Web Extra Material for "Dynamics of the HIV outbreak and response in Scott County, Indiana, 2011-2015"

Gregg S. Gonsalves and Forrest W. Crawford

1 Detailed methods

1.1 A deterministic model of HIV transmission in Scott County, IN

Consider a population of size N at risk for HIV infection, in which each individual can be classified into one of four categories: susceptible HIV-, infectious HIV+ but undiagnosed, infectious HIV+ diagnosed, and removed. This model is a generalization of the classical susceptible-infectious-removed (SIR) model commonly used to describe infectious disease transmission in a population[1]. In this context, removal means cessation of epidemiologic contact sufficient to transmit HIV infection, including successful virologic suppression following diagnosis, use of clean needles, practicing safer sex, or otherwise preventing transmission.

Let $S(t)$ be the number of susceptible individuals, $I_{udx}(t)$ the number of undiagnosed infectious individuals, $I_{dx}(t)$ the number of diagnosed infections, and $R(t)$ the number of "removed" individuals at time t, with $S(t) + I_{udx}(t) + I_{dx}(t) + R(t) = N$. Suppose at time t, each susceptible individual becomes infected with rate proportional to the number of infectious individuals in the population, $I_{udx}(t) + I_{dx}(t)$. Suppose each undiagnosed HIV+ individual is diagnosed with rate $\gamma(t)$; we call $\gamma(t)$ the casefinding/diagnosis rate. Each diagnosed HIV+ individual removed from the group of infectious individuals with rate ρ. The model dynamics are given by the ordinary differential equations

$$\begin{aligned}
\frac{\mathrm{dS}}{\mathrm{dt}} &= -\beta S(t)(I_{udx}(t) + I_{dx}(t)) \\
\frac{\mathrm{dI_{udx}}}{\mathrm{dt}} &= \beta S(t)(I_{udx}(t) + I_{dx}(t)) - \gamma(t) I_{udx}(t) \\
\frac{\mathrm{dI_{dx}}}{\mathrm{dt}} &= \gamma(t) I_{udx}(t) - \rho I_{dx}(t) \\
\frac{\mathrm{dR}}{\mathrm{dt}} &= \rho I_{dx}(t)
\end{aligned} \tag{1}$$

for $\beta > 0$, $\rho > 0$, and a possibly time-varying non-negative function $\gamma(t)$, with initial conditions $S(0)$, $I_{udx}(0)$, $I_{dx}(0)$ and $R(0)$ and conservation equation $S(t) + I_{dx}(t) + I_{udx}(t) + R(t) = N$. As ρ becomes larger, the model dynamics approach those of a traditional time-varying SIR model with time-varying removal rate $\gamma(t)$. The infection rate β has units "infections per susceptible-infectious pair, per day". The casefinding/diagnosis rate $\gamma(t)$ has units "diagnoses per undiagnosed HIV+ person per day". The removal rate ρ has units "removals per diagnosed HIV+ person per day".

1.2 Reconstruction of bounds for SIR dynamics from Scott County outbreak data

From the published data[2] we obtained the cumulative HIV diagnoses $D(t)$. From the recency assay data[3] we obtained lower and upper bounds $\underline{C}(t)$ and $\overline{C}(t)$ for the cumulative HIV incidence $C(t)$. Limited information is available on the number of individuals N at risk for HIV infection during the Scott County Outbreak; for now we will assume N is fixed and known. The number of undiagnosed HIV infections at time t is the cumulative number of infections by time t minus the number of diagnosed infections,

$$I_{udx}(t) = C(t) - D(t),$$

and the number of susceptible individuals at time t is

$$S(t) = N - C(t).$$

We obtain lower and upper bounds for $I_{udx}(t)$, $I_{dx}(t)$, and $S(t)$ as

$$\begin{aligned}
\underline{I}_{udx}(t) &= \underline{C}(t) - D(t), \\
\overline{I}_{udx}(t) &= \overline{C}(t) - D(t),
\end{aligned}$$

$$\underline{S}(t) = N - \overline{C}(t),$$

and

$$\overline{S}(t) = N - \underline{C}(t).$$

The time-varying diagnosis rate $\gamma(t)$ can be reconstructed by considering the rate of diagnoses as a function of the number of undiagnosed infections,

$$\gamma(t) = \frac{dD(t)}{I_{udx}(t)dt}.$$

We compute lower and upper bounds for $\gamma(t)$ as

$$\underline{\gamma}(t) = \frac{dD(t)}{\overline{I}_{udx}(t)dt},$$

and

$$\overline{\gamma}(t) = \frac{dD(t)}{\underline{I}_{udx}(t)dt}.$$

The transmission rate β can be estimated by dividing the number of new infections by the person-time of infectiousness,

$$\beta = \frac{C(t) - C(0)}{\int_0^t S(u)(I_{udx}(u) + I_{dx}(u))du}.$$

We compute lower and upper bounds for β as

$$\underline{\beta} = \frac{\underline{C}(t) - \overline{C}(0)}{\int_0^t \overline{S}(u)(\overline{I}_{udx}(u) + \overline{I}_{dx}(u))du} \tag{2}$$

and

$$\overline{\beta} = \frac{\overline{C}(t) - \underline{C}(0)}{\int_0^t \underline{S}(u)(\underline{I}_{udx}(u) + \underline{I}_{dx}(u))du}. \tag{3}$$

1.3 Evaluation of counterfactual intervention scenarios

Let t_s be the time of the first HIV diagnosis in Scott County, IN, and let t_e be a later time, for example the time at which diagnostic scaleup ceased. For a hypothetical earlier first diagnosis time $t_s^* < t_s$, define the counterfactual diagnosis rate as

$$\gamma^*(t) = \begin{cases} \gamma(t) & \text{if } t < t_s^* \\ \gamma(t_s - t_s^* + t) & \text{if } t_s^* \le t < t_s^* - t_s + t_e \\ \gamma(t_e) & \text{if } t_s^* - t_s + t_e \le t \end{cases}$$

Define $\underline{\gamma}^*(t)$ and $\overline{\gamma}^*(t)$ by substituting $\underline{\gamma}(t)$ and $\overline{\gamma}(t)$ for $\gamma(t)$ above. The resulting diagnosis rate is a replica of the true diagnosis rate during the outbreak response, shifted to the earlier starting time t_s^*.

Define the counterfactual SIR dynamics under earlier intervention at t_s^* as

$$\begin{aligned} \frac{\mathrm{dS}^*}{\mathrm{dt}} &= -\beta S^*(t)(I_{udx}^*(t) + I_{dx}^*(t)) \\ \frac{\mathrm{dI}_{\mathrm{udx}}^*}{\mathrm{dt}} &= \beta S^*(t)(I_{udx}^*(t) + I_{dx}^*(t)) - \gamma(t) I_{udx}^*(t) \\ \frac{\mathrm{dI}_{\mathrm{dx}}^*}{\mathrm{dt}} &= \gamma(t) I_{udx}^*(t) - \rho I_{dx}^*(t) \\ \frac{\mathrm{dR}^*}{\mathrm{dt}} &= \rho I_{dx}^*(t) \end{aligned} \tag{4}$$

with initial conditions $S^*(0)$, $I_{udx}^*(0)$, $I_{dx}^*(0)$ and $R^*(0)$. This system has identical dynamics to (1) before t_s^*. That is, for $t < t_s^*$, $S^*(t) = S(t)$, $I_{udx}^*(t) = I_{udx}(t)$, $I_{dx}^*(t) = I_{dx}(t)$, and $R^*(t) = R(t)$. Subsequently, for $t > t_s^*$, the system evolves under the counterfactual rate $\gamma^*(t)$. Define the lower bounds $\underline{S}^*(t)$, $\underline{I}_{udx}^*(t)$, $\underline{I}_{dx}^*(t)$, and $\underline{R}^*(t)$, by substituting

Term	Description
$S(t)$	Number of susceptibles at time t
$I_{udx}(t)$	Number of HIV+ undiagnosed individuals at time t
$I_{dx}(t)$	Number of HIV+ diagnosed individuals at time t
$R(t)$	Number of removed individuals at time t
$\beta(t)$	Transmission rate
$\gamma(t)$	Diagnosis rate
ρ	Removal rate for diagnosed individuals
$\gamma^*(t)$	Counterfactual diagnosis rate

Table 1: Notation and symbols used to define the HIV transmission model.

$\underline{\gamma}^*(t)$ for $\gamma^*(t)$ in (4). Likewise, define the upper bounds $\overline{S}^*(t)$, $\overline{I}^*_{udx}(t)$, $\overline{I}^*_{dx}(t)$, and $\overline{R}^*(t)$ by substituting $\overline{\gamma}^*(t)$ for $\overline{\gamma}^*(t)$ in (4). Under the model specification (4), the model output delivers bounds for the dynamics that would have occured if, contrary to fact, the public health intervention program had been implemented at an earlier date. Table 1 summarizes the notation and symbols used to define the mathematical model.

1.4 Evaluation of reduction in transmission rate

Here we define an intervention on the transmission rate β for use in the sensitivity analysis below. Define $0 \leq \epsilon \leq 1$ as the proportion reduction in β following implementation of an intervention at time t^*_s. Fix a pre-intervention value β and let

$$\beta^*(t) = \begin{cases} \beta & \text{if } t < t^*_s \\ \epsilon\beta & \text{if } t \geq t^*_s \end{cases}$$

Define $\underline{\beta}^*(t)$ and $\overline{\beta}^*(t)$ by substituting $\underline{\beta}$ and $\overline{\beta}$ for β above. The resulting diagnosis rate is a replica of the true diagnosis rate during the outbreak response, shifted to the earlier starting time t^*_s.

1.5 Evaluation of increasing uncertainty in cumulative HIV incidence

Define the midpoint of the HIV incidence interval

$$m(t) = \frac{\overline{C}(t) + \underline{C}(t)}{2}.$$

To evaluate the effect of increasing uncertainty about cumulative incidence $C(t)$ at time t, let $f > 1$ and define

$$\overline{C}^*(t) = m(t) + f \times (\overline{C}(t) - m(t))$$

and

$$\underline{C}^*(t) = m(t) - f \times (m(t) - \overline{C}(t)).$$

Then the new interval defined by $[\underline{C}^*(t), \overline{C}^*(t)]$ contains interval $[\underline{C}(t), \overline{C}(t)]$ for every t. We evaluate the sensitivity of results to values of f in the interval $[1, 2]$ below.

1.6 Discrete-time system

In the analysis presented in the main manuscript, we discretize time in units of one day. At day t, we produce the model output at day $t + 1$ by the iteration

$$\begin{aligned} S(t+1) &= S(t) - \beta S(t)(I_{udx}(t) + I_{dx}(t)) \\ I_{udx}(t+1) &= I_{udx}(t) + \beta S(t)(I_{udx}(t) + I_{dx}(t)) - \gamma(t) I_{udx}(t) \\ I_{dx}(t+1) &= I_{dx}(t) + \gamma(t) I_{udx}(t) - \rho I_{dx}(t) \\ R(t+1) &= R(t) + \rho I_{dx}(t). \end{aligned} \tag{5}$$

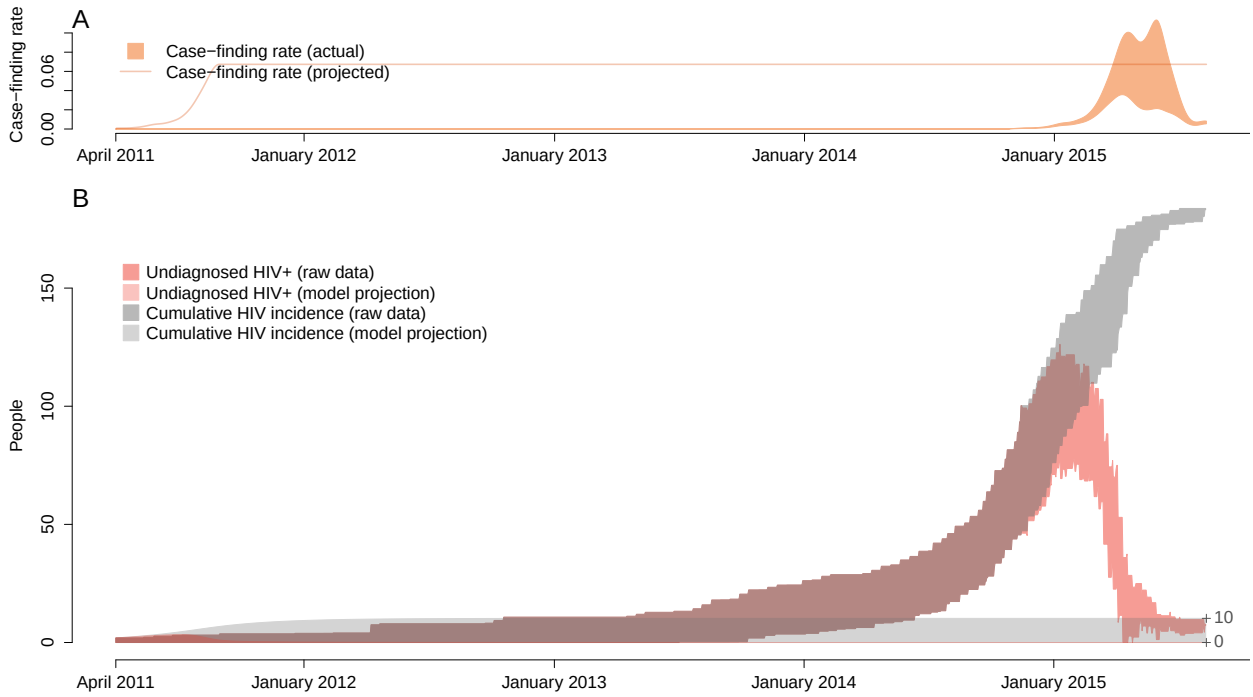

Figure 1: Evaluation of projected outbreak dynamics under a counterfactual intervention date of April 4, 2011. (A) Counterfactual case-finding rate. (B) Cumulative HIV incidence (gray) and undiagnosed HIV infections (red) in the actual outbreak and under earlier intervention. In this scenario, cumulative HIV incidence by August 2015 is projected to be at most 10 people, compared to the actual number183-184.

The available outbreak data $\underline{C}(t)$, $\overline{C}(t)$, $\underline{R}(t)$, and $\overline{R}(t)$ are discretely observed on different timescales. To facilitate projection of trajectories over time, we used smoothing techniques to interpolate trajectories at the daily level. We implemented spline, loess, and kernel smoothers.

2 Further results

2.1 Intervention in April 2011

We analyzed the impact of intervening on April 4, 2011. Figure 1 shows the results. The number of undiagnosed infections stays very lowe througout the epidemic, while the cumulative incidence by August 2015 is reduced to at most 10 infections.

2.2 Model calibration and sensitivity

Some prior information is available about expected HIV incidence rates among injection drug users. When the raw data $C(t)$ and $D(t)$ do not identify model parameters, we seek to calibrate these parameters using data in the literature, and evaluate how the results change as a function of these parameters.

2.2.1 Transmission rate β

Upper and lower bounds for β can be calculated from the raw data $C(t)$ and $D(t)$, in conjunction with the model assumptions (1). For $\rho = 0.024$, we calculate that estimates for β (presented in the main text). We can validate this interval estimate as follows. Peters et al.[2] report 841 syringe sharing relationships among 536 individuals in a contact tracing investigation (comprising $\binom{536}{2}$ possible connections) of the HIV risk population in Scott County. Peters et al.[2] report that individuals in this population engage in between 2 and 15 injections per day. The probability of sharing a

needle is between 0.565 and 0.75, and the probability of HIV transmission between serodiscordant injection partners per sharing event is estimated to be approximately 0.0067[4–7]. Multiplying these proportions, we find lower and upper bounds for β,

$$\underline{\beta} = \frac{841}{\binom{536}{2}} \times 2 \times 0.565 \times 0.0067 \approx 5.1 \times 10^{-5}$$

$$\overline{\beta} = \frac{841}{\binom{536}{2}} \times 15 \times 0.75 \times 0.0067 \approx 4.3 \times 10^{-4}$$

new HIV infections per susceptible-infectious pair.

2.2.2 Removal rate ρ

The removal rate ρ controls the rate of viral suppression or cessation of infectious contact in HIV+ diagnosed individuals. While information on the diagnosis rate is readily available from the raw data, ρ cannot be directly estimated. An upper bound for the rate of removal can be calculated from known data from individuals who are diagnosed and immediately receive ART. The viral suppression interval is roughly 6 weeks, or 42 days, leading to an estimate of

$$\overline{\rho} = 1/42 \approx 0.024$$

removals per HIV+ diagnosed person per day. Information from the Scott County outbreak can give clues to a possible lower bound for ρ. Despite logistical challenges, many HIV+ diagnosed individuals eventually received comprehensive HIV care, including ART. Janowicz[8] writes,

> At the end of 2015, among 176 individuals who were eligible for HIV treatment, 86% had been engaged in care, 74% had undergone care coordination, 59% had been prescribed antiretroviral therapy, and 32% had achieved virologic suppression.

First, 32% of 175 is approximately 56. If individuals achieved viral suppression approximately in the order in which they were diagnosed, then the time to viral suppression is at least the time required for the 56th diagnosed individual to achieve suppression on or before January 1, 2016. The 56th diagnosis occured around March 8, 2015, leaving at most 299 days to January 1, 2016. The rate of viral suppression for HIV+ diagnosed individuals can therefore be estimated by

$$\underline{\rho} = 1/299 \approx 0.0033$$

removals per HIV+ diagnosed person per day. Since both of these estimates only take into account viral suppression, and not behavioral change leading to loss of infectious ness or cessation of infectious contact, they are likely to be under-estimates of the true removal rate ρ. In the main text, we fix $\rho = 0.024$, but analyze sensitivity of results to different choices of ρ below.

2.3 Sensitivity analysis: N and ρ

Figure 2 shows projected cumulative HIV infections by August 2015, as a function of the risk population size N. The top panel shows cumulative HIV incidence for intervention on April 5, 2011; the middle panel shows intervention on January 1, 2013; and the bottom panel shows intervention under actual circumstances. In both early intervention scenarios, cumulative HIV incidence is largely invariant to assumed population size N. In the actual scenario shown at bottom, there is greater variation in cumulative incidence – larger assumed population sizes lead to more infections.

Figure 3 shows projected cumulative HIV infections by August 2015, as a function of the removal rate ρ for diagnosed individuals, for three intervention scenarios. The top panel shows cumulative HIV incidence for intervention on April 5, 2011; the middle panel shows intervention on January 1, 2013; and the bottom panel shows intervention under actual circumstances. In all cases, implausibly low values of ρ lead to HIV infections in the entire risk population. For plausible larger values of ρ, both early intervention scenarios show dramatic reduction in cumulative HIV incidence. In the actual intervention scenario shown at bottom, cumulative incidence is not substantially modified by larger values of ρ and increasing ρ has little effect on cumulative infections. This happens because intervention in late 2014 occurrs too late to avert most infections, and a higher rate of removals does little to decrease transmission.

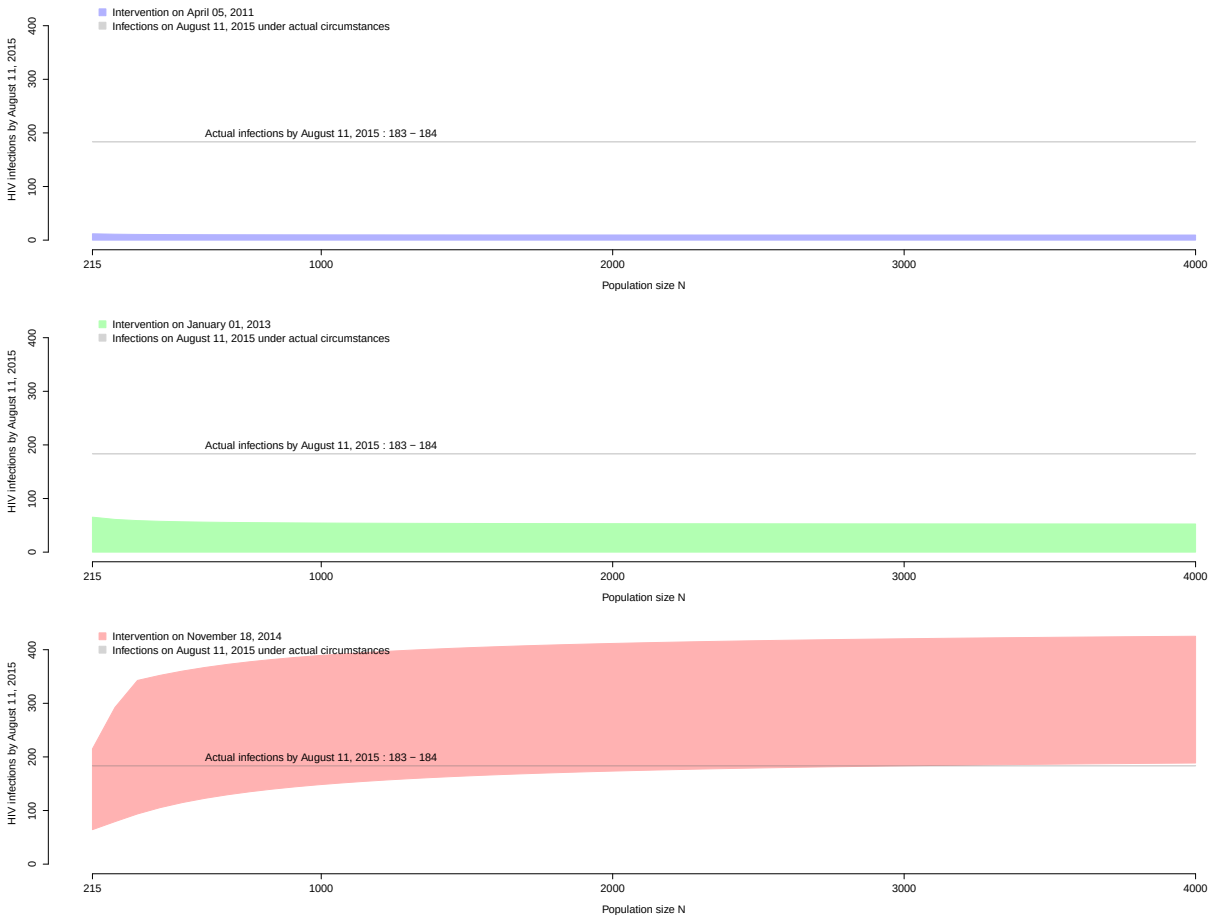

Figure 2: Projected cumulative HIV infections by August 2015, for three intervention date scenarios, as a function of the assumed risk population size N.

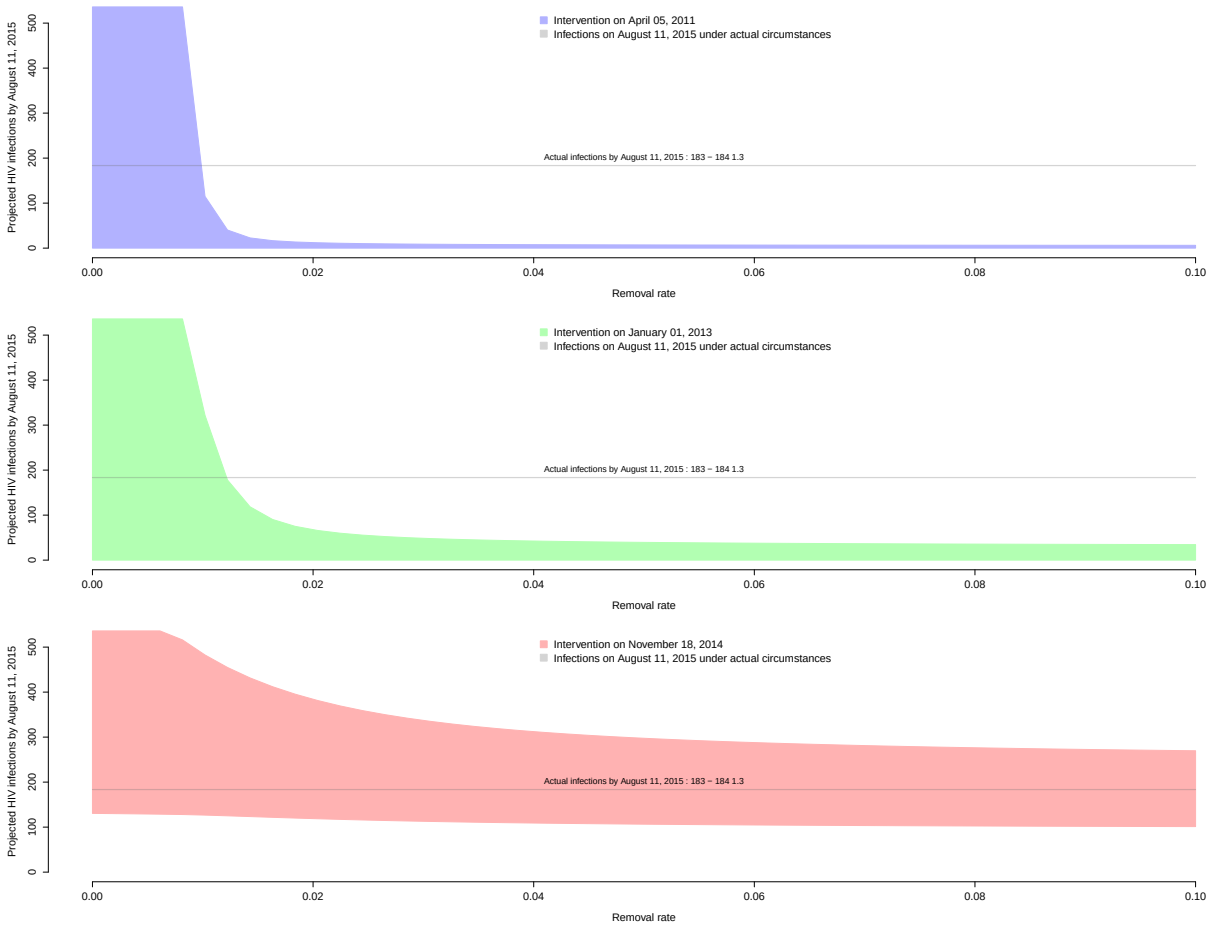

Figure 3: Projected cumulative HIV infections by August 2015, for three intervention date scenarios, as a function of the assumed removal rate ρ for diagnosed individuals.

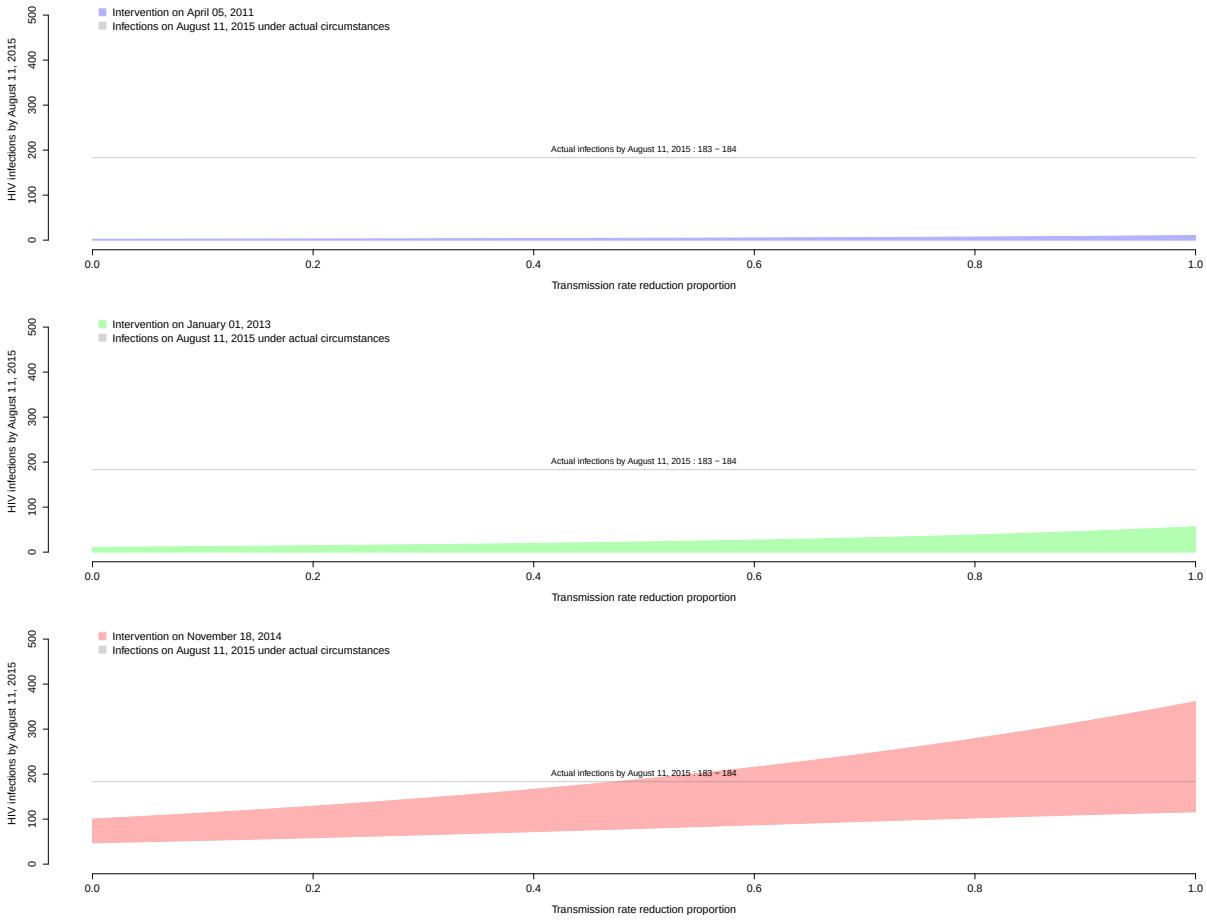

Figure 4: Projected cumulative HIV infections by October 2015, for three intervention date scenarios, as a function of the transmission reduction portion.

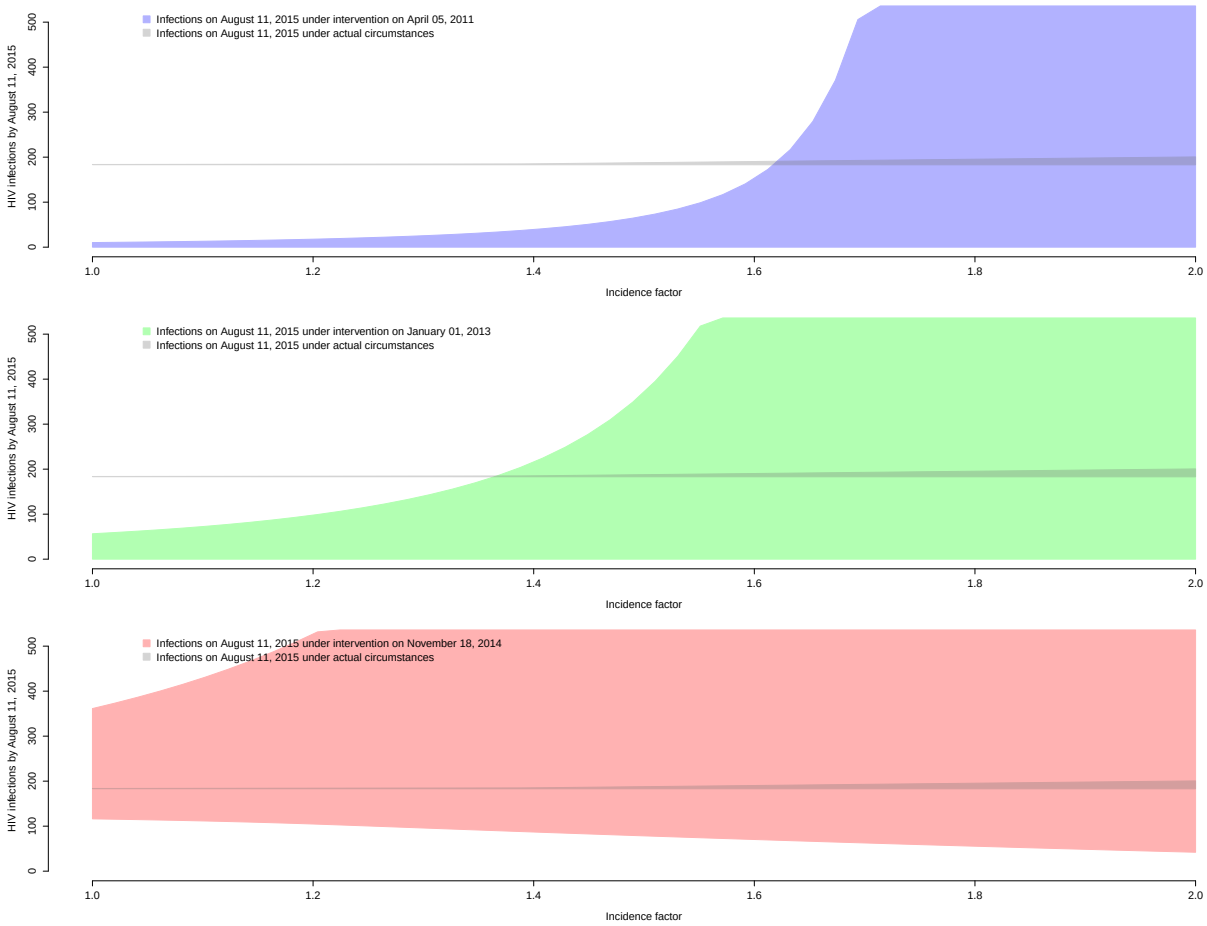

Figure 5: Projected cumulative HIV infections by October 2015, for three intervention date scenarios, as a function of the incidence uncertainty scale factor.

2.4 Sensitivity analysis: transmission reduction and incidence uncertainty

Figure 4 shows projected cumulative HIV infections by October 2015, as a function of the transmission rate reduction factor ϵ, defined above in Section 1.4. The top panel shows cumulative HIV incidence for intervention on April 5, 2011; the middle panel shows intervention on January 1, 2013; and the bottom panel shows intervention under actual circumstances. In both early intervention scenarios, even large values of ϵ – corresponding to little reduction in transmission risk – lead to substantial reductions in cumulative HIV incidence. In the actual intervention scenario, very small values of ϵ lead to projections below the actual number of HIV infections. Larger values of ϵ in this case lead to incidence bounds matching the observed incidence.

Figure 5 shows projected cumulative HIV infections by October 2015, as a function of the incidence uncertainty scale-up factor f, defined in Section 1.5 above. The top panel shows cumulative HIV incidence for intervention on April 5, 2011; the middle panel shows intervention on January 1, 2013; and the bottom panel shows intervention under actual circumstances. For both early intervention scenarios, projected cumulative incidence is at or below the actual incidence for f smaller than approsximately 1.4. In the actual intervention scenario, even modest increases in uncertainty in HIV incidence lead to dramatic increases in proejcted incidence.

References

[1] Roy M Anderson, Robert M May, and B Anderson. *Infectious Diseases of Humans: Dynamics and Control*, volume 28. Wiley Online Library, 1992.

[2] Philip J Peters, Pamela Pontones, Karen W Hoover, Monita R Patel, Romeo R Galang, Jessica Shields, Sara J Blosser, Michael W Spiller, Brittany Combs, William M Switzer, et al. HIV infection linked to injection use of oxymorphone in Indiana, 2014–2015. *New England Journal of Medicine*, 375(3):229–239, 2016.

[3] Ellsworth M Campbell, Hongwei Jia, Anupama Shankar, Debra Hanson, Wei Luo, Silvina Masciotra, S Michele Owen, Alexandra M Oster, Romeo R Galang, Michael W Spiller, et al. Detailed transmission network analysis of a large opiate-driven outbreak of HIV infection in the United States. *The Journal of Infectious Diseases*, 216(9): 1053–1062, 2017.

[4] Gerald H Friedland and Robert S Klein. Transmission of the human immunodeficiency virus. *New England Journal of Medicine*, 317(18):1125–1135, 1987.

[5] Matt D Gaughwin, Eric Gowans, Robert Ali, and Christopher Burrell. Bloody needles: the volumes of blood transferred in simulations of needlestick injuries and shared use of syringes for injection of intravenous drugs. *AIDS*, 5(8):1025–1027, 1991.

[6] Edward H Kaplan and Robert Heimer. A model-based estimate of HIV infectivity via needle sharing. *Journal of Acquired Immune Deficiency Syndromes*, 5(11):1116–1118, 1992.

[7] Edward H Kaplan and Elaine O'Keefe. Let the needles do the talking! Evaluating the New Haven needle exchange. *Interfaces*, 23(1):7–26, 1993.

[8] DM Janowicz. Hiv transmission and injection drug use: Lessons from the indiana outbreak. *Topics in antiviral medicine*, 24(2):90, 2016.