

Structural sensitivity in HIV modeling: A case study of vaccination

Cora L. Bernard*, Margaret L. Brandeau

Department of Management Science and Engineering, Stanford University, Stanford, CA, USA



ARTICLE INFO

Article history:

Received 3 July 2017

Accepted 23 August 2017

Available online 11 November 2017

Keywords:

HIV transmission

HIV vaccine

Structural sensitivity analysis

Inference robustness assessment

Network model

Dynamic compartmental model

ABSTRACT

Structural assumptions in infectious disease models, such as the choice of network or compartmental model type or the inclusion of different types of heterogeneity across individuals, might affect model predictions as much as or more than the choice of input parameters. We explore the potential implications of structural assumptions on HIV model predictions and policy conclusions. We illustrate the value of inference robustness assessment through a case study of the effects of a hypothetical HIV vaccine in multiple population subgroups over eight related transmission models, which we sequentially modify to vary over two dimensions: parameter complexity (e.g., the inclusion of age and HCV comorbidity) and contact/simulation complexity (e.g., aggregated compartmental vs. individual/disaggregated compartmental vs. network models). We find that estimates of HIV incidence reductions from network models and individual compartmental models vary, but those differences are overwhelmed by the differences in HIV incidence between such models and the aggregated compartmental models (which aggregate groups of individuals into compartments). Complexities such as age structure appear to buffer the effects of aggregation and increase the threshold of net vaccine effectiveness at which aggregated models begin to overestimate reductions. The differences introduced by parameter complexity in estimated incidence reduction also translate into substantial differences in cost-effectiveness estimates. Parameter complexity does not appear to play a consistent role in differentiating the projections of network models.

© 2017 The Authors. Production and hosting by Elsevier B.V. on behalf of KeAi Communications Co., Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

1. Introduction

A guiding principle in public health modeling stipulates that a model should be only as complex as is necessary. Some argue that, “Unnecessary complexity ... is almost as undesirable as over-simplification” (Basu & Andrews, 2013; Grassly & Fraser, 2008), yet there is little consensus as to what qualifies as “necessary.” A model should be simple but “not so simple that realistic violation of simplifying assumptions will change an inference” (Koopman, 2005). As a field, we have not yet

Abbreviations: PWID, people who inject drugs; MSM, men who have sex with men; PWUD, people who abuse but do not inject drugs; HIV, human immunodeficiency virus; HCV, hepatitis C virus; NSP, needle/syringe exchange program; SAT, substance abuse treatment.

* Corresponding author. Department of Management Science and Engineering, Huang Engineering Center, 475 Via Ortega, Stanford University, Stanford, CA 94305-4026, USA.

E-mail address: cibernard@stanford.edu (C.L. Bernard).

Peer review under responsibility of KeAi Communications Co., Ltd.

<https://doi.org/10.1016/j.idm.2017.08.002>

2468-0427/© 2017 The Authors. Production and hosting by Elsevier B.V. on behalf of KeAi Communications Co., Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

adopted a consistent evaluation of model complexity nor a protocol for structural sensitivity analysis; consequently, although taxonomies can provide guidance (Brennan, Chick, & Davies, 2006), structural choices for models are often chosen based on guesswork, intuition, and computational convenience. While most single-model studies can thoroughly assess first-order uncertainty (i.e., model sensitivity to input values), they cannot address second-order uncertainty (i.e., how structural assumptions and the construction of the model itself influence model behavior).

Although there is no consensus about *how* to address structural sensitivity, there is a growing consensus that it is an essential component of sensitivity analysis that *should* be addressed (Basu & Andrews, 2013; Bilcke, Beutels, Brisson, & Jit, 2011; Bojke, Claxton, Sculpher, & Palmer, 2009; Caro, Briggs, Siebert, & Kuntz, 2012; Jackson, Bojke, Thompson, Claxton, & Sharples, 2011; Suen, Goldhaber-Fiebert, & Brandeau, 2017). Consortia of modeling groups like CISNET (the Cancer Intervention and Surveillance Modeling Network) and the Mt. Hood Diabetes Challenge Network have made it common practice to collect and compare predictions across models that vary in scope, parameters, and structure (e.g., Eaton et al., 2012; van de Vijver et al., 2013). Bilcke et al. (2011) propose a standardized framework for addressing and presenting structural uncertainty in decision-analytic models, and Jackson et al. (2011) advocate for a comprehensive model that includes all possible parameters, which can then be pruned back to explore different simplifying assumptions.

In infectious disease modeling, a number of researchers have studied the effect of contact structure assumptions on epidemic control (Bansal, Grenfell, & Meyers, 2007; Hamilton, Handcock, & Morris, 2008; Huppert & Katriel, 2013; Kong, Wang, Han, & Cao, 2016). For instance, Bansal et al. (2007) characterize the “epidemiological distance” between network models and homogeneous-mixing compartmental models and Hamilton et al. (2008) explore how the properties of network degree distributions affect their ability to replicate disease propagation. The majority of such analyses use stylized examples to illustrate how different structural make-up affects model prediction and fitting to epidemics but they do not translate these differences into policy implications.

In the policy context, Rahmandad and Sterman (2008) compare multiple agent-based models with varying network structures to a dynamic compartmental model, all calibrated to the same targets, and Suen et al. (2017) analyze intervention effects for compartmental epidemic models with different levels of risk stratification. Both studies find that modeling choices may lead to different policy choices. Other infectious disease modeling studies have called for diversifying contact structures to improve modeling accuracy (Chick, Adams, & Koopman, 2000; Hellard et al., 2014; Scott, Hellard, & McBryde, 2016) and improve understanding of models' predictive capacities (Lee et al., 2017) or for adding individual heterogeneity to reduce predictive bias (Monteiro et al., 2016). Some studies compare parameter complexity across models with otherwise similar compartmental structure (Sital, Little, Barnes, & White, 2016) or compare the outcomes of simple compartmental models to the predictions of more complex models (White et al., 2009), demonstrating that models can and should be as simple as possible.

These are mostly isolated analyses; researchers have not employed a consistent framework for categorizing the exploration of structural assumptions in infectious disease modeling. Many studies point out the overwhelming use of dynamic compartmental models when evaluating policies and emphasize the important role of network effects in disease spread (Dombrowski et al., 2017; Hellard et al., 2014; Koopman, 2004; Scott et al., 2016) but do not offer further insight into when the use of various structures might be more or less appropriate or even how one might establish such criteria.

Towards this goal, Koopman develops a structural taxonomy for transmission models (Koopman, 2004; Koopman, Jacquez, & Chick, 2001) and a systematized approach, inference robustness assessment, for isolating the effects of structural choices in a model (Koopman, 2005, 2007; Koopman, Singh, & Ionides, 2016). Inference robustness assessment tests the validity of a given structural assumption, and the inference it produces, by relaxing that assumption gradually over a family of linked models. The experiment is thus designed to explicitly assess the extent to which inferences are robust to simplifying assumptions.

In this paper, we present a case study focused on HIV modeling that utilizes some of the key concepts of inference robustness assessment. Our case study compares the effects of a hypothetical HIV vaccine in multiple subpopulations over eight related transmission models, all calibrated to the HIV epidemic in King County, Washington, which we sequentially modified to vary over two dimensions: parameter complexity (e.g., the inclusion of age and hepatitis C virus (HCV) comorbidity) and contact/simulation complexity (e.g., compartmental vs. network models).

The introduction of an HIV vaccine could radically shift HIV control policy, and infectious disease modelers can play an important role in informing policy makers about population-level effectiveness and cost-effectiveness. As of June 2017, multiple ongoing HIV vaccine trials were in the phase II stage (Choi et al., 2016; National Institutes of Health, 2017). No model-based study of vaccine effectiveness or cost-effectiveness analysis has addressed structural uncertainty (Adamson, Dimitrov, Devine, & Barnabas, 2017). Additionally, no study that we are aware of has applied the principles of inference robustness assessment in infectious disease modeling, nor has any study that we are aware of compared models along the dimensions of parameter and contact complexity simultaneously.

In Section 2 we describe our modeling approach. In Section 3 we present the results of our analyses. We conclude with discussion in Section 4.

2. Methods

2.1. Overview

Fig. 1 provides a schematic overview of the eight models in our case study, organized according to the two dimensions of contact/simulation complexity and parameter complexity. Section 2.2 discusses each model in detail and Table 1 summarizes

component differences between models. In brief, a dynamic compartmental model typically captures mixing at an aggregated level between groups of individuals (Brennan et al., 2006; Koopman, 2004; Koopman et al., 2001). Every compartment/group mixes with every other compartment/group with some time-specific probability. An individual compartmental model disaggregates compartments to the individual level but still assumes that every compartment/individual mixes with every other compartment/individual. A network model captures individuals mixing selectively with some subset of other individuals according to network structure. These models differ both in terms of contact complexity (how individuals or groups of individuals mix with each other) and simulation complexity (whether there are multiple, stochastic trials, resulting from selecting individual transition probabilities from distributions, or fixed, point-estimate parameters pertaining to each compartment).

The Appendix contains all input data and technical details of the models, consistent with international model reporting guidelines (Rahmandad & Sterman, 2012). All models were populated with data from King County, Washington and calibrated to key epidemiologic targets, such as overdose deaths, HIV incidence, awareness, and treatment in people who inject drugs (PWID), men who have sex with men (MSM), the PWID/MSM subpopulation, and heterosexual groups, considered lower risk in terms of their sexual and injecting behaviors (Aleccia, 2016; Buskin, Hanrahan, & Jaenicke, 2016). Each model included these risk groups and had consistent “macro-network” structure (Koopman, 2007); that is, the distributions over risk group and sex/sexual orientation by which individuals choose their partners were the same across all models (Table 2). An illustrative subset of demographic parameters is presented in Table 3. Tables A.1–A.16 document all parameters and sources.

We consider two broad types of models along the dimension of contact/simulation complexity: compartmental models (Models 1a – 4) and network models (Models 5a – 5c). Other, subtler differences exist as we move along each dimension of complexity. The compartmental models vary according to the level at which transmission rates are calculated and applied: Models 1a, 1b, and 2 assume that contact is aggregated at the subpopulation (risk group) level; Models 3 and 4 assume that contact is disaggregated at the individual level. The compartmental models also vary according to heterogeneity within subpopulations: Models 1a, 1b, 2, and 3 assume that individuals within a subpopulation are homogeneous, while Model 4 allows individuals to be distinguished by their event histories.

We consider three levels along the dimension of parameter complexity: models in which we distinguish three population subgroups (PWID, MSM, and low-risk) and we include the single disease HIV (Models 1a and 5a); models in which individuals are additionally distinguished by age and HCV infection status (Models 1b, 2, 3, 4 and 5b); and a model that includes persons who abuse drugs but do not inject (PWUD), distinguishes individuals by race, and tracks incarceration dynamics (Model 5c).

In practice, we began by implementing the most complex model, Model 5c, then iteratively constructed the remaining models so that each successively less complex model either completely removed structural components (e.g., incarceration)

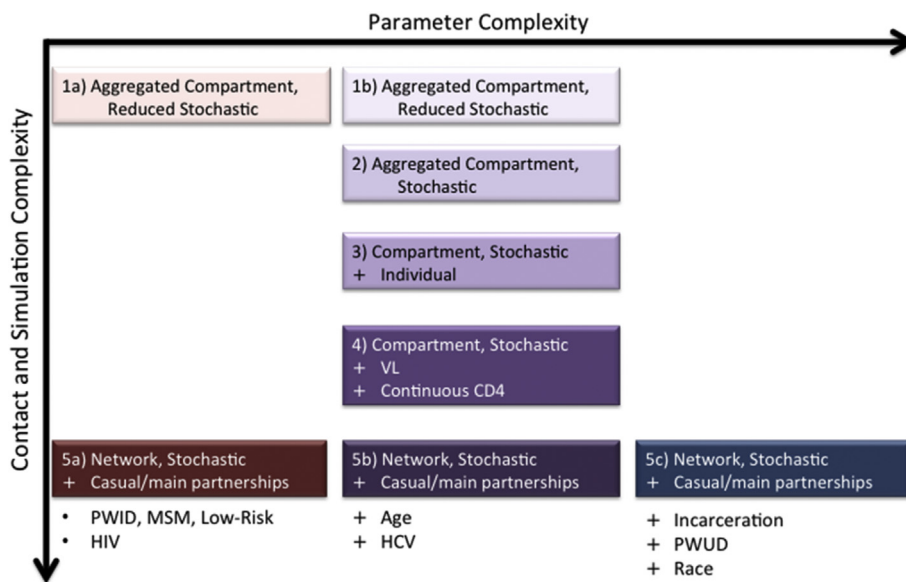


Fig. 1. Schematic of model types. PWID = people who inject drugs. MSM = men who have sex with men. HCV = hepatitis C virus. HIV = human immunodeficiency virus. PWUD = people who abuse drugs but do not inject. The eight models in our case study are organized along two dimensions: contact/simulation complexity and parameter complexity. Moving from left to right, along the parameter complexity direction, we begin with three basic risk groups (PWID, MSM, and low risk) and HIV; add age structure and HCV comorbidity; and finally add incarceration structure, the PWUD risk group, and race to the models. Moving from top to bottom, along the contact/simulation complexity dimension, we begin with a close approximation of a dynamic compartmental model; introduce heightened stochasticity; move compartment size from a subpopulation to a (homogeneous) individual; introduce heterogeneity of individuals; and finally include network structure.

Table 1

Model complexity: Inclusion of different aspects of parameter, structural, and simulation complexity in the considered models.

	Model							
	1a	1b	2	3	4	5b	5a	5c
Parameter Complexity								
CD4 count	×	×	×	×	×	×	×	×
PWID, MSM, and low-risk groups	×	×	×	×	×	×	×	×
HCV comorbidity		×	×	×	×	×	×	×
Age structure		×	×	×	×	×		×
Incarceration						×		
PWUD risk group						×		
Race						×		
Contact Complexity								
Compartment	×	×	×	×	×			
Aggregated	×	×	×					
Individual/disaggregated				×	×	×	×	×
Continuous viral load					×	×	×	×
Continuous CD4 count					×	×	×	×
Network						×	×	×
Simulation Complexity								
Deterministic/reduced stochastic	×	×						
Stochastic			×	×	×	×	×	×

PWID = people who inject drugs. MSM = men who have sex with men. PWUD = people who abuse but do not inject drugs. A × indicates that the indicated form of complexity is included in the model.

Table 2

Models' macro-network structure: Partner choice distribution.

		PWID			Non-PWID		
		HF	HM	MSM	HF	HM	MSM
PWID	HF	0.0	0.80	0.0	0.0	0.20	0.0
	HM	0.80	0.0	0.0	0.20	0.0	0.0
	MSM	0.12	0.0	0.68	0.03	0.0	0.17
Non-PWID	HF	0.0	0.0	0.0	0.0	1.0	0.0
	HM	0.0	0.0	0.0	1.0	0.0	0.0
	MSM	0.0	0.0	0.0	0.15	0.0	0.85

HF = heterosexual female. HM = heterosexual male. MSM = men who have sex with men. To identify how a member of a risk group chooses a partner, read a row from left to right: for example, a female PWID has an 80% chance of choosing a heterosexual male PWID and a 20% chance of choosing a low-risk heterosexual male. All models except Models 1a and 5a include an additional selection distribution by age. Data sources are shown in [Table A.4](#).

Table 3

Baseline HIV prevalence by risk group, sex/sexual orientation and age.

PWID (0.017)														
HF (0.326)					HM (0.579)					MSM (0.095)				
18–29	30–39	40–49	50–59	60–74	18–29	30–39	40–49	50–59	60–74	18–29	30–39	40–49	50–59	60–74
(0.151)	(0.249)	(0.267)	(0.250)	(0.083)	(0.153)	(0.248)	(0.263)	(0.252)	(0.084)	(0.126)	(0.268)	(0.323)	(0.212)	(0.071)
0.038	0.070	0.087	0.041	0.041	0.019	0.035	0.044	0.020	0.020	0.385	0.541	0.599	0.399	0.399
Non-PWID (0.983)														
HF (0.503)					HM (0.472)					MSM (0.025)				
18–29	30–39	40–49	50–59	60–74	18–29	30–39	40–49	50–59	60–74	18–29	30–39	40–49	50–59	60–74
(0.229)	(0.215)	(0.211)	(0.194)	(0.150)	(0.228)	(0.216)	(0.212)	(0.195)	(0.149)	(0.343)	(0.281)	(0.192)	(0.128)	(0.056)
0.0004	0.0008	0.0012	0.0014	0.0008	0.0006	0.0013	0.0021	0.0026	0.0015	0.0813	0.1667	0.1995	0.2212	0.2252

HF = heterosexual female. HM = heterosexual male. MSM = men who have sex with men. Parenthetical values correspond to the percent that the current row's subpopulation makes up of the next highest (sub)population: for example, PWID make up 1.7% of the model population and 32.6% of PWID are heterosexual female. Models 1a and 5a do not include the age structure. Within each subpopulation, those models take the expected value of HIV prevalence over its age distributions as the starting prevalence. Data sources can be found in [Appendix Tables A.1 and A.2](#).

or took the expected value across them (e.g. we took a weighted average of mortality over age groups for models that did not include age) to determine the appropriate parameters in a simplified context. For instance, in Model 5c, we assumed that no overdose deaths occurred in the lowest-risk population because we considered a separate PWUD population. In all other models, the low-risk population subsumed the PWUD population, which meant that net mortality increased across this low-risk group compared to that in Model 5c. In this way, each model was designed in direct correspondence to its neighbors on the parameter complexity/contact and simulation complexity plane and explicitly included the same categories, if not the same sets, of assumptions. The models were programmed in Python™ version 2.7.11. Model code can be obtained by writing to the authors.

In summary, Models 2, 3, and 4 are microsimulations that capture the intermediate steps between dynamic compartmental models, which Model 1b closely approximates, and network models such as Model 5b. Models 1a and Models 5a and 5c expand or contract Models 1 and 5, respectively, in terms of parameter complexity. Using each of the eight models, we examined the effect of a hypothetical HIV vaccine with varying levels of coverage and efficacy. We simulated each model over a five-year time horizon and measured the number of HIV infections averted in each population subgroup.

2.2. Description of models

Model 1a closely approximates the transmission and transition dynamics of a traditional compartmental model. The model's subpopulations are categorized by risk group (PWID, MSM, or lower-risk) and sex/sexual orientation (heterosexual male, heterosexual female, or MSM) with specific demographics supplied from King County, Washington data (Broz et al., 2014; Buskin et al., 2016; Kingston & Banta-Green, 2016; Monteiro et al., 2016; Seattle and King County Public Health Department, 2009; US Census Bureau, 2017; Vance-Sherman, 2015). Individuals enter the model as low-risk, HIV-uninfected 18-year-olds and mature out of it at age 75. While active in the model, individuals in uninfected compartments mix with individuals in infected compartments according to subpopulation characteristics. HIV transmission can occur through either sexual or injection-based contact. The compartment-specific rates of infection depend on the probability of contact with each infected compartment and the probability of infection from that compartment given that contact occurred. Community programs for PWID such as needle/syringe exchange programs (NSP) or substance abuse therapy (SAT), which broadly includes both drug rehabilitation and opioid agonist therapies, moderate the use and sharing of injecting equipment and therefore the risk of HIV infection (Bernard et al., 2016). Upon infection with HIV, individuals enter a brief but highly infectious acute phase and then progress through stages characterized by decreasing CD4 count, increased probability of transmitting to partners, and increased risk of opportunistic infections and mortality (Bendavid, Brandeau, Wood, & Owens, 2010; Bendavid et al., 2008; Monteiro et al., 2016; Public Health Agency of Canada, 2012; Zhong, Arjmand, Brandeau, & Bendavid, 2017). HIV-infected individuals can be tested and, once aware of their infection status, are eligible for antiretroviral therapy (ART), which reduces infectivity and extends life expectancy (Bernard et al., 2016). Background mortality rates adjust for deaths from overdose (Aleccia, 2016; Bird, Parmar, & Strang, 2015; Degenhardt et al., 2014; Larney, Toson, Burns, & Dolan, 2012; Lee et al., 2015; Maryland Department of Health and Mental Hygiene, 2014; National Institute on Drug Abuse, 2017; Rich et al., 2015), comorbidities (Liu, Cipriano, Holodniy, Owens, & Goldhaber-Fiebert, 2012), and differences between sexes (Arias, 2014). Aside from the characteristics detailed above, individuals are homogenous and cannot be distinguished from one another in the model. Details of transmission calculations and model equations can be found in the Appendix and in the appendix of two previously published analyses (Bernard, Owens, Goldhaber-Fiebert, & Brandeau, 2017; Bernard et al., 2016).

Model 1b introduces a greater level of parameter complexity to Model 1a: subpopulations are further divided by age (Buskin et al., 2016; US Census Bureau, 2017) and HCV comorbidity (Buskin, Hanrahan, & Jaenicke, 2013; Buskin et al., 2016; Hellard et al., 2014; University of Washington, 2017; Valdiserri & Koh, 2014) according to the demographics of King County, Washington. HCV infection is modeled as a Markov process. At the time of HCV infection, individuals enter a six-month acute phase, during which they are not infectious (Chen & Morgan, 2006). If they do not spontaneously clear the HCV infection (Liu, Watcha, Holodniy, & Goldhaber-Fiebert, 2014; Scott & Chew, 2017), it progresses to a chronic, and infectious, F0 stage, which is followed by the F1, F2, F3, and F4 stages before end-stage liver disease (ESLD) (Liu et al., 2014). Progression rates depend on age, sex, HIV comorbidity, and HIV treatment status. HCV treatment is available, although quitting may occur before completion, and even at the time of completion not everyone achieves a sustained virologic response (Cipriano et al., 2012; Liu et al., 2014). Mortality rates are disaggregated to separately account for background risk by age and HCV-specific risk, which is especially high in ESLD (Liu et al., 2012). HCV transmission is modeled concurrently with HIV infection using the same contact structure as Model 1a with additional macro-network complexity to capture age-based selection; that is, the distribution over which individuals choose a partner now depends on age as well as risk group and sex/sexual orientation.

Model 2 is identical to Model 1b in parameter complexity. Like Model 1b, it aggregates the calculation of disease transmission at the subpopulation level, but unlike Model 1b, after calculating the transmission probability it applies it at the individual level, rather than as a compartment-specific rate, thus creating enhanced stochasticity. We introduced Model 2 to isolate the effects of aggregation from the effects of deterministic modeling and rounding. That is, if the aggregated models overestimated vaccine effectiveness, as we predicted they would, we wanted to demonstrate with clarity that this came from structural assumptions and not from the potentially arbitrary effects of model implementation.

Model 3 is identical to Model 2 in parameter complexity, but unlike Model 2, in which compartments are composed of aggregated individuals, the compartment and the individual are synonymous in Model 3. As in any compartment model, all compartments (in this case, therefore, all individuals) can potentially mix with all other compartments (individuals), but here the calculations of transmission probabilities are necessarily disaggregated. While we explicitly track individuals in Model 3, they remain homogeneous up to the same subpopulation levels as in Model 2.

Model 4 captures the same mixing and transmission structures as Model 3 but relaxes the assumption that individuals/compartments are homogeneous. In Model 4, individuals have heterogeneous event history and unchangeable characteristics such as adherence to ART (Chen et al., 2013; Monteiro et al., 2016) and genotypes that make it easier or harder to clear HCV (Cipriano et al., 2012). A random integer is drawn from a negative binomial distribution (Hamilton et al., 2008; Monteiro et al., 2016) to determine the number of partners an individual has at each time step, which means that, in practice, many individuals do not mix because they have 0 partners that month. (In Models 1a, 1b, 2, and 3, the number of partners is taken as

the mean of the partnership distribution.) HIV is no longer modeled as a discrete Markov process. Instead, individuals have continuous viral load and CD4 count trackers. Starting viral loads are assigned from a distribution with mean 4.02 log₁₀ copies/ml (Bendavid et al., 2010), and CD4 counts begin in a 500–1600 cells/mm³ range. For an HIV-infected individual on ART, CD4 count rises and viral load decreases during an initial starting period (Zhong et al., 2017). In the absence of ART, CD4 count decreases with a magnitude that depends on viral load.

Model 5b is identical to Model 4 except for contact structure. Model 5b relaxes the assumption of homogeneous compartmental mixing, and connects individuals through partnerships that form sexual and injecting networks. Partnerships can be casual or main, which signify different risk behaviors (Broz et al., 2014; Kapadia et al., 2007; Monteiro et al., 2016; Sionean et al., 2014). (In Models 1a – 4, risk behavior is averaged across all partnerships.) Model 5a contracts the parameter complexity of Model 5b in the same way that Model 1a contracts Model 1b: parameters are averaged over age and HCV comorbidity and these characteristics are no longer explicitly tracked.

Model 5c extends the parameter complexity of Model 5b. A recent systematic review of HCV models for PWID noted the potential importance for disease transmission dynamics of high incarceration rates that interrupt injecting networks and called for modeling incarceration (Scott et al., 2016). No infectious disease models for PWID that we are aware of have incorporated incarceration dynamics. To capture the relevant aspects of the King County, Washington incarceration system, we introduced the PWID risk group, disaggregating it from the non-injecting, low-risk population. We further categorized individuals by race, and explicitly modeled criminal activity and jail and prison structures. In the model, crimes occur on a weekly basis. When a crime is identified, it is determined to be a misdemeanor or a felony and an appropriate sentence length is assigned (Berk Consulting, 2014; King County Department of Adult and Juvenile Detention, 2015). All crimes are processed through the jail structure, with felons awaiting trial and potentially serving a less than one-year sentence there. Otherwise, after conviction, felons are sent to prison. We assumed there is no HIV or HCV transmission during incarceration. At the time of release, due to lowered tolerance, PWID have an increased risk of death from overdose (Aleccia, 2016; Degenhardt et al., 2014; Larney et al., 2012; Maryland Department of Health and Mental Hygiene, 2014; Rich et al., 2015).

2.3. Model instantiation and calibration

We instantiated the models with data to reflect the population of King County, Washington (Table 3, Appendix Tables A.1–A.15). We manually calibrated Model 5c to multiple epidemiologic targets in the county such as jail and prison populations, HCV and HIV incidence, prevalence, awareness, and treatment, and overdose deaths, both in the population as a whole and in each population subgroup (Appendix Figures A.1–A.10). We then calibrated all other models to the output of Model 5c. Fig. 2 shows the distribution of cumulative five-year HIV incidence across subpopulations (PWID, MSM, MSM/PWID, and low-risk) for each model under the status quo (in the absence of a vaccine intervention).

We associated each state and some transitions with a cost and quality-of-life value and calculated lifetime costs and quality-adjusted years (QALYs) (Appendix Table A.16). We discounted costs and QALYs at 3% annually (Gold, 1996; Weinstein et al., 2003). Costs in each model included, where relevant, background healthcare costs by sex and age, costs of criminal activity and incarceration, community program costs, and disease-specific costs. QALYs included, where relevant, background quality of life by sex, age, and risk group, and multipliers greater or less than 1 which improved or reduced quality of life depending on whether they were associated with community programs or infection. At the end of each model run we included projected future lifetime costs and QALYs for each subpopulation by infection status to capture long-term costs and health outcomes (details in Appendix).

2.4. HIV vaccine scenarios

We considered a hypothetical protective HIV vaccine given to individuals at the start of the modeled 5-year time horizon. (Individuals subsequently entering into the model were not vaccinated.) Similar to other analyses of the potential effectiveness of HIV vaccines (e.g., Long & Owens, 2011; Long, Brandeau, & Owens, 2009) we assumed that the vaccine has a fixed level of efficacy in reducing the chance that an individual acquires HIV (we considered efficacy levels of 25% and 75%) and that vaccine efficacy does not diminish over the 5-year time horizon. As illustrative examples we considered coverage levels of 25% and 100% of susceptible individuals in the target population receiving the vaccine. We considered three target populations: PWID, MSM/PWID, and MSM. For each vaccine scenario (efficacy, coverage level, and target population) and each model, we calculated the number of HIV infections averted over the time horizon.

Given the different structural complexities of each model, the transmission calculations, and thus total run time, varied greatly. Computational constraints dictated that we could not run every model for the standard 10,000 runs typical for microsimulations. Consequently, we employed a different standard: for each scenario, we ran enough trials for the number of HIV infections averted to weakly converge. In particular, we ran two sets of trials until the means of the sets (projected number of HIV infections averted) converged to within 1% of each other. This was largely insufficient for the costs and QALYs to converge. Therefore, we performed an economic analysis with a limited subset of models (Models 1a and 1b) to examine whether parameter complexity affects not only a dynamic compartmental model's predictive capacities but also its ultimate policy recommendations, which we report in terms of the incremental cost-effectiveness ratio (ICER), the additional cost for each additional unit of benefit gained by a scenario compared to the next best alternative (Gold, 1996). As most infectious disease policy models rely on dynamic compartmental models, this analysis has particular relevance.

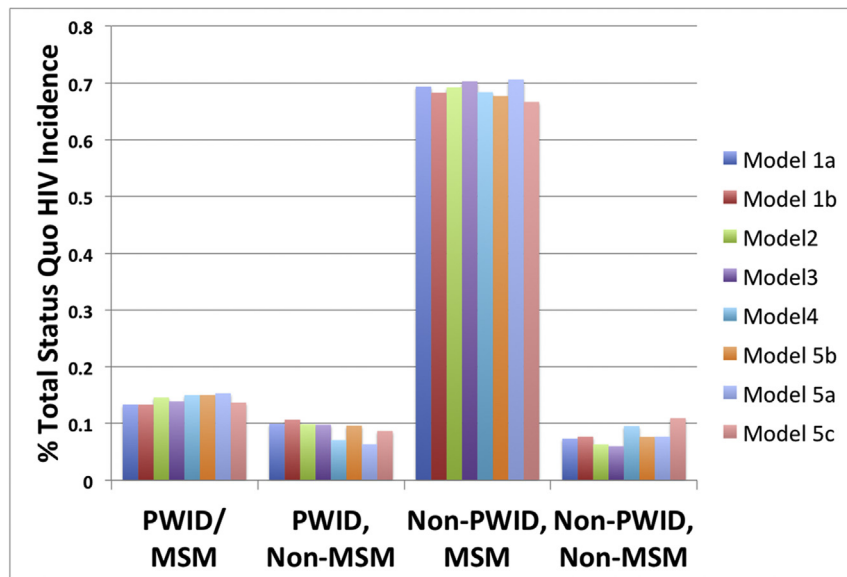


Fig. 2. Model calibration: Distribution of projected HIV incidence over five years across subpopulations in the absence of a vaccine intervention. PWID = people who inject drugs. MSM = men who have sex with men. The eight models were calibrated to a number of epidemiologic targets from King County, Washington (Appendix Figures A.1–A.10), including the distribution of HIV incidence across risk groups as shown here.

3. Results

3.1. Overview

Figs. 3–5 summarize the results from the scenario runs. Because the models are calibrated to similar targets but differ slightly in terms of total HIV incidence predicted under the status quo, we present standardized results in terms of percent of HIV infections averted rather than absolute number of infections averted. Because the models also differ slightly in terms of the distribution of HIV incidence over risk groups, we also calculated the percent of infections averted divided by the percent of infections contributed under the status quo by each risk group (Appendix Figures A.11–A.13); results were qualitatively similar to those in Figs. 3–5. From Figs. 3–5, we observe that the level at which infection force is calculated is the most substantial differentiator between model performance. While the network models and individual compartmental models vary within and across their categories, those differences are overwhelmed by the differences between them and the aggregated compartmental models (Models 1a, 1b, and 2). This difference is especially pronounced in risk groups with smaller population size (the PWID and the PWID/MSM populations). Sections 3.2–3.4 provide detailed analysis and model-to-model comparisons.

3.2. Contact/simulation complexity

Moving along the dimension of contact/simulation complexity (Models 2, 3, 4, and 5b), projected vaccine effectiveness in the target population generally increases as we increase complexity from network to aggregated compartmental models. However, this trend is not consistent. Across the majority of vaccine scenarios and target populations, Model 4, which includes individual event histories, projects higher reductions in HIV incidence than both network Model 5b (more complex) and homogeneous individual compartmental Model 3 (less complex). In fact, under most circumstances, the network model (Model 5b) and the homogeneous individual compartmental model (Model 3) perform similarly. Thus, we cannot conclude from these results that adding simulation/complexity would necessarily improve or change results.

The most consistent differences in projected HIV incidence occur in the transition from disaggregated to aggregated models. Models 1b and 2 (which are aggregated compartment models) often project reductions in HIV incidence far surpassing estimates from the other models, regardless of other forms of complexity, especially in smaller target populations and under assumptions of higher vaccine coverage or efficacy. At lower coverage or efficacy, Models 1b and 2 may generate incidence reductions in line with the disaggregated models or over-predict to a lesser extent (e.g., when evaluating vaccination targeted to the MSM population at 25% coverage and 25% efficacy). This suggests the notion of a threshold net effectiveness of a vaccination program below which an aggregated model performs similarly to a disaggregated model and above which transmission is over-damped and an aggregated model substantially over-predicts HIV incidence reductions relative to a disaggregated model. The threshold also appears to depend on population size (e.g., note the performance of Models 2 and 5b at 100% coverage and 75% efficacy in the MSM versus PWID populations).

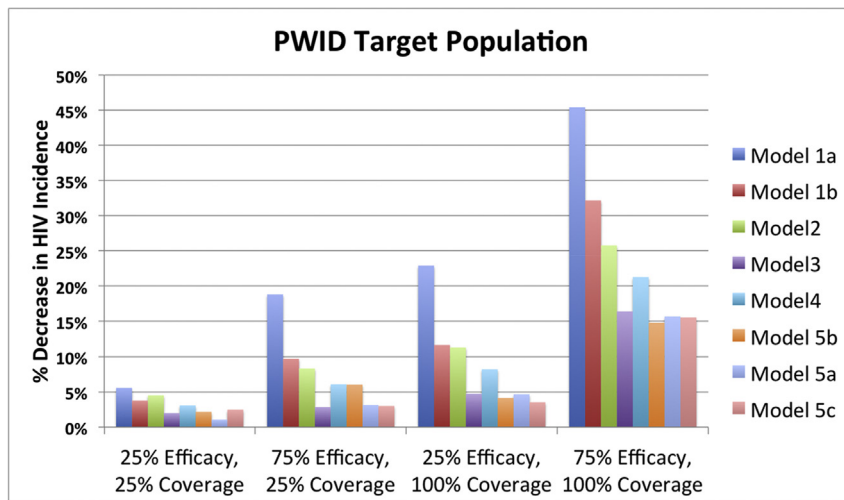


Fig. 3. Vaccine scenarios in the PWID target population: percent decrease in HIV incidence over five years. PWID = people who inject drugs. The figure presents the results of all vaccine scenarios targeted to the PWID population across the eight models. The figure shows the percent of total HIV infections averted.

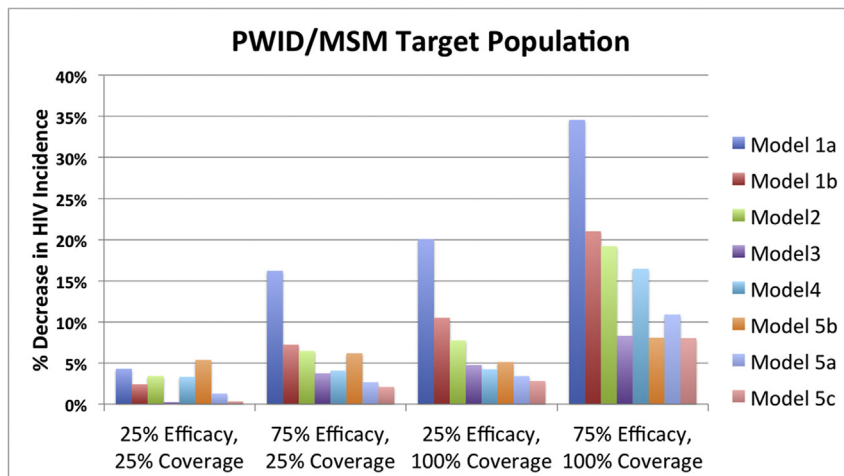


Fig. 4. Vaccine scenarios in the PWID/MSM target population: percent decrease in HIV incidence over five years. PWID = people who inject drugs. MSM = men who have sex with men. The figure presents the results of all vaccine scenarios targeted to the PWID/MSM population (that is, individuals who fall into both categories) across the eight models. The figure shows the percent of total HIV infections averted.

Finally, we note the role of stochasticity in the projections of aggregated models. While Models 1b and 2 both over-predict the reduction in HIV incidence relative to the disaggregated models, Model 2 does so to a lesser extent in the smaller target populations. Recall that Models 1b and 2 approximate the differences between a dynamic compartmental model with and without rounding, respectively (Model 1b only picks up integer values of infections while Model 2's stochasticity allows infection to occur even when population size and transmission probability are low). Thus, we see the large, and compounding, role that rounding – a relatively minor assumption – can play in estimating the effectiveness of an intervention (in this case, HIV vaccination).

3.3. Parameter complexity in aggregated compartmental models

Models 1a and 1b differ only in their parameter complexity. It is useful to compare their “thresholds” directly. While there are scenarios under which Model 1b generates projections in line with the disaggregated models, Model 1a over-predicts the reduction in HIV incidence in every scenario; and when both Models 1a and 1b over-predict, Model 1a tends to over-predict to a greater extent (e.g., observe the progression of net vaccine effectiveness along the x-axis in the PWID target population). We hypothesize that age structure, rather than HCV comorbidity, makes the key difference in terms of parameter complexity in our models. Age structure can pool prevalence into highly concentrated groups so that high transmission probabilities

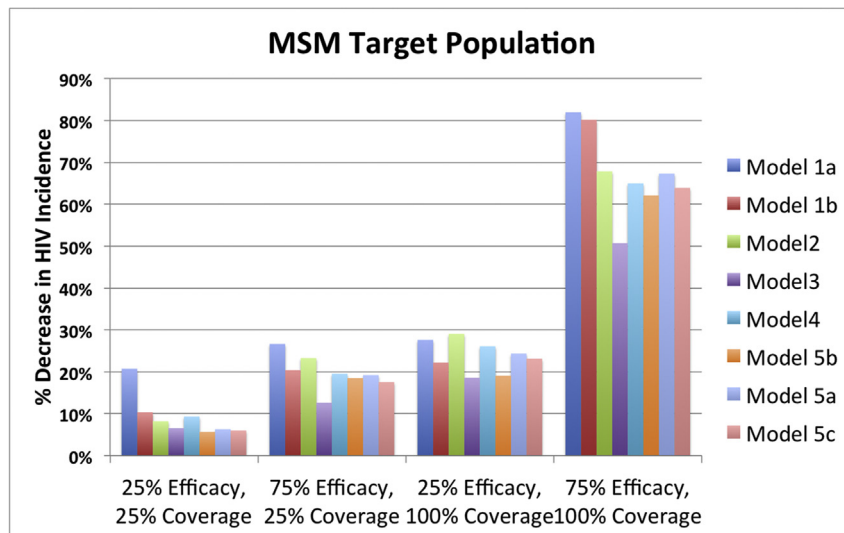


Fig. 5. Vaccine scenarios in the MSM target population: percent decrease in HIV incidence over five years. MSM = men who have sex with men. The figure presents the results of all vaccine scenarios targeted to the MSM population across the eight models. The figure shows the percent of total HIV infections averted.

compensate for small compartment size. Dispersing prevalence throughout each subpopulation in Model 1a exacerbates the over-prediction of Model 1b. It appears that parameter complexity in some circumstances has the potential to act as a buffer against the effects of model aggregation.

While Models 1a and 1b differ in their projections of HIV incidence in the presence of vaccination, it does not necessarily follow that this difference matters from a policy perspective. If the purpose of infectious disease modeling is to inform investment in public health investments, then the cost and QALY findings of an analysis are more important for decision making than potentially arbitrary incidence estimates. Table 4 shows an illustrative subset of economic comparisons between Models 1a and 1b. Although we performed a relatively small number of trial runs, results converged sufficiently to detect distinguishing features between the two models.

Overall, we observe trends that are consistent with the projected reductions in HIV incidence: for example, in the larger target population (MSM) and with higher net vaccine efficacy, parameter complexity differentiates the estimates to a lesser extent. We also observe that ICERs can vary substantially, in terms of both absolute value and sign. In the scenarios we considered, these differences would not lead to substantially different policy recommendations: while Model 1b is less likely to find the interventions cost-saving, all scenarios are considered highly cost-effective at a \$50,000 willingness-to-pay threshold. However, the example illustrates the potential for different policy recommendations arising from models with different structures. Note also that, while the findings in Table 4 are intuitive, they cannot necessarily be predicted in the absence of simulation. The difference in projected reductions in HIV incidence suggests that incremental QALYs would be lower in Model 1b, but it is not obvious that incremental costs would be higher as well, since the inclusion of costs associated with HCV comorbidity and age might have made the value of preventing an HIV infection higher, thus displacing the difference in savings that comes from preventing more infections.

3.4. Parameter complexity in network models

While we found parameter complexity to strongly differentiate aggregated models, we did not find this to be the case for network models. Not only did parameter complexity fail to consistently distinguish between mean projected reductions, it also did not substantially alter the standard deviations of each model's projections (Table A.17). Models 5a, 5b, and 5c do not always agree in their predictions, but they do not differ in a clearly recognizable pattern that suggests that, given this level of contact/simulation complexity, increased parameter complexity refines model predictions.

4. Discussion

It has been conjectured that structural assumptions in infectious disease models, such as the choice of network or compartmental model type or increased heterogeneity across individuals, might affect model predictions as much as or more than the choice of input parameters (Brisson & Edmunds, 2006; Koopman, 2004, 2005, 2007). Yet, because of computational and time constraints, and because of the current focus of most modeling studies on calibration and first-order uncertainty, these structural assumptions are rarely explored in a policy context.

Koopman (2007) underscores the importance of exploring the effects of both contact and simulation complexity (e.g., aggregated compartmental model versus stochastic network model with gradations in between) as well as parameter

Table 4
Effects of parameter complexity on estimated cost-effectiveness of an HIV vaccine.

Scenario	Model 1a				Model 1b			
	Fraction of HIV Infections Averted	Incremental Costs (\$, Thousand)	Incremental QALYs (Thousand)	ICER	Fraction of HIV Infections Averted	Incremental Costs (\$, Thousand)	Incremental QALYs (Thousand)	ICER
PWID: [L, H, L]	0.23	-1645	203	-\$8103	0.12	638	77	\$8286
PWID: [L, H, H]	0.23	120	203	\$591	0.12	1935	77	\$25,130
PWID: [H, H, H]	0.45	-337	362	-\$931	0.32	2058	188	\$10,947
MSM: [L, H, L]	0.28	-2565	243	-\$10,512	0.22	-1804	166	-\$10,867
MSM: [L, H, H]	0.28	-557	243	-\$2292	0.22	361	166	\$2175
MSM: [H, H, H]	0.82	-7018	743	-\$9445	0.80	-5712	664	-\$8602

PWID = people who inject drugs. MSM = men who have sex with men. ICER = incremental cost-effectiveness ratio, expressed as incremental cost per QALY gained. Scenarios are presented as “target population: [vaccine efficacy, vaccine coverage, vaccine cost].” “L” and “H” denote low or high. Low vaccine efficacy is 25%; high is 75%. Low coverage is 25%; high is 100%. Low cost is \$300; high is \$1000. All results are compared to the status quo over a five-year time horizon.

complexity (i.e., increasing model realism by including progressively more risk factors). In our case study, we consider one policy, HIV vaccination, and explore the potential implications of such structural assumptions (contact/simulation complexity and parameter complexity) on our conclusions using the protocols of inference robustness assessment (Koopman, 2005, 2007; Koopman et al., 2016).

We found that complexity did not consistently differentiate network models but had a substantial effect in aggregated compartmental models. The aggregated compartmental models we considered over-predicted the incidence reductions of an HIV vaccine compared to the disaggregated models. These effects were especially pronounced in smaller target populations, but were mitigated by the inclusion of parameter complexity such as age structure. These findings have several implications for infectious disease modelers in the policy context. The concept of a threshold intervention effectiveness below which aggregated and disaggregated models have substantial overlap suggests the possibility of future model-selection guidelines. If a modeler’s intended analysis explores the cost-effectiveness of a relatively low-impact intervention or a higher-impact intervention in a large population, a compartmental, mass-action model may be sufficient.

The finding that parameter complexity can reduce the over-prediction of an aggregated compartmental model suggests that modelers should attempt to capture mixing-relevant aspects of their population, as we hypothesize that appropriately capturing mixing patterns may have greater impact than incorporating other differentiating characteristics such as comorbidities. At a certain point, the amount of complexity required to appropriately capture the macro-network of a population might make defaulting to a network model more efficient in terms of implementation, although computational constraints still point to the value of a dynamic compartmental model. Finally, depending on the intervention being analyzed, modelers already using a network model may not necessarily want to include further parameter complexity given the substantial time investment versus the potential gains in model prediction, and this may in turn guide surveillance data collection.

Our case study has relatively narrow scope, and we emphasize that our findings are illustrative and not prescriptive. We consider a set of stylized policy scenarios under a constrained number of trial runs. We made every effort to isolate the structural assumptions between linked models, but no two models, even ones as closely developed as ours, can entirely control for certain inherent differences that arise during their implementation. Our models are closely but not perfectly calibrated to each other, and slightly differently tuned parameters may have outsized effects not accounted for given the “black-box” nature of model dynamics (more discussion in Appendix). Finally, Models 1a and 1b were designed to approximate aggregated compartmental models but they have a slightly different implementation and their projections are therefore not definitive.

Nonetheless, while our quantitative findings are specific to the particular set of models and scenarios we considered, our qualitative findings are illustrative of broader phenomena and can provide insight for modelers as they consider the appropriate balance of simplicity versus complexity in model structure. Comparative model simulation is the first step in a long process of engagement around structural sensitivity. Our case study suggests several directions for future work.

First, although the models we created are specific to HIV, further simulations can help illustrate the importance of structural uncertainty analysis in other contexts. For example, the World Health Organization has goals for eradication of certain infectious diseases. Comparing the level of intervention required to eliminate transmission across models could be of particular interest and help to inform the appropriate complexity of models used for future goal-setting. Models are also commonly used to estimate real-world parameters from observed data in non-endemic settings (Koopman et al., 2016). A case study similar to ours could “work backwards,” fitting each model to data to estimate the same known parameter in order to document how structural complexity can affect a model’s capacity for inference.

Second, we hypothesize that the value of parameter complexity in network models may be higher when analyzing other interventions, such as NSP, whose value would likely be directly affected both by the inclusion of HCV comorbidity and by quit-rates associated with incarceration. Our current analysis suggests that the impacts of differences in parameter complexity are subsumed by differences in contact/simulation complexity and this further analysis would help to elucidate whether that inference holds across varied policies. Inference robustness assessment can also be applied to modeled network distributions themselves to inform epidemiologists about the value of collecting precise data on social networks.

Finally, our case study provides descriptive analysis of *how* model inferences can vary across dimensions of complexity but it does not answer *why*. We observe from our simulations that smaller populations are more vulnerable to over-estimates of intervention efficacy in models of reduced complexity (Figs. 3–5). Aggregated compartmental models predict that transmission can be suppressed in those populations to such an extent that falling prevalence drives down transmission to and within other, larger populations (Figures A.11–A.13). One hypothesis for this is that aggregation weakens the extent to which small but high-prevalence compartments can concentrate transmission within their own group. Focused exploration between Models 2 and 3 would help test this hypothesis. By examining individual trajectories, one could track how different populations “experience” infection risk over time (e.g., from which groups and to what extent), in the presence and absence of intervention.

More generally, experimentation on how macro-network structure affects the performance of aggregated compartmental models can help to isolate the underlying mechanisms of the threshold property that we observed. A macro-network structure where small populations like PWID/MSM are more or less insular in their mixing or an assumption of sero-sorting might mitigate or worsen the over-prediction effects we currently observe. Increasing our understanding of the link between macro-network assumptions, population size, and thresholds for aggregated compartmental models through simulation can help to inform modelers as to when a given macro-network property in a population would necessitate the use of a non-aggregated model. It may also inform when it is appropriate to ignore geographical heterogeneities and reduce compartments and when it is necessary to explicitly model subpopulations in order not to overestimate intervention effectiveness. Further inference robustness assessment targeted to the areas where we observe the most striking disparities between aggregated compartmental models can help tease apart the relative impact of often bundled assumptions and may lead to valuable analytic insights.

Our analysis contributes to the effort of identifying what makes a model “as detailed as necessary to represent all potentially relevant aspects of the [modeled] process” (Jackson et al., 2011). In models with sufficient contact/simulation complexity, a high level of parameter complexity may not be needed, whereas in models with lower contact/simulation complexity, higher parameter complexity may be called for. Our experiment elucidates the implications of structural assumptions in policy analysis of infectious disease control interventions, and illustrates the value of inference robustness assessment and the importance of further work in this direction.

Competing interest

The authors declare they have no competing interests.

Authors' contributions

CB and MB were responsible for study design and planning. CB performed all analyses. CB and MB wrote the manuscript and approved the final version.

Acknowledgments

Financial support for this study was provided by Grant Number R01-DA15612 from the National Institute on Drug Abuse. CB was supported by a PACCAR Inc. Stanford Graduate Fellowship and National Science Foundation Graduate Fellowship DGE-114747. We thank Stanford University and the Stanford Research Computing Center for providing computational resources and support. CB thanks Sze-Chuan Suen and Rui Fu for productive conversations and insights around structural sensitivity and Konner Robison for his help compiling references.

Appendix A. Supplementary data

Supplementary data related to this article can be found at <https://doi.org/10.1016/j.idm.2017.08.002>.

References

- Adamson, B., Dimitrov, D., Devine, B., & Barnabas, R. (2017). The potential cost-effectiveness of HIV vaccines: A systematic review. *Pharmacoeconomics Open*, 1, 1–12.
- Aleccia, J. (2016). Heroin deaths drop in King County, but addiction still a serious problem. *The Seattle Times*. July 19.
- Arias, E. (2014). National Vital Statistics Report. *United States life tables, 2010* (Vol. 63, pp. 1–63).
- Bansal, S., Grenfell, B. T., & Meyers, L. A. (2007). When individual behaviour matters: Homogeneous and network models in epidemiology. *Journal of The Royal Society Interface*, 4, 879–891.

- Basu, S., & Andrews, J. (2013). Complexity in mathematical models of public health policies: A guide for consumers of models. *PLoS Medicine*, *10*, e1001540.
- Bendavid, E., Brandeau, M. L., Wood, R., & Owens, D. K. (2010). Comparative effectiveness of HIV testing and treatment in highly endemic regions. *Archives of Internal Medicine*, *170*, 1347–1354.
- Bendavid, E., Young, S. D., Katzenstein, D. A., Bayoumi, A. M., Sanders, G. D., & Owens, D. K. (2008). Cost-effectiveness of HIV monitoring strategies in resource-limited settings – a Southern African analysis. *Archives of Internal Medicine*, *168*, 1910–1918.
- Berk Consulting. (2014). *Analysis of statewide adult correctional needs and costs*. Washington State: Final report submitted to the Office of Financial Management, November 6, 2014. http://www.ofm.wa.gov/reports/Correctional_Needs_and_Costs_Study2014.pdf. (Accessed 20 January 2017).
- Bernard, C. L., Brandeau, M. L., Owens, D. K., Humphreys, K., Bendavid, E., Weyant, C., et al. (2016). Cost-effectiveness of HIV preexposure prophylaxis for people who inject drugs in the United States. *Annals of Internal Medicine*, *165*, 10–19.
- Bernard, C. L., Owens, D. K., Goldhaber-Fiebert, J. D., & Brandeau, M. L. (2017). Estimation of the cost-effectiveness of HIV prevention portfolios for people who inject drugs in the United States: A model-based analysis. *PLoS Medicine*, *14*, e1002312.
- Bilcke, J., Beutels, P., Brisson, M., & Jit, M. (2011). Accounting for methodological, structural, and parameter uncertainty in decision-analytic models: A practical guide. *Medical Decision Making*, *31*, 675–692.
- Bird, S. M., Parmar, M. K., & Strang, J. (2015). Take-home naloxone to prevent fatalities from opiate-overdose: Protocol for Scotland's public health policy evaluation, and a new measure to assess impact. *Drugs (Abingdon Engl)*, *22*, 66–76.
- Bojke, L., Claxton, K., Sculpher, M., & Palmer, S. (2009). Characterizing structural uncertainty in decision analytic models: A review and application of methods. *Value in Health*, *12*, 739–749.
- Brennan, A., Chick, S. E., & Davies, R. (2006). A taxonomy of model structures for economic evaluation of health technologies. *Health Economics*, *15*, 1295–1310.
- Brisson, M., & Edmunds, W. J. (2006). Impact of model, methodological, and parameter uncertainty in the economic analysis of vaccination programs. *Medical Decision Making*, *26*, 434–446.
- Broz, D., Wejnert, C., Pham, H. T., DiNenno, E., Heffelfinger, J. D., Cribbin, M., ... Paz-Bailey, G. (2014). Morbidity and Mortality Weekly Report: Surveillance Summaries. *HIV infection and risk, prevention, and testing behaviors among injecting drug users – National HIV Behavioral Surveillance System, 20 U.S. cities, 2009* (Vol. 63, pp. 1–51).
- Buskin, S., Hanrahan, M., & Jaenicke, T. (2013). *HIV/AIDS epidemiology report, first half 2013* (Vol. 82). Seattle, WA: HIV/AIDS Epidemiology Unit, Public Health – Seattle & King County and the Infectious Disease Assessment Unit, Washington State Department of Health.
- Buskin, S., Hanrahan, M., & Jaenicke, T. (2016). *HIV/AIDS epidemiology report 2015* (Vol. 84). Seattle, WA: HIV/AIDS Epidemiology Unit, Public Health – Seattle & King County and the Infectious Disease Assessment Unit, Washington State Department of Health.
- Caro, J. J., Briggs, A. H., Siebert, U., & Kuntz, K. M. (2012). Modeling good research practices—overview: A report of the ISPOR-SMDM modeling good research practices task Force—1. *Value in Health*, *15*, 796–803.
- Chen, N. E., Meyer, J. P., Avery, A. K., Drain, J., Flanigan, T. P., Lincoln, T., ... Altice, F. L. (2013). Adherence to HIV treatment and care among previously homeless jail detainees. *AIDS and Behavior*, *17*, 2654–2666.
- Chen, S. L., & Morgan, T. R. (2006). The natural history of hepatitis C virus (HCV) infection. *International Journal of Medical Science*, *3*, 47–52.
- Chick, S. E., Adams, A. L., & Koopman, J. S. (2000). Analysis and simulation of a stochastic, discrete-individual model of STD transmission with partnership concurrency. *Mathematical Biosciences*, *166*, 45–68.
- Choi, E., Michalski, C. J., Choo, S. H., Kim, G. N., Banasikowska, E., Lee, S., ... Kang, C. Y. (2016). First Phase I human clinical trial of a killed whole-HIV-1 vaccine: Demonstration of its safety and enhancement of anti-HIV antibody responses. *Retrovirology*, *13*, 82.
- Cipriano, L. E., Zanic, G. S., Holodniy, M., Bendavid, E., Owens, D. K., & Brandeau, M. L. (2012). Cost effectiveness of screening strategies for early identification of HIV and HCV infection in injection drug users. *PLoS One*, *7*, e45176.
- Degenhardt, L., Larney, S., Kimber, J., Gisev, N., Farrell, M., Dobbins, T., ... Burns, L. (2014). The impact of opioid substitution therapy on mortality post-release from prison: Retrospective data linkage study. *Addiction*, *109*, 1306–1317.
- Dombrowski, K., Khan, B., Habecker, P., Hagan, H., Friedman, S. R., & Saad, M. (2017). The interaction of risk network structures and virus natural history in the non-spreading of HIV among people who inject drugs in the early stages of the epidemic. *AIDS and Behavior*, *21*, 1004–1015.
- Eaton, J. W., Johnson, L. F., Salomon, J. A., Barnighausen, T., Bendavid, E., Bershteyn, A., ... Hallett, T. B. (2012). HIV treatment as prevention: Systematic comparison of mathematical models of the potential impact of antiretroviral therapy on HIV incidence in South Africa. *PLoS Medicine*, *9*, e1001245.
- Gold, M. (1996). Panel on cost-effectiveness in health and medicine. *Medical Care*, *34*(Suppl), DS197–DS199.
- Grassly, N. C., & Fraser, C. (2008). Mathematical models of infectious disease transmission. *Nature Reviews Microbiology*, *6*, 477–487.
- Hamilton, D. T., Handcock, M. S., & Morris, M. (2008). Degree distributions in sexual networks: A framework for evaluating evidence. *Sexually Transmitted Diseases*, *35*, 30–40.
- Hellard, M., Rolls, D. A., Sacks-Davis, R., Robins, G., Pattison, P., Higgs, P., ... McBryde, E. (2014). The impact of injecting networks on hepatitis C transmission and treatment in people who inject drugs. *Hepatology*, *60*, 1861–1870.
- Huppert, A., & Katriel, G. (2013). Mathematical modelling and prediction in infectious disease epidemiology. *Clinical Microbiology and Infection*, *19*, 999–1005.
- Jackson, C. H., Bojke, L., Thompson, S. G., Claxton, K., & Sharples, L. D. (2011). A framework for addressing structural uncertainty in decision models. *Medical Decision Making*, *31*, 662–674.
- Kapadia, F., Latka, M. H., Hudson, S. M., Golub, E. T., Campbell, J. V., Bailey, S., ... Garfein, R. S. (2007). Correlates of consistent condom use with main partners by partnership patterns among young adult male injection drug users from five US cities. *Drug and Alcohol Dependence*, *1*(91 Suppl), S56–S63.
- King County Department of Adult and Juvenile Detention. (2015). Detention and alternatives report. http://www.kingcounty.gov/~media/courts/detention/documents/KC_DAR_12_2015.ashx?la=en. (Accessed 1 July 2016).
- Kingston, S., & Banta-Green, C. (2016). Results from the 2015 Washington state drug injector health survey. <http://adai.uw.edu/pubs/infobriefs/2015DrugInjectorHealthSurvey.pdf>. (Accessed 5 January 2017).
- Kong, L., Wang, J., Han, W., & Cao, Z. (2016). Modeling heterogeneity in direct infectious disease transmission in a compartmental model. *International Journal of Environmental Research and Public Health*, *13*, E253.
- Koopman, J. (2004). Modeling infection transmission. *Annual Review of Public Health*, *25*, 303–326.
- Koopman, J. S. (2005). Infection transmission science and models. *Japanese Journal of Infectious Diseases*, *58*, S3–S8.
- Koopman, J. S. (2007). Infection transmission through networks. In F. Képès (Ed.), *Biological networks* (pp. 449–505). Singapore: World Scientific Publishing.
- Koopman, J. S., Jacquez, G., & Chick, S. E. (2001). New data and tools for integrating discrete and continuous population modeling strategies. *Annals of the New York Academy of Sciences*, *954*, 268–294.
- Koopman, J. S., Singh, P., & Ionides, E. I. (2016). Public health modeling and data mining. In S. C. J. M. McNabb, L. Ferland, W. MacWright, Z. Memish, S. Okutani, et al. (Eds.), *Transforming public health surveillance* (pp. 327–345). Amman, Jordan: Elsevier.
- Larney, S., Toson, B., Burns, L., & Dolan, K. (2012). Effect of prison-based opioid substitution treatment and post-release retention in treatment on risk of re-incarceration. *Addiction*, *107*, 372–380.
- Lee, E. C., Kelly, M. R., Jr., Ochocki, B. M., Akinwumi, S. M., Hamre, K. E., Tien, J. H., et al. (2017). Model distinguishability and inference robustness in mechanisms of cholera transmission and loss of immunity. *Journal of Theoretical Biology*, *420*, 68–81.
- Lee, J. D., McDonald, R., Grossman, E., McNeely, J., Laska, E., Rotrosen, J., et al. (2015). Opioid treatment at release from jail using extended-release naltrexone: A pilot proof-of-concept randomized effectiveness trial. *Addiction*, *110*, 1008–1014.
- Liu, S., Cipriano, L. E., Holodniy, M., Owens, D. K., & Goldhaber-Fiebert, J. D. (2012). New protease inhibitors for the treatment of chronic hepatitis C: A cost-effectiveness analysis. *Annals of Internal Medicine*, *156*, 279–290.

- Liu, S., Watcha, D., Holodniy, M., & Goldhaber-Fiebert, J. D. (2014). Sofosbuvir-based treatment regimens for chronic, genotype 1 hepatitis C virus infection in U.S. incarcerated populations: A cost-effectiveness analysis. *Annals of Internal Medicine*, *161*, 546–553.
- Long, E. F., Brandeau, M. L., & Owens, D. K. (2009). Potential population health outcomes and expenditures of HIV vaccination strategies in the United States. *Vaccine*, *27*, 5402–5410.
- Long, E. F., & Owens, D. K. (2011). The cost-effectiveness of a modestly effective HIV vaccine in the United States. *Vaccine*, *29*, 6113–6124.
- Maryland Department of Health and Mental Hygiene. (2014). Risk of overdose death following release from prison or jail. https://bha.health.maryland.gov/OVERDOSE_PREVENTION/Documents/corrections%20brief_V3.pdf. (Accessed 1 October 2016).
- Monteiro, J. F., Escudero, D. J., Weinreb, C., Flanigan, T., Galea, S., Friedman, S. R., et al. (2016). Understanding the effects of different HIV transmission models in individual-based microsimulation of HIV epidemic dynamics in people who inject drugs. *Epidemiology and Infection*, *144*, 1683–1700.
- National Institute on Drug Abuse. (2017). Overdose death rates, revised 2017. <https://www.drugabuse.gov/related-topics/trends-statistics/overdose-death-rates>. (Accessed 29 June 2017).
- National Institutes of Health. (2017). Large-scale HIV vaccine trial to launch in South Africa: NIH-funded study will test safety, efficacy of vaccine regimen. Press release. <https://www.nih.gov/news-events/news-releases/large-scale-hiv-vaccine-trial-launch-south-africa>. (Accessed 18 May 2017).
- Public Health Agency of Canada. (2012). HIV transmission risk: A summary of the evidence. <http://www.catie.ca/sites/default/files/HIV-TRANSMISSION-RISK-EN.pdf>. (Accessed 29 June 2017).
- Rahmandad, H., & Sterman, J. D. (2008). Heterogeneity and network structure in the dynamics of diffusion: Comparing agent-based and differential equation models. *Management Science*, *54*, 998–1014.
- Rahmandad, H., & Sterman, J. D. (2012). Reporting guidelines for simulation-based research in social sciences. *System Dynamics Review*, *28*, 396–411.
- Rich, J. D., McKenzie, M., Larney, S., Wong, J. B., Tran, L., Clarke, J., ... Zaller, N. (2015). Methadone continuation versus forced withdrawal on incarceration in a combined US prison and jail: A randomised, open-label trial. *The Lancet*, *386*, 350–359.
- Scott, J. A., & Chew, K. W. (2017). Treatment optimization for HIV/HCV co-infected patients. *Therapeutic Advances in Infectious Disease*, *4*, 18–36.
- Scott, N., Hellard, M., & McBryde, E. S. (2016). Modeling hepatitis C virus transmission among people who inject drugs: Assumptions, limitations and future challenges. *Virulence*, *7*, 201–208.
- Seattle and King County Public Health Department. (2009). Facts about HIV/AIDS in incarcerated people. <http://www.kingcounty.gov/healthservices/health/communicable/hiv/epi/-/media/health/publichealth/documents/hiv/HIVIncarcerated.ashx>. (Accessed 15 June 2016).
- Silal, S. P., Little, F., Barnes, K. I., & White, L. J. (2016). Sensitivity to model structure: A comparison of compartmental models in epidemiology. *Health Systems*, *5*, 178–191.
- Sionean, C., Le, B. C., Hageman, K., Oster, A. M., Wejnert, C., Hess, K. L., et al. (2014). Morbidity and Mortality Weekly Report: Surveillance Summaries. *HIV risk, prevention, and testing behaviors among heterosexuals at increased risk for HIV infection – National HIV Behavioral Surveillance System, 21 U.S. Cities, 2010* (Vol. 63, pp. 1–39).
- Suen, S. C., Goldhaber-Fiebert, J. D., & Brandeau, M. L. (2017). Risk stratification in compartmental epidemic models. Where to draw the line? *Journal of Theoretical Biology* (in press).
- University of Washington. (2017). Hepatitis C. <http://www.hepatitisc.uw.edu/>. (Accessed 3 April 2017).
- US Census Bureau. (2017). Population demographics for king County, Washington in 2016 and 2017. <https://suburbanstats.org/population/washington/how-many-people-live-in-king-county>. (Accessed 29 June 2017).
- Valdiserri, R. O., & Koh, H. K. (2014). Breaking the silence on viral hepatitis. *Annals of Internal Medicine*, *161*, 147–148.
- Vance-Sherman, A. (2015). King County profile. <https://fortress.wa.gov/esd/employmentdata/reports-publications/regional-reports/county-profiles/king-county-profile>. (Accessed 1 June 2016).
- van de Vijver, D. A., Nichols, B. E., Abbas, U. L., Boucher, C. A., Cambiano, V., Eaton, J. W., ... Hallett, T. B. (2013). Preexposure prophylaxis will have a limited impact on HIV-1 drug resistance in sub-saharan africa: A comparison of mathematical models. *AIDS*, *27*, 2943–2951.
- Weinstein, M. C., O'Brien, B., Hornberger, J., Jackson, J., Johannesson, M., McCabe, C., et al. (2003). Principles of good practice for decision analytic modeling in health-care evaluation: Report of the ISPOR task force on good research practices—modeling studies. *Value in Health*, *6*, 9–17.
- White, L. J., Maude, R. J., Pongtavornpinyo, W., Saralamba, S., Aguas, R., Van Effelterre, T., ... White, N. J. (2009). The role of simple mathematical models in malaria elimination strategy design. *Malaria Journal*, *8*, 212.
- Zhong, H., Arjmand, I. K., Brandeau, M. L., & Bendavid, E. (2017). Clinical outcomes and cost-effectiveness of treating depression in HIV-infected populations in sub-saharan africa: A model-based analysis. *Working Paper*.