0.046	0.028	0.102	0.117	4	48	0.018	0.033	0.222	38.289	27.674	0.134	0.430	0.046	0.072	0.134	0.430	4
24	0.026	0.030	0.121	23.521	10.510	0.114	0.340	0.036	0.049	0.114	0.340	4	49	0.030	0.022	0.272	25.403	9.279	0.021	0.315	0.054	0.052	0.021	0.315	4
25	0.029	0.030	0.169	35.260	8.151	0.122	0.496	0.035	0.070	0.122	0.496	4	50	0.023	0.030	0.119	2.163	27.532	0.135	0.126	0.051	0.037	0.135	0.126	4

To create our calibration system, we note that HIV incidence (i.e. the rate of new HIV infections) is given by $\lambda e^{-\alpha i^n} (1-i)i$. To guarantee that HIV incidence is \hat{I} at the time when HIV prevalence is \hat{P} , we require that $\lambda e^{-\alpha \hat{P}^n} (1-\hat{P})\hat{P} = \hat{I}$. To ensure that HIV incidence is maximized at this time, we require that $\frac{d}{dP} (\lambda e^{-\alpha P^n} (1-P)P) = \lambda e^{-\alpha \hat{P}^n} (-\alpha \hat{P}^n n (1-\hat{P}) + 1 - 2\hat{P}) = 0$. Adding these two additional conditions (noting that $\lambda e^{-\alpha \hat{P}^n} \neq 0$), we arrive at the following calibration system of equations:

$$0 = \lambda e^{-\alpha i^n} (1-i)i - \rho i_1 - \mu i_1, \tag{36}$$

$$0 = \rho i_1 - \rho i_2 - \mu i_2, \tag{37}$$

$$0 = \rho i_2 - \rho i_3 - \mu i_3, \tag{38}$$

$$0 = \rho i_3 - \rho i_4 - \mu i_4, \quad (39)$$

$$0 = i_1 + i_2 + i_3 + i_4 - \hat{A}, \quad (40)$$

$$0 = \lambda e^{-\alpha \hat{P}^n} (1 - \hat{P}) \hat{P} - \hat{I}, \quad (41)$$

$$0 = 1 - \alpha \hat{P}^n n (1 - \hat{P}) - 2\hat{P}. \quad (42)$$

To reiterate, the calibration system imposes that the disease is at equilibrium in (36)–(39), that the HIV prevalence at endemic equilibrium is \hat{A} in (40), that HIV incidence is \hat{I} at the time when HIV prevalence is \hat{P} in (41) and that HIV incidence is maximized at the time when HIV prevalence is \hat{P} in (42).

In order to complete the calibration, we use Newton's Method to find the solution $(i_1^*, i_2^*, i_3^*, i_4^*, \lambda^*, \alpha^*, n^*)$ of the calibration system (36)–(42). Unlike the previous calibration examples, establishing a suitably robust initial approximation to the root is nontrivial. To do so, we note that at equilibrium we must have that $i = \hat{A}$, so from (36) and (41), we have that

$$\lambda e^{-\alpha \hat{A}^n} = \frac{i_1(\rho + \mu)}{(1 - \hat{A})\hat{A}} \quad \text{and} \quad \lambda e^{-\alpha \hat{P}^n} = \frac{\hat{I}}{(1 - \hat{P})\hat{P}}, \quad (43)$$

which implies that

$$e^{-\alpha(\hat{A}^n - \hat{P}^n)} = \frac{i_1(\rho + \mu)(1 - \hat{P})\hat{P}}{(1 - \hat{A})\hat{A}\hat{I}}. \quad (44)$$

Solving (42) for α , we get that

$$\alpha = \frac{1 - 2\hat{P}}{n\hat{P}^n(1 - \hat{P})}. \quad (45)$$

Substituting (45) into (44), we get

$$e^{-\frac{(1-2\hat{P})(\hat{A}^n - \hat{P}^n)}{n\hat{P}^n(1-\hat{P})}} - \frac{i_1(\rho + \mu)(1 - \hat{P})\hat{P}}{(1 - \hat{A})\hat{A}\hat{I}} = e^{-\frac{(1-2\hat{P})((\hat{A}/\hat{P})^n - 1)}{n(1-\hat{P})}} - \frac{i_1(\rho + \mu)(1 - \hat{P})\hat{P}}{(1 - \hat{A})\hat{A}\hat{I}} = 0 \quad (46)$$

which can be solved numerically to obtain an initial approximation for n . Once an initial approximation n_0 is established, (45) and (43) are used to obtain an initial approximation for α and λ , respectively. Initial approximations for i_1, i_2, i_3 and i_4 can be found by noting that Equations (37)–(40) consist of 4 linear conditions on i_1, i_2, i_3, i_4 .

To summarize, we calibrate our model to HIV prevalence and peak incidence by using Newton's Method to find the solution $(i_1^*, i_2^*, i_3^*, i_4^*, \lambda^*, \alpha^*, n^*)$ of the calibration system (36)–(42) using the following initial approximation:

$$i_{1,0} = \frac{\hat{A}(\rho + \mu)^3}{(2\rho + \mu)(2\rho^2 + 2\rho\mu + \mu^2)}, \quad i_{2,0} = \frac{(\rho + \mu)^2 \rho \hat{A}}{(2\rho + \mu)(2\rho^2 + 2\rho\mu + \mu^2)} \quad (47)$$

$$i_{3,0} = \frac{(\rho + \mu) \rho^2 \hat{A}}{(2\rho + \mu)(2\rho^2 + 2\rho\mu + \mu^2)}, \quad i_{4,0} = \frac{\rho^3 \hat{A}}{(2\rho + \mu)(2\rho^2 + 2\rho\mu + \mu^2)}, \quad (48)$$

$$n_0 = \text{numerical solution to (46)}, \quad \alpha_0 = \frac{1 - 2\hat{P}}{n_0 \hat{P}^{n_0} (1 - \hat{P})}, \quad \lambda_0 = \frac{e^{\alpha \hat{A}^{n_0}} (\rho + \mu) i_{1,0}}{(1 - \hat{A}) \hat{A}}. \quad (49)$$

7.1.1. *Demonstration of calibration to HIV prevalence, peak HIV incidence and timing of peak HIV incidence.* In this section, we use our calibration technique to parameterize the *Granich et al.* HIV model to HIV prevalence at equilibrium, peak HIV incidence and timing of peak incidence. To demonstrate the robustness of our approach, we calibrate the model for random sets of model parameters and calibration conditions. Specifically, we obtain 50 independent samples of parameters and calibration conditions from uniform distributions over the ranges specified in Table 7(a). Notably, restrictions were placed on the peak HIV incidence, \hat{I} , and the HIV prevalence at time of peak HIV incidence, \hat{P} , for epidemiological plausibility. Specifically, we require peak HIV incidence and the HIV prevalence at the time of peak incidence to be from 10–50% and 5–95% of the HIV prevalence at equilibrium, respectively (see Table 7(a)). This is done to avoid identifiability issues that could arise from having epidemiologically implausible calibration conditions such as a very high peak HIV incidence followed by very low prevalence at equilibrium.

The calibration system given by (36)–(42) was then solved using Newton’s Method with the initial conditions given in (47)–(49). For each sample, the calibrated values of λ , α and n were noted along with the resulting HIV prevalence (\hat{A}^*), peak HIV incidence (\hat{I}^*), HIV prevalence at the time of peak incidence (\hat{P}^*) and the necessary number of iterations of Newton’s Method (N). The results are shown in Table 6(b).

For all 50 samples, the calibration approach successfully matched the three calibration conditions while requiring no more than 3 iterations of Newton’s Method (see Table 6(b)). To illustrate the variation in the nature of the epidemic, Figure 7 shows the dynamics of HIV prevalence and HIV incidence of the first 12 calibrations. In each calibration plot, the horizontal dashed blue line indicates the sampled HIV prevalence at equilibrium (\hat{A}), the horizontal dashed red line the sampled peak HIV incidence (\hat{I}) and the horizontal dashed gray line the sampled HIV prevalence at the time of peak incidence (\hat{P}). Thus, the plot indicates a successful calibration if HIV prevalence (solid blue curve) approaches \hat{A} (dashed blue line) as $t \rightarrow \infty$; HIV incidence (solid red curve) peaks at \hat{I} (dashed red line); and HIV prevalence at time of peak incidence is \hat{P} (solid blue curve, horizontal dashed gray line and the dashed vertical gray line indicating time of peak incidence all intersect at the same point).

The minor discrepancies between the sampled and calibrated values of HIV prevalence at peak incidence (\hat{P}) observed in Table 6(b) are due to the transient nature of the system at the time of peak incidence. Consequently, the HIV prevalence at peak incidence is influenced by the imprecision of choosing the time of peak incidence from the discrete set of time points t_0, t_1, \dots, t_K that results from the numerical solution to the dynamical systems model. Table 6(b) also shows a troubling level of variability in the magnitude of calibrated values that determine the population’s response to disease prevalence, α^* and n^* . Most notably, we see in Table 6(b) that α^* ranges from values of order 10^1 to 10^{149} . In Figure 8, we plot the behavioral response to HIV prevalence via the transmission term $e^{-\alpha P^n}$ for the calibrations with largest α values. We note that the shape of the behavioral responses remain biologically feasible despite the magnitude of the α terms. Also in Figure 8, we plot the prevalence and incidence curves for Calibrations #20, 30, 33, 40 and 46 (noting that those of #4 and #10 are in Figure 7). In each calibration, we observe that the scenario of a late-peaking epidemic (i.e. one where the HIV prevalence at the time of peak incidence is very near the HIV prevalence at equilibrium) seems to be producing the large α^* values.

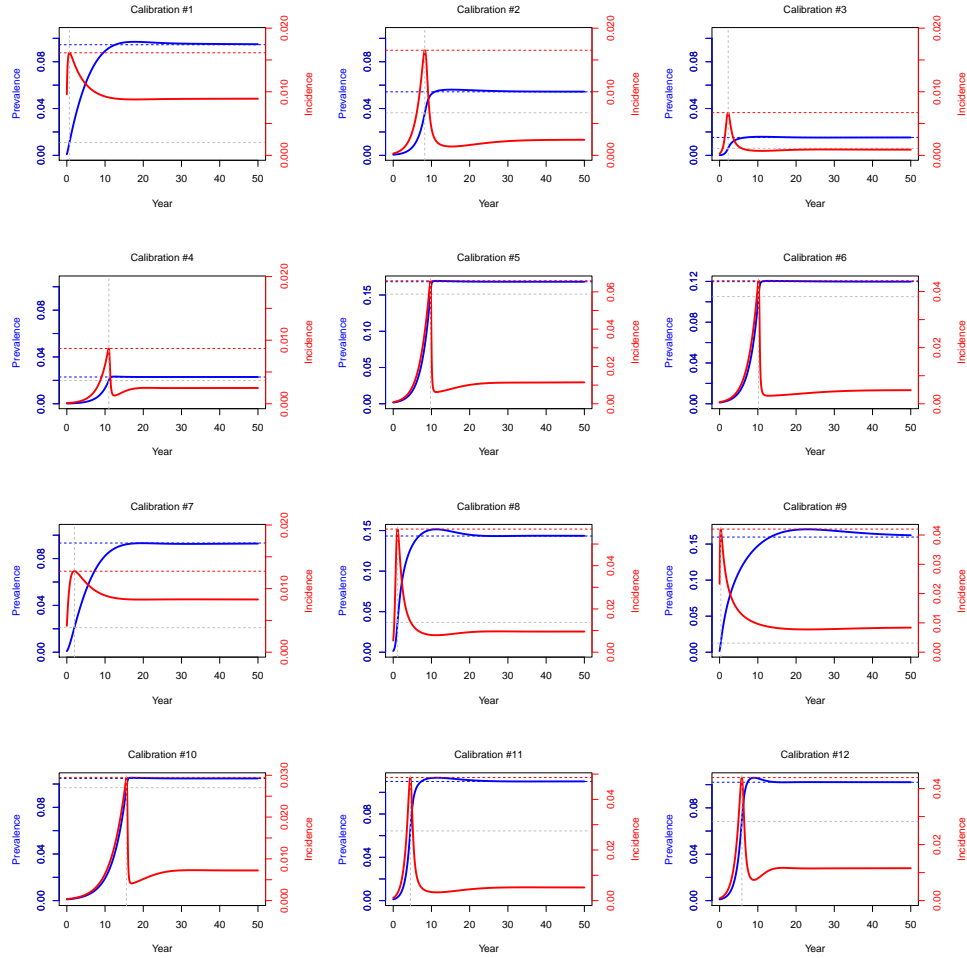


FIGURE 7. Dynamics of HIV prevalence and HIV incidence of the first 12 calibrations: horizontal dashed blue line is sampled HIV prevalence at equilibrium (\hat{A}), horizontal dashed red line is sampled peak HIV incidence (\hat{I}) and horizontal dashed gray line is sampled HIV prevalence at the time of peak incidence (\hat{P}). Calibration is successful if HIV prevalence (solid blue curve) approaches \hat{A} (dashed blue line) as $t \rightarrow \infty$; HIV incidence (solid red curve) peaks at \hat{I} (dashed red line); and HIV prevalence at time of peak incidence is \hat{P} (solid blue curve, horizontal dashed gray line and the dashed vertical gray line indicating time of peak incidence all intersect at the same point).

7.2. Approximate calibration of HIV/HSV-2 coinfection model to HIV prevalence at equilibrium, peak HIV incidence and timing of peak incidence. In Section 6.2, the HIV/HSV-2 model with behavioral response was calibrated to HIV and HSV-2 prevalence at the endemic equilibrium. In the previous

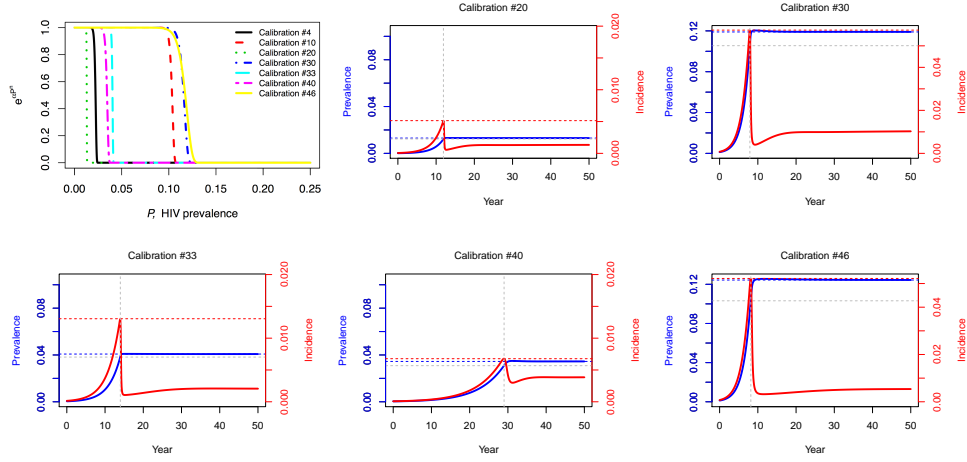


FIGURE 8. Behavioral response to HIV prevalence via the transmission term $e^{-\alpha P^n}$ for the calibrations with largest α values and the prevalence/incidence curves for those calibrations. Note that the prevalence/incidence curves of Calibrations #4 and #10 are shown in Figure 7.

section, the *Granich et al.* HIV model was calibrated to different types of epidemic curves (e.g. early, mid and late peaking epidemics) using peak HIV incidence and the timing of peak HIV incidence. In this section, we attempt to calibrate the expanded HIV/HSV-2 coinfection model with behavioral response (i.e. System (20) in Section 6.1) to HIV and HSV-2 prevalence conditions as well as the the shape of the HIV epidemic (i.e. peak HIV incidence and its timing). The conditions are given by

$$\begin{aligned}
 \hat{A} &= \text{HIV prevalence at endemic equilibrium,} \\
 \hat{H} &= \text{HSV-2 prevalence at endemic equilibrium,} \\
 \hat{I} &= \text{peak HIV incidence,} \\
 \hat{P} &= \text{HIV prevalence at time of peak HIV incidence,}
 \end{aligned} \tag{50}$$

and the calibration is achieved using the initial values of the HIV and HSV-2 transmission terms (i.e. λ and κ) and the location and shape of the transmission terms (i.e. α and n) as calibration parameters.

To impose the peak incidence condition in the same manner as the previous section, we note that the rate of new HIV infections in the HIV/HSV-2 model with behavioral response is given by $\lambda e^{-\alpha P^n} (a+r_2c)(s+r_1h)$. Unfortunately, we immediately see that this rate is no longer a function of only P as was the case in (40) in our previous calibration. This is a major issue in setting up a calibration system because peak HIV incidence occurs during the transient phase of the epidemic. Consequently the state variables a, c, s and h at this time are not the endemic equilibrium values that would be enforced in the rest of the calibration system. Overcoming this would require an incidence formulation such as $\lambda e^{-\alpha P^n} \hat{P}(1+r_2\omega_1)(1-\hat{P})(1+r_1\omega_2)$, where ω_1 (ω_2) are the proportion of HIV positive (negative) individuals that are infected

with HSV-2 at the time of peak HIV incidence. Unfortunately, such ω values would depend on the particular initial conditions assumed and would not be available *a priori*.

As a simpler alternative, we present the following approximate calibration of the model to the four conditions given in (50) using λ , κ , n and α as calibration variables. This two step approach relies on previous calibration systems in the following way:

Step 1. Calibrate the *Granich et al.* model to \hat{A} , \hat{I} and \hat{P} using calibration parameters λ , n and α by solving system (36)–(42) in the previous section.

Step 2. Using the resulting values for HIV transmission terms n and α from Step 1, calibrate the HIV/HSV-2 model with behavioral response to \hat{A} and \hat{H} using λ and κ as calibration variables by solving system (22)–(32) from Section 6.2.

The calibrated parameter values are then n^* , α^* (outputs from Step 1 which determine the shape of the behavioral response) and λ^* , κ^* (outputs from Step 2 which impose disease prevalence at endemic equilibrium).

7.2.1. *Demonstration of calibration to HIV and HSV-2 prevalence at equilibrium, peak HIV incidence and timing of peak HIV incidence.* To demonstrate the performance of this approach, we calibrate the model to 50 independent samples of parameters and calibration conditions from uniform distributions over the ranges specified in Table 7(a). The results are shown in Table 7(b). As might be expected given Step 2 of this approximate approach, we see that the model is exactly calibrated to both HIV and HSV-2 prevalence at equilibrium (see \hat{A} vs \hat{A}^* and \hat{H} vs \hat{H}^* in Table 7(b)). On the other hand, the calibration is not able to exactly match the conditions related to the transient phase of the epidemic (i.e. peak HIV incidence and timing of peak HIV incidence). Specifically, we see in Table 7(b) that the peak HIV incidence realized via calibration, \hat{I}^* , is higher than the desired value, \hat{I} , and quite often by a substantial amount. This makes sense because the terms that control the behavioral response to HIV prevalence, namely n and α , are determined in Step 1 before HIV's interaction with HSV-2 infection is included. Including the interaction produces a higher peak incidence level due to the fact that HSV-2 infection increases the rate of new HIV infections from both sides of transmission (i.e. $r_1, r_2 \geq 1$). While the approximate calibration considerably overshoots peak HIV incidence, it is worth noting that the approach does much better in terms of the timing of peak HIV incidence (i.e. \hat{P}^* vs \hat{P} in Table 7(b)).

8. Discussion. The model calibrations conducted in this work have demonstrated that an exact approach to calibration can be a straightforward and efficient means to parameterize an epidemiological model to surveillance data. On the other hand, the similarities of the scenarios examined herein (e.g. disease prevalence at endemic equilibrium) may cast doubt on the general applicability of this approach in other situations. Hence, it is important to identify limitations and generalizations of the approach that could influence its applicability.

From the examples presented, the exact calibration approach may appear to be limited to diseases at endemic equilibria. However, it is worth noting that there is no such limitation from a technical perspective. As an example, recall our first calibration system (4)–(8) from Section 4.2. If reliable data were available that suggested a given HIV and HSV-2 prevalence of \hat{A} and \hat{H} , respectively, and that the proportions of HIV only, HSV-2 only and coinfecting individuals were changing at rates $\frac{da}{dt}$, $\frac{dh}{dt}$ and $\frac{dc}{dt}$, respectively, one could simply insert these nonzero rates

as the left-hand sides of (4)–(6). The calibration procedure would work exactly the same. However, a practical limitation is evident if we attempt to consider an analogous situation for the HIV/HSV-2 coinfection model with behavior change from Section 6. While it is reasonable that estimates could exist for the rates of change for the aggregate HIV-infected, HSV-2 infected and coinfecting proportions of the population, it is unlikely that specific estimates exist for each subdivision (i.e. $\frac{da_1}{dt}, \dots, \frac{da_4}{dt}, \frac{dc_1}{dt}, \dots, \frac{dc_4}{dt}$).

An exact calibration to estimates of disease prevalence, incidence, etc. may leave the impression that such indicators are known data rather than the imprecise estimates that they are in reality. However, this approach should not be viewed as an alternative to incorporating uncertainty into modeling work. In fact, the ideas presented here could compliment the standard approach of sampling parameter values and filtering described in Section 2. In the standard approach, all model parameters would be independently sampled before running simulations and filtering out the parameter sets which do not meet calibration conditions. In a combined approach, model parameters for which strong empirical estimates exist could be randomly sampled and values of calibration parameters (for which reliable empirical estimates do not exist) would be obtained via the exact procedure to match calibration conditions. Parameter sets with infeasible values for the calibration parameters could then be filtered out. With such an approach, the ultimate number of parameter sets that are filtered out (i.e. computational time wasted) would be significantly reduced. Moreover, uncertainty in the calibration conditions (e.g. disease prevalence, incidence) could also be incorporated by sampling values from assumed distributions as was done in each robustness demonstration in this work.

Lastly, it is clear that the exact calibration approach can not be used to fit a model to time series data. While approaches were shown to calibrate to different types of epidemics in Section 7, the approach remains limited to calibration conditions at a few time steps, at most. Hence, numerical fitting algorithms will be required whenever a model must be fit to time series data. In such situations, our approach could improve the performance of numerical fitting algorithms whose efficiency and precision are largely determined by the quality of the initial approximation to the solution. To do this, an exact fitting (analogous to that of Section 7.1) would be used to obtain parameter values for which the model behavior matches the general shape of the data. Those values would then be used as an initial approximation for a numerical fitting algorithm.

While not universally applicable or intended to replace other calibration procedures, the exact calibration process presented in this work streamlines the process of calibrating mathematical models of infectious disease to surveillance data while increasing efficiency and reducing reliance on computational methods.

REFERENCES

- [1] L. J. Abu-Raddad, A. S. Magaret, C. Celum, A. Wald, I. M. Longini, S. G. Self and L. Corey, Genital herpes has played a more important role than any other sexually transmitted infection in driving hiv prevalence in africa, *PLoS ONE*, **3** (2008), e2230.
- [2] B. Auvert, R. Ballard, C. Campbell, M. Caraël, M. Carton, G. Fehler, E. Gouws, C. MacPhail, D. Taljaard, J. V. Dam and B. Williams, Hiv infection among youth in a south african mining town is associated with herpes simplex virus-2 seropositivity and sexual behaviour, *AIDS*, **15** (2001), 885–898.
- [3] J. M. Baeten, D. Donnell, P. Ndase, N. R. Mugo, J. D. Campbell, J. Wangisi, J. W. Tappero, E. A. Bukusi, C. R. Cohen, E. Katabira, A. Ronald, E. Tumwesigye, E. Were, K. H. Fife, J. Kiarie, C. Farquhar, G. John-Stewart, A. Kania, J. Odoyo, A. Mucunguzi, E. Nakku-Joloba, R. Twesigye, K. Ngure, C. Apaka, H. Tamoooh, F. Gabona, A. Mujugira, D. Panteleeff, K. K. Thomas, L. Kidoguchi, M. Krows, J. Revall, S. Morrison, H. Haugen, M. Emmanuel-Ogier, L. Ondrejcek, R. W. Coombs, L. Frenkel, C. Hendrix, N. N. Bumpus, D. Bangsberg, J. E. Haberer, W. S. Stevens, J. R. Lingappa and C. Celum, Antiretroviral prophylaxis for HIV prevention in heterosexual men and women, *N Engl J Med*, **367** (2012), 399–410.
- [4] R. V. Barnabas, E. L. Webb, H. A. Weiss and J. N. Wasserheit, The role of coinfections in HIV epidemic trajectory and positive prevention, *AIDS*, **25** (2011), 1559–1573.
- [5] M.-C. Boily, R. F. Baggaley, L. Wang, B. Masse, R. G. White, R. J. Hayes and M. Alary, Heterosexual risk of hiv-1 infection per sexual act: Systematic review and meta-analysis of observational studies, *Lancet Infect Dis*, **9** (2009), 118–129, URL [http://linkinghub.elsevier.com/retrieve/pii/S1473-3099\(09\)70021-0](http://linkinghub.elsevier.com/retrieve/pii/S1473-3099(09)70021-0).
- [6] R. Burden and J. Faires, *Numerical Analysis*, Cengage Learning, 2010.
- [7] M. S. Cohen, Y. Q. Chen, M. McCauley, T. Gamble, M. C. Hosseinipour, N. Kumarasamy, J. G. Hakim, J. Kumwenda, B. Grinsztejn, J. H. S. Pilotto, S. V. Godbole, S. Mehendale, S. Chariyalertsak, B. R. Santos, K. H. Mayer, I. F. Hoffman, S. H. Eshleman, E. Piwowar-Manning, L. Wang, J. Makhema, L. A. Mills, G. de Bruyn, I. Sanne, J. Eron, J. Gallant, D. Havlir, S. Swindells, H. Ribaud, V. Elharrar, D. Burns, T. E. Taha, K. Nielsen-Saines, D. Celentano, M. Essex, T. R. Fleming and HPTN 052 Study Team, Prevention of HIV-1 infection with early antiretroviral therapy., *N Engl J Med*, **365** (2011), 493–505.
- [8] M. Das, P. L. Chu, G.-M. Santos, S. Scheer, E. Vittinghoff, W. McFarland and G. N. Colfax, Decreases in community viral load are accompanied by reductions in new HIV infections in san francisco, *PLoS ONE*, **5** (2010), e11068.
- [9] E. E. Freeman, H. A. Weiss, J. R. Glynn, P. L. Cross, J. A. Whitworth and R. J. Hayes, Herpes simplex virus 2 infection increases hiv acquisition in men and women: Systematic review and meta-analysis of longitudinal studies, *AIDS*, **20** (2006), 73–83.
- [10] D. Gerberry, B. G. Wagner, J. G. García-Lerma, W. Heneine and S. Blower, Using geospatial modelling to optimize the rollout of antiretroviral-based pre-exposure HIV interventions in Sub-Saharan Africa, *Nature Communications*, **5** (2014), 1–15.
- [11] R. M. Granich, C. F. Gilks, C. Dye, K. M. De Cock and B. G. Williams, Universal voluntary HIV testing with immediate antiretroviral therapy as a strategy for elimination of HIV transmission: A mathematical model, *Lancet*, **373** (2009), 48–57.
- [12] R. H. Gray, M. J. Wawer, R. Brookmeyer, N. K. Sewankambo, D. Serwadda, F. Wabwire-Mangen, T. Lutalo, X. Li, T. vanCott, T. C. Quinn and R. P. Team, Probability of hiv-1 transmission per coital act in monogamous, heterosexual, hiv-1-discordant couples in rakai, uganda, *Lancet*, **357** (2001), 1149–1153.
- [13] R. J. Hayes, K. F. Schulz and F. A. Plummer, The cofactor effect of genital ulcers on the per-exposure risk of hiv transmission in sub-saharan africa, *J Trop Med Hyg*, **98** (1995), 1–8.
- [14] K. A. Powers, C. Poole, A. E. Pettifor and M. S. Cohen, [Rethinking the heterosexual infectivity of hiv-1: A systematic review and meta-analysis](#), *Lancet Infect Dis*, **8** (2008), 553–563.
- [15] J. A. Røttingen, D. W. Cameron and G. P. Garnett, A systematic review of the epidemiologic interactions between classic sexually transmitted diseases and hiv: How much really is known?, *Sex Transm Dis*, **28** (2001), 579–597.
- [16] J. S. Smith and N. J. Robinson, Age-specific prevalence of infection with herpes simplex virus types 2 and 1: A global review, *J Infect Dis*, **186** (2002), 3–28.
- [17] R. J. Smith and S. M. Blower, [Could disease-modifying HIV vaccines cause population-level perversity?](#), *Lancet Infect Dis*, **4** (2004), 636–639.

- [18] The World Bank, World development indicators, <http://data.worldbank.org/data-catalog/world-development-indicators> [Accessed: 25 Oct 2011], URL <http://data.worldbank.org/data-catalog/world-development-indicators>.
- [19] M. C. Thigpen, P. M. Kebaabetswe, L. A. Paxton, D. K. Smith, C. E. Rose, T. M. Segolodi, F. L. Henderson, S. R. Pathak, F. A. Soud, K. L. Chillag, R. Mutanhourwa, L. I. Chirwa, M. Kasonde, D. Abebe, E. Buliva, R. J. Gvetadze, S. Johnson, T. Sukalac, V. T. Thomas, C. Hart, J. A. Johnson, C. K. Malotte, C. W. Hendrix and J. T. Brooks, [Antiretroviral preexposure prophylaxis for heterosexual HIV transmission in botswana](#), *N Engl J Med*, **367** (2012), 423–434.
- [20] A. J. Vyse, N. J. Gay, M. J. Slomka, R. Gopal, T. Gibbs, P. Morgan-Capner and D. W. Brown, [The burden of infection with HSV-1 and HSV-2 in England and Wales: Implications for the changing epidemiology of genital herpes.](#), *Sex Transm Infect*, **76** (2000), 183–187.
- [21] A. Wald and K. Link, [Risk of human immunodeficiency virus infection in herpes simplex virus type 2-seropositive persons: a meta-analysis](#), *J Infect Dis*, **185** (2002), 45–52.
- [22] H. Weiss, [Epidemiology of herpes simplex virus type 2 infection in the developing world](#), *Herpes*, **11** (2004), 24A–35A.

Received September 15, 2016; Accepted March 2, 2017.

E-mail address: david.gerberry@xavier.edu

TABLE 6. Calibration of *Granich et al.* model to HIV prevalence, peak HIV incidence and timing of peak HIV incidence.

(a) Model parameters and calibration conditions for parameterizing the *Granich et al.* model to HIV prevalence, peak HIV incidence and timing of peak HIV incidence. ‡Recruitment rate range is same as background mortality to give a total population of 1 in the absence of HIV, without loss of generality.

	Symbol	Range	References
Model Parameters			
Background mortality rate (/yr)	μ	$\frac{1}{60} - \frac{1}{30}$	[11]
Recruitment rate into the population (/yr)	Λ	$\frac{1}{60} - \frac{1}{30}$	‡
Rate of HIV progression (/yr)	ρ	0.1 – 0.4	[11]
Calibration Conditions			
HIV prevalence at endemic equilibrium	\hat{A}	0.3 – 17	[18]
Peak HIV incidence	\hat{I}	10% – 50% of \hat{A}	
HIV prevalence at time of peak HIV incidence	\hat{P}	5% – 95% of \hat{A}	
Calibration Parameters			
Initial value of transmission term (/yr)	λ	determined by calibration procedure	
Location of transmission term	α	determined by calibration procedure	
Shape of transmission term	n	determined by calibration procedure	

(b) Results of calibrating the *Granich et al.* model to 50 independent sets of parameters and calibration conditions (randomly sampled from ranges above). For each sample, the resulting values of the calibration parameters (λ^*, α^*, n^*) and the resulting calibration values ($\hat{A}^*, \hat{I}^*, \hat{P}^*$) are given. The number of iterations of Newton's Method required to achieve a stopping criterion of $\max\{|\lambda_i - \lambda_{i-1}|, |\alpha_i - \alpha_{i-1}|, |n_i - n_{i-1}|\} < 10^{-9}$ is denoted by m .

j	Sampled parameters and calibration condition						Calibration results						
	Λ	μ	ρ	\hat{A}	\hat{I}	\hat{P}	λ^*	α^*	n^*	\hat{A}^*	\hat{I}^*	\hat{P}^*	m
1	0.031	0.028	0.306	0.095	0.016	0.011	224.91	1.22212E+01	0.2	0.095	0.016	0.011	2
2	0.017	0.027	0.115	0.054	0.017	0.036	0.54	2.80696E+09	7.2	0.054	0.017	0.036	2
3	0.031	0.022	0.171	0.015	0.007	0.006	1.82	2.22785E+04	2.1	0.015	0.007	0.006	2
4	0.024	0.026	0.372	0.023	0.009	0.020	0.46	1.09674E+42	25.5	0.023	0.009	0.020	3
5	0.032	0.032	0.192	0.168	0.066	0.151	0.52	1.91294E+33	42.7	0.168	0.066	0.152	3
6	0.022	0.023	0.102	0.120	0.044	0.105	0.48	1.30078E+32	34.4	0.120	0.044	0.104	3
7	0.031	0.018	0.307	0.093	0.013	0.021	15.58	1.04219E+01	0.3	0.093	0.013	0.021	2
8	0.024	0.017	0.224	0.143	0.057	0.037	3.86	3.28571E+01	1.1	0.143	0.057	0.039	2
9	0.027	0.033	0.124	0.160	0.042	0.013	62.63	1.27961E+01	0.3	0.160	0.042	0.014	2
10	0.032	0.023	0.215	0.105	0.029	0.097	0.34	1.15023E+57	58.1	0.105	0.029	0.096	3
11	0.023	0.020	0.138	0.110	0.049	0.064	0.97	2.69922E+05	5.2	0.110	0.049	0.066	2
12	0.032	0.024	0.393	0.102	0.044	0.068	0.80	3.07050E+06	6.3	0.102	0.044	0.070	2
13	0.031	0.018	0.199	0.074	0.013	0.055	0.27	5.60380E+09	8.5	0.074	0.013	0.055	2
14	0.026	0.023	0.386	0.049	0.019	0.018	2.04	2.41472E+02	1.5	0.049	0.019	0.018	2
15	0.033	0.032	0.364	0.138	0.067	0.073	1.32	1.62810E+03	3.3	0.138	0.067	0.071	2
16	0.018	0.032	0.378	0.055	0.010	0.024	1.11	4.61171E+01	1.0	0.055	0.010	0.024	2
17	0.027	0.019	0.254	0.170	0.064	0.075	1.36	1.83155E+02	2.4	0.170	0.064	0.075	2
18	0.018	0.025	0.223	0.062	0.028	0.051	0.62	5.34456E+22	18.6	0.062	0.028	0.050	3
19	0.025	0.017	0.177	0.147	0.058	0.051	2.07	9.90795E+01	1.7	0.147	0.058	0.049	2
20	0.017	0.032	0.357	0.013	0.005	0.012	0.43	2.57848E+149	79.2	0.013	0.005	0.012	3
21	0.031	0.028	0.367	0.083	0.016	0.050	0.48	1.12025E+03	2.7	0.083	0.016	0.050	2
22	0.017	0.023	0.213	0.157	0.035	0.030	6.57	1.27404E+01	0.6	0.157	0.035	0.028	2
23	0.023	0.022	0.278	0.124	0.051	0.010	104.17	1.40918E+01	0.3	0.124	0.051	0.009	2
24	0.020	0.018	0.280	0.044	0.013	0.025	0.69	2.06649E+05	3.7	0.044	0.013	0.025	2
25	0.027	0.023	0.234	0.169	0.018	0.084	0.54	1.16327E+01	1.0	0.169	0.018	0.084	2
26	0.025	0.017	0.112	0.041	0.008	0.018	0.69	5.50369E+03	2.4	0.041	0.008	0.018	2
27	0.021	0.030	0.135	0.146	0.067	0.094	0.90	9.15850E+05	6.6	0.146	0.067	0.093	2
28	0.026	0.022	0.257	0.076	0.014	0.021	2.34	2.52938E+01	0.8	0.076	0.014	0.021	2
29	0.021	0.025	0.245	0.105	0.038	0.092	0.47	2.27510E+28	28.8	0.105	0.038	0.091	3
30	0.021	0.030	0.277	0.119	0.057	0.106	0.62	2.40327E+33	35.8	0.119	0.057	0.107	3
31	0.024	0.017	0.382	0.092	0.017	0.013	137.54	1.16973E+01	0.2	0.092	0.017	0.013	2
32	0.017	0.033	0.111	0.097	0.020	0.034	1.18	9.80157E+01	1.5	0.097	0.020	0.034	2
33	0.023	0.019	0.150	0.041	0.013	0.038	0.36	3.52296E+107	77.2	0.041	0.013	0.039	3
34	0.026	0.019	0.115	0.065	0.020	0.058	0.38	9.88557E+45	38.4	0.065	0.020	0.058	3
35	0.018	0.029	0.133	0.123	0.039	0.046	1.45	1.60470E+02	1.9	0.123	0.039	0.047	2
36	0.018	0.023	0.212	0.128	0.048	0.049	1.72	1.44867E+02	1.9	0.128	0.048	0.050	2
37	0.029	0.022	0.194	0.131	0.036	0.021	9.28	1.59178E+01	0.6	0.131	0.036	0.021	2
38	0.023	0.029	0.113	0.147	0.030	0.120	0.30	2.46362E+14	17.0	0.147	0.030	0.120	2
39	0.030	0.030	0.132	0.093	0.041	0.071	0.66	5.17855E+13	12.9	0.093	0.041	0.071	2
40	0.028	0.032	0.370	0.034	0.007	0.031	0.23	6.39613E+38	26.6	0.034	0.007	0.031	2
41	0.026	0.028	0.264	0.012	0.002	0.010	0.26	7.91021E+25	13.5	0.012	0.002	0.010	2
42	0.022	0.021	0.139	0.038	0.018	0.003	60.79	2.99343E+01	0.4	0.038	0.018	0.003	2
43	0.029	0.025	0.175	0.131	0.044	0.017	16.96	1.61755E+01	0.5	0.131	0.044	0.015	2
44	0.031	0.019	0.163	0.061	0.013	0.042	0.38	6.60195E+08	7.0	0.061	0.013	0.041	2
45	0.021	0.017	0.238	0.141	0.053	0.106	0.61	7.72114E+09	11.3	0.141	0.053	0.104	2
46	0.023	0.027	0.104	0.124	0.052	0.103	0.59	1.88292E+20	22.0	0.124	0.052	0.105	3
47	0.024	0.029	0.295	0.076	0.025	0.052	0.57	9.65496E+07	6.9	0.076	0.024	0.053	2
48	0.027	0.029	0.378	0.106	0.012	0.091	0.21	2.25508E+02	2.7	0.106	0.012	0.091	2
49	0.030	0.033	0.202	0.024	0.003	0.016	0.24	6.01293E+05	3.5	0.024	0.003	0.016	2
50	0.029	0.020	0.116	0.028	0.008	0.007	2.62	2.59513E+02	1.2	0.028	0.008	0.007	2

