Technical Appendix:
HIV transmission and conception estimation in serodiscordant couples

This is a supplemental document explaining the mathematical model developed to calculate probabilities of different possible outcomes during conception in serodiscordant couples where the man is HIV-infected and the woman is HIV-uninfected. Among various options, the model incorporates the ability to use pre-exposure prophylaxis (PrEP) and/or anti-retroviral therapy (ART) to reduce infection risk. From the model output, the benefit of PrEP and ART can be evaluated both separately and in conjunction with additional risk inputs. The model is developed for simulations using the statistical software R and prototyped as an interactive tool in Excel.

Contents

1 Introduction

We developed a model to estimate the likelihood of possible outcomes defined in terms of HIV infection and successful pregnancy of uninfected women who engage in unprotected sex with infected male partners. The likelihoods depend upon the number of condomless sex acts, biological parameters defining HIV infectivity (such as the male using ART or having other sexually transmitted infections (STIs)), and age-based female fertility. This technical appendix explains two transmission scenarios, mathematical details of the model outcomes, the model equation inputs and outputs, and results of running the model across various clinical settings.

2 Transmission Risk Scenarios Explored

The likelihood of a women remaining HIV-uninfected and having a child via condomless sex with an HIV-infected male is explored under two transmission scenarios that depend upon the frequency of condomless sex. The first is an *optimal* clinical situation where condomless sex acts are limited to the women's ovulation window. From a clinical perspective, the second scenario is less ideal and termed *suboptimal*; this scenario assumes the condomless sex acts are not limited to the ovulation window. The *optimal* scenario represents an adherent serodiscordant couple exhibiting ideal behavior and following clinical counseling recommendations to limit condomless sex to an approximate three-day window (the ovulation day and two previous days when conception is most likely to occur) [1, 2]. In this setting we assume that no sex occurs outside the ovulation window, such that HIV transmission and the chance of pregnancy are reduced to zero. On the contrary, HIV transmission can occur over the entire month in the *suboptimal* scenario, yet pregnancy is only possible over the same three-day ovulation window. In both cases, the number of condomless sex is uniformly distributed over the time-frame of condomless sexual activity (i.e., three days for the *optimal* scenario and 30 days for the *suboptimal* scenario).

3 Modeling the Outcomes

3.1 Parameters

The parameters that enter the model equations used to estimate the outcome probabilities are:

- α: The per-unprotected-sex-act base male-to-female transmissibility when the male is in the early stage of HIV.
- h_L: The multiplicative factor that multiples α to provide the per-unprotected-sex-act base male to female transmissibility when the male is in the late stage of HIV.
- h_{ART}: The multiplicative factor that multiplies α to provide the per-unprotected-sex-act base male-to-female transmissibility when the male is on ART.
- h_{STIs}: The multiplicative factor that multiplies α to provide the per-unprotected-sex-act base male-to-female transmissibility when either partner has other STIs (e.g., genital herpes).
- h_{PrEP}: The multiplicative factor that multiplies α to provide the per-unprotected-sex act base male-to-female transmissibility when the female is taking PrEP.
- $p_{c(a)}$: The probability per unprotected sex act over the fertility period that the female will conceive and become pregnant at age a.
- $p_{d(a)}$: The probability that a pregnant female will deliver a baby if she becomes pregnant at age a.
- p_{MTCT}: The probability that an HIV-infected (during conception) and pregnant female will infect her baby during pregnancy.
- h_{TxMTCT}: The multiplicative factor that multiplies p_{MTCT} to provide the effective mother to child transmission (MTCT) probability when the pregnant female is placed on ART for the duration of her pregnancy.

- M: The number of unprotected sex acts over the ovulation window that an uninfected women engages with an infected male partner before being tested again for either pregnancy or HIV status.
- N: The number of unprotected sex acts (inside and outside the ovulation window) that an uninfected women engages with an infected male partner before being tested again for either pregnancy or HIV status.

The probability of successfully delivering a child depends upon the number of condomless sex acts over the ovulation window, M. The HIV transmissibility varies based upon the number of condomless sex acts over the entire month, N. In the optimal scenario, M and N are the same.

Thus, the overall unprotected sex act base male-to-female transmissibility is given by:

$$\tilde{\alpha} = h_{PrEP} \times h_{STIs} \times h_{ART} \times h_L \times \alpha, \tag{1}$$

where each multiplicative factor is taken to be equal to 1 if the condition does not apply (e.g., $h_L = 1$ if the male not in the late stage of HIV and late stage is effectively turned "off"). Similarly, the effective MTCT transmission is:

$$\tilde{p}_{MTCT} = h_{TxMTCT} \times p_{MTCT}, \tag{2}$$

where h_{TxMTCT} is taken to be equal to 1 if the the pregnant female is not on ART during pregnancy. The probability that the female stays uninfected after N condomless sex acts is:

$$[1 - \tilde{\alpha}]^N. \tag{3}$$

Similarly, the probability that the female never becomes pregnant after M condomless sex acts is:

$$[1 - p_c(a)]^M. \tag{4}$$

3.2 The Outcome Equations

Following are the five possible outcomes relevant to the female and the associated equations:

1. Female stays HIV-uninfected, becomes pregnant, and successfully delivers a child. This occurs with probability:
$$P_{-\surd} = p_d(a)\{1 - [1 - p_c(a)]^M\}[1 - \tilde{\alpha}]^N. \tag{5}$$
2. Female becomes HIV-infected, becomes pregnant, and successfully gives birth to an HIV-uninfected child. This occurs with probability:
$$P_{+\surd-} = p_d(a)(1 - \tilde{p}_{MTCT})\{1 - [1 - p_c(a)]^M\}\{1 - [1 - \tilde{\alpha}]^N\}. \tag{6}$$
3. Female becomes HIV-infected, becomes pregnant, and successfully gives birth to an HIV-infected child. This occurs with probability:
$$P_{+\surd+} = p_d(a)\tilde{p}_{MTCT}\{1 - [1 - p_c(a)]^M\}\{1 - [1 - \tilde{\alpha}]^N\}. \tag{7}$$
4. Female stays HIV-uninfected and does not become pregnant (resulting in no child). This occurs with probability:
$$P_{-\times} = [1 - p_c(a)]^M[1 - \tilde{\alpha}]^N + [1 - p_d(a)]\{1 - [1 - p_c(a)]^M\}[1 - \tilde{\alpha}]^N, \tag{8}$$
where the first term represents the case where the female stays HIV-uninfected and does not become pregnant and the second term is the probability she stays HIV-uninfected and becomes pregnant but does not deliver a child.
5. Female becomes HIV-infected, and does not become pregnant. This occurs with probability:
$$P_{+\times} = [1 - p_c(a)]^M\{1 - [1 - \tilde{\alpha}]^N\} + [1 - p_d(a)]\{1 - [1 - p_c(a)]^M\}\{1 - [1 - \tilde{\alpha}]^N\}, \tag{9}$$
where the first term represents the case where the female becomes HIV-infected and does not become pregnant and the second term is the probability she becomes HIV-infected and becomes pregnant but loses the baby before birth.

To get the total probability of the female's HIV status or whether she gives birth to a child the probabilities can be added as follows:

- Female stays HIV-uninfected:

$$P_{HIV-} = P_{-\surd} + P_{-\times} \tag{10}$$

- Female becomes HIV-infected:

$$P_{HIV+} = P_{+\surd-} + P_{+\surd+} + P_{+\times} \tag{11}$$

- Female becomes pregnant and successfully gives birth:

$$P_{child} = P_{-\surd} + P_{+\surd-} + P_{+\surd+} \tag{12}$$

- Female does not become pregnant:

$$P_{nochild} = P_{-\times} + P_{+\times} \tag{13}$$

3.3 Annual Probability Estimate

The model equations are designed to quantify the likelihood of the outcome until the next pregnancy or HIV test. When we sample N for both the *optimal* and *suboptimal* scenarios we assume the condomless sex acts occur over a 30-day period (representative of one-month and approximately one menstrual cycle). Using the sampled N the outcome equations provide probabilities that represent monthly estimates. To obtain annual probability estimates, we evaluate the model equations over an entire sample of parameters, as described in Section 4.1, for the number of condomless sex acts that would occur over one month (i.e., using N) as well as those for two to twelve months (i.e., a year). In this case the model equations are evaluated twelve times, each with N multiplied by respective month number. Then the average probability, over all twelve months, is reported; in this instance we assume that outcomes are equally likely in the first month as month twelve and also assume the serodiscordant couple will cease condomless sex once the woman becomes pregnant or becomes HIV-infected. For the purpose of robustness, we also created annual probabilities using a decision tree approach, where the decision to engage in condomless sex for one more month is conditioned on the fact that pregnancy of HIV-infection did not occur in the previous month and each of the leaves in the trees are propagated out for twelve months; this method produced results with similar trends. The time risk of pregnancy and HIV infection is contemporaneous and a limitation of the model.

4 Experimental Design and Parameter Details

4.1 Running the Model

The model was constructed using two different software tools:

1. In the statistical software R, where we sample the parameters via a Latin Hypercube Sample (LHS) and use these to run many thousands of scenarios (each using a different sampled parameter set) and obtain a sensitivity of the outcomes to each parameter, for both the *optimal* (M = N) and *suboptimal* (M not necessarily equal to N) scenarios.

2. As an Excel tool where users can vary parameter values within given ranges and see a pie chart split by the five outcomes for a given value of N and two further plots showing how these five probabilities change with N (see Figure 1), this assumes an *optimal* setting where M = N (the couple is following clinical counseling recommendations);

The reference parameters and their ranges are specified similarly in both the Excel tool and the R code. Values for these parameters are described in section 4.3. When sampling parameters via a Latin Hypercube approach, we specify a range of values (i.e., a lower and upper bound) and a most likely reference value (the mode). Each parameter may be sampled using a different probability distribution function (pdf). We consider both a Uniform and a beta PERT distribution (see Appendix A). In the Uniform pdf, every parameter value within the range of specified values is equally likely to be selected. In the beta PERT pdf, the reference value is more likely to be sampled relative to the boundary values of the parameter range.

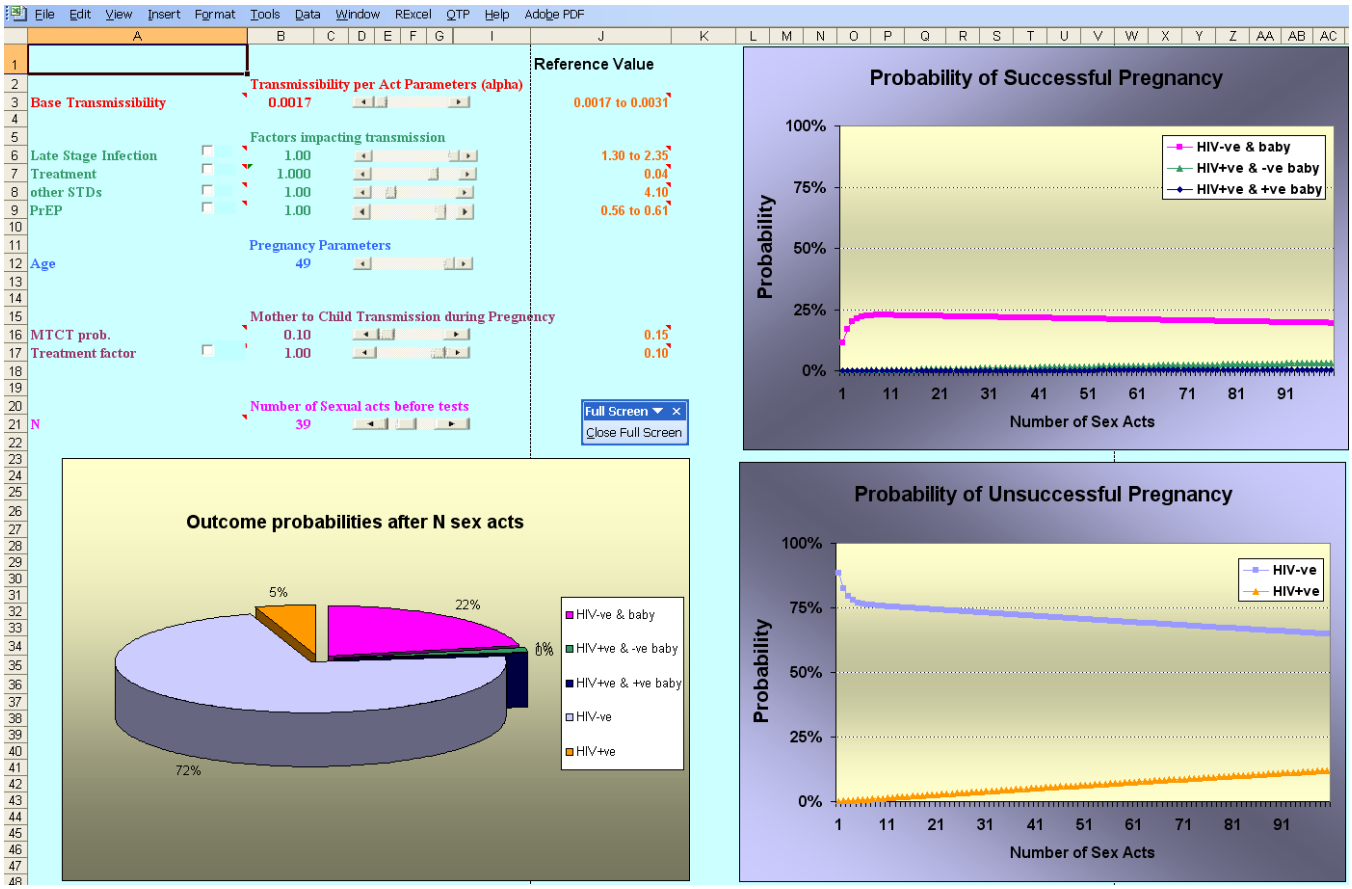

Figure 1: Snapshot of the Excel tool

4.2 Experimental Design

Our experimental design considers the following five binary variables (i.e., the states can be either ”on” or ”off”):

1. Late Stage: Whether the infected male partner is in the late stage of HIV;
2. Treatment: Whether the infected male partner is on treatment;
3. Other STIs: Whether either partner has other STIs;
4. PrEP: Whether the uninfected female partner is on PrEP;
5. ART: Whether the infected pregnant female partner is on on ART.

There are $2^5 = 32$ combinations of the various binary settings and we can consider all 32 of these combinations. In the setting where we assume all STIs are treated or under treatment, there are $2^4 = 16$ combinations. For each combination, we run $10,000$ runs by sampling the parameters values via our LHS. Therefore, in total we run $32 \times 10^4 = 320,000$ scenarios, assuming STIs are not treated; the $160,000$ scenarios, that assume STIs are treated, are a subsample of the entire binary setting permutations (i.e, only those samples where $h_{STIs} = 1$, implying this parameter is ”off”).

4.3 Ranges for Sampled Parameters

All parameters except the pregnancy probabilities are sampled over a range of possible values. Table 1 provides parameter values, their ranges, their sampling pdf’s, and the references to literature used to quantify the values.

- The value for α was obtained using values given by another modeling paper by Smith *et. al.* [3] as described in Table S2 in the supplementary material. These values were estimated by analyzing viral load data from Gray *et al.* [4] and using relationships obtained by Quinn *et.al.* [5] where viral loads are linked to transmission probabilities. The lower bound and the distributional peak for α are reduced slightly from the values obtained from Smith *et. al.* [3] to coincide with the unadjusted per-act risk of unprotected male-to-female transmission found by Gray *et.al.* [6]. Confirming this range, is a study by Hughes *et. al.* [7] based in Africa, where the reported transmissibiities lie within a similar range (0.0010 to 0.0037) when covariates such as age and STI status are unadjusted . We maintain an upper bound slightly lower than that found by Hughes *et. al.* to account for the covariates as well as the decreased infectivity found in industrialized countries in separate studies [8, 9].

- For the value of h_L, we also use the value provided in table S2 of the supplementary material in the modeling paper by Smith *et. al.* [3]. This range is confirmed within Uganda by Wawer *et. al.* who estimated the average rate of HIV transmission to be 0.0015 per coital act within 6 to 15 months (i.e., after primary infection) and 0.0028 per coital act (a 95% confidence interval of 0.0015 to 0.0041) at the late stage [10]. We are aware that by using the data from Wawer *et. al* and Hollingsworth *et. al.* estimates in the late stage transmission are seven times more infectious than the asymptomatic stage [11].

- The value of h_{ART} was estimated using a finding of a 96% overall reduction in HIV transmission in discordant heterosexual couples randomized to early HIV treatment [12]; This study provides strong evidence for the dramatic effectiveness of antiretroviral therapy in reducing HIV infectiousness.

- The multiplicative impact of a sexually transmitted infections (STIs) in either partner varies greatly based upon the study and this large uncertainty corresponds with the large range; Fleming and Wasserheit [13] show the effect of STI(s) on base transmission from numerous studies, covering four continents, to range between from 2 to 23, with a larger proportion at the lower-end and variation based upon the type of STI (i.e., ulcerative and non-ulcerative) [13]. The Fleming study incorporates primary studies like the randomized control trial of Grosskurth *et. al.* [14] who show that in Tanzania sharp reductions in HIV incidence are associated with STI program intervention. Based upon the concentration of studies at the low-end by Fleming *et. al.* [13] and the results from Gray *et. al.* [6] that show transmission increase of 2 to 2.7 with STIs, we set our peak at 2.3.

- The value for h_{PrEP} is based on a recent (2012) study by Jared M. Beaten *et. al.* demonstrating a reduction in HIV transmission of about 66% (efficacy range of 28% to 84%) for the female population using oral teofovir disporxil fumrate coformulated with emtricitabine (TDF-FTC) - Note: $h_{PrEP} = 1 - \epsilon_{PrEP}$, where ϵ_{PrEP} represents efficacy. These data are supported by similar results, yet a larger range (0.125 to 0.846), from additional studies that are less recent; the first by Karim *et. al.* [15] showed that Tenofovir Gel reduced transmission to 39% when used as an antiviral microbicide for the prevention of HIV infection in women; the second is the IPREX study by Myers *et. al.* where PrEP was shown to have efficacy ranges from 15.4% to 87.5% with peak at 43.8% [16].

- The value of p_{MTCT} determined using a study by Conner *et. al.* [17], a study by Cock *et. al.* [18] on the prevention of Mother-to-Child HIV transmission, and a more recent publication by Cooper *et al.* using various antiretroviral strategies [19]. The proportion of children infected at 18 months on treatment (zidovudine) is 25.5 percent (95 percent confidence intervals of 18.4 to 32.5) as compared to the placebo group [17]. The summarized ranges for resource-constrained countries are similar at 15 to 30 percent [18] and Cooper *et al.* find the probability of transmission to be 20% [19]. The reduction factor associated with treatment for a mother with HIV h_{TxMTCT} was found based on the same study by Conner *et. al.* [17] and on the study by Zutlevics *et. al.* [20] - A 67.5% (95 percent confidence intervals 40.7% to 82.1%) relative reduction in the risk of HIV transmission from treatment corresponds to a multiplicative value of 32.5% (1 - 0.675). The total risk of a mother with HIV transmitting to a child while on treatment is p_{MTCT} * h_{TxMTCT} and Zutlevics *et al.* find this to be 1 to 2% with elective cesarean and Cooper *et al* find 3.8% with dual antiretroviral therapy and 1.2% with highly active antiretroviral therapy. The multiplied modes we select result in a slightly higher total risk of transmission while on treatment than these two numbers, however, we believe this represents a conservative case where non-cesarean and breastfeeding might not be controlled. Furthermore, our results

indicate how unlikely this outcome is on an annual basis and using an even lower range for h_{TxMTCT} would further minimize this outcome.

Parameter	Value	Lower Bound	Upper Bound	Distribution	Source
α	0.0022	0.0010	0.0031	PERT	[3, 5, 6]
h_L	1.82	1.29	2.35	Uniform	[3, 10, 11]
h_{ART}	0.04	0.01	0.27	PERT	[12]
h_{STIs}	2.3	2.0	23.0	PERT	[6, 13, 14]
h_{PrEP}	0.34	0.16	0.72	PERT	[15, 16, 21]
p_{MTCT}	0.255	0.184	0.325	Uniform	[17–19]
h_{TxMTCT}	0.325	0.179	0.593	Uniform	[17, 19, 20]
$N_{optimal}$	3	1	12	PERT	
$N_{suboptimal}$	15	1	60	PERT	

Table 1: Parameter values, ranges, distribution, and sources

For both the optimal and suboptimal case, the parameter M is derived from N (i.e., for the *optimal* scenario N = $N_{optimal}$ and for the *suboptimal* scenario N = $N_{suboptimal}$). As mentioned in Section 3.1, M and N are equal in the *optimal* scenario. As we assume the condomless sex acts (N) are uniformly distributed over a 30-day period and the ovulation window occurs over 3-days, the ovulation window represents a tenth of the month and so for the *suboptimal* scenario $M = (1/10) * N_{suboptimal}$. Ranges of N for both scenarios were obtained from clinical practitioner best estimates.

Based upon a clinical practitioner knowledge, the number of monthly condomless sex acts during the ovulation window, M, is highly variable. This knowledge is supported by Hughes *et. al.* where they find the median number of monthly sex acts to be four [7], however this number does not differentiate between those trying to conceive; it is reasonable that the number of acts will increase slightly when trying to conceive. It is important to note, that each month the female can only conceive during the ovulation window and so we assume that this window is known.

4.4 Age-Based Pregnancy Values

Similar to Table 1, Table 2 provides the values of $p_c(a)$ and $p_d(a)$ and how these parameters change with age, a, for the female. These two probabilities are provided by Van Noord-Zaastra *et. al* [22] in terms of the number of cycles before pregnancy (Table 1, used to obtain $p_c(a)$) and the result of the pregnancy (Table 2, used to obtain $p_d(a)$). The percent of women who became pregnant after 12-cycles (equating to approximately one year) are provided for the following age ranges: 18-24, 25-29, 30-34, and 35-42 years. Conception declines with age and to obtain conception probabilities for each age for 18 to 49 years we interpolate the percents using a simple non-linear model using the mean age of each group, A, as an independent variable, and the annual probability of conception as the dependent variable. We use the outcome of the pregnancy by age-groups as the probability of delivery, $p_d(a)$ and assume this is the same on the per act basis. In order for the units to match in the model equations, $p_c(a)$ and $p_d(a)$ must be expressed on a per act basis. We use the following relationships (and model relating P_C and A) to obtain the per act probability of conceiving and delivering a baby.

$$\hat{P_C} = 0.2067 * (50 - A)^{-0.4092} \tag{14}$$

We take one minus the annual probability of conception, $\hat{P_C}$, to obtain the annual probability of not conceiving, P_{NC}:

$$P_{NC} = 1 - \hat{P_C}. \tag{15}$$

The monthly probability of not conceiving, $p_{nc}(a)$, is independent each month and equal to the annual probability of not conceiving as follows:

$$P_{NC} = p_{nc}(a)^{12}. \tag{16}$$

As such, the monthly probability of conceiving over all sexual acts during month's fertility window is:

$$p_c(a)^* = 1 - p_{nc}(a) = 1 - {P_{NC}}^{1/12}. \tag{17}$$

To obtain the probability of conception per sex act, the monthly probability of conceiving, $p_c(a)$ is divided by the number of sex acts occurring during the ovulation window, M:

$$p_c(a) = \frac{p_c(a)^*}{M} \tag{18}$$

Each month the female can only conceive over the ovulation window and so in calculating $p_c(a)$ we assume that these three days are known and all condomless sex acts (M) occur within this three-day window. It is important to note that the M is not the same as N; the former represents the number or acts over the female's monthly fertility window where as N represents all sex acts engaged in before the next HIV or pregnancy test (these acts can occur outside the three-day fertility window).

age: a	**Interpolate Annual Conception:** $\hat{P_C}$	**Interpolated Monthly Conception:** $p_c(a)^*$	**Delivery:** $p_d(a)$
18	0.85	0.793	0.89
19	0.84	0.726	0.89
20	0.83	0.676	0.89
21	0.82	0.635	0.89
22	0.81	0.601	0.89
23	0.80	0.570	0.89
24	0.78	0.542	0.89
25	0.77	0.516	0.89
26	0.76	0.492	0.89
27	0.75	0.470	0.89
28	0.73	0.449	0.89
29	0.72	0.429	0.89
30	0.70	0.410	0.86
31	0.69	0.391	0.82
32	0.67	0.374	0.79
33	0.66	0.357	0.75
34	0.64	0.341	0.72
35	0.63	0.326	0.68
36	0.61	0.310	0.65
37	0.59	0.296	0.61
38	0.57	0.282	0.58
39	0.55	0.268	0.54
40	0.53	0.254	0.51
41	0.51	0.241	0.47
42	0.48	0.228	0.44
43	0.46	0.216	0.40
44	0.43	0.204	0.37
45	0.40	0.192	0.38
46	0.36	0.180	0.41
47	0.32	0.169	0.43
48	0.27	0.157	0.45
49	0.21	0.146	0.48

Table 2: Interpolated monthly probability of becoming pregnant, $p_c(a)^*$, and of giving birth if pregnant, $p_d(a)$, as a function of age, a. The per act probability of conception, $p_c(a)$, and of conceiving plus delivering, $p_d(a) * p_c(a)$, use the number of sampled condomless sex acts, M derived from N.

5 Model Results

5.1 Descriptive Statisics, Percent of Model Runs Within Each Outcome Range

Table 3 displays the percentage of simulations for each outcome that fall within the given annual probability range (this is for the 160,000 model runs that assume STIs are treated). For example, in the suboptimal scenario, of the 160,000 model runs 26 percent result of the outcomes where the female is HIV-uninfected and has a successful pregnancy fall within a probability of 0 to 10 percent and 44 percent result in a probability of 40 to 60 percent for the same outcome.

Outcome	Very Unlikely (0 to 0.1)	Unlikely (0.1 to 0.4)	Possible (0.4 to 0.6)	Likely (0.6 to 1.0)
Female HIV-, Pregnancy Success, Child HIV-	S=26%, O=17%	S=28%, O=26%	S=44%, O=54%	S=2%, O=3%
Female HIV+, Pregnancy Success, Child HIV-	S=91%, O=100%	S=9%, O=0%	–	–
Female HIV+, Pregnancy Success, Child HIV+	S=100%, O=100%	S=0%, O=0%	–	–
Female HIV-, Pregnancy Unsuccessful	–	S=0%, O=0%	S = 22%, O=9%	S=77%, O=91%
Female HIV+, Pregnancy Unsuccessful	S=63%, O=94%	S = 20%, O=6%	S=14%, O=0%	S=3%, O=0%

Table 3: S = suboptimal, O = optimal. Percentage of model runs by outcome (all ages).

An example of the parameter sampling distribution is below for the number of condomless sex acts in a month for the optimal and suboptimal scenarios:

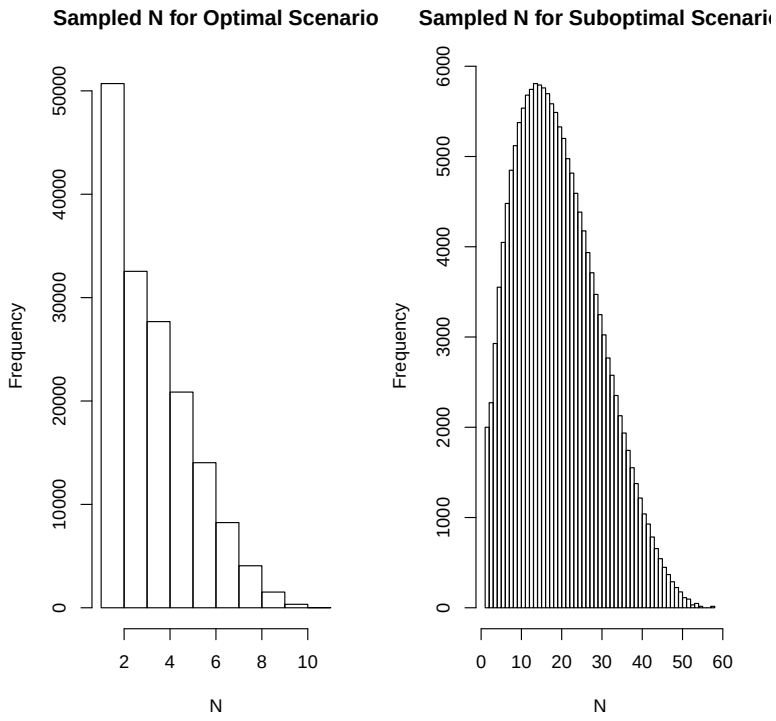

Figure 2: Distribution of sampled number of condomless sex acts

For each intervention of interest and scenario, we show the average annual probabilities combined by outcome in Figure 3. We see across all situations the unsuccessful outcome is most likely and the outcomes where the female is HIV+ are more likely to occur in the suboptimal scenario.

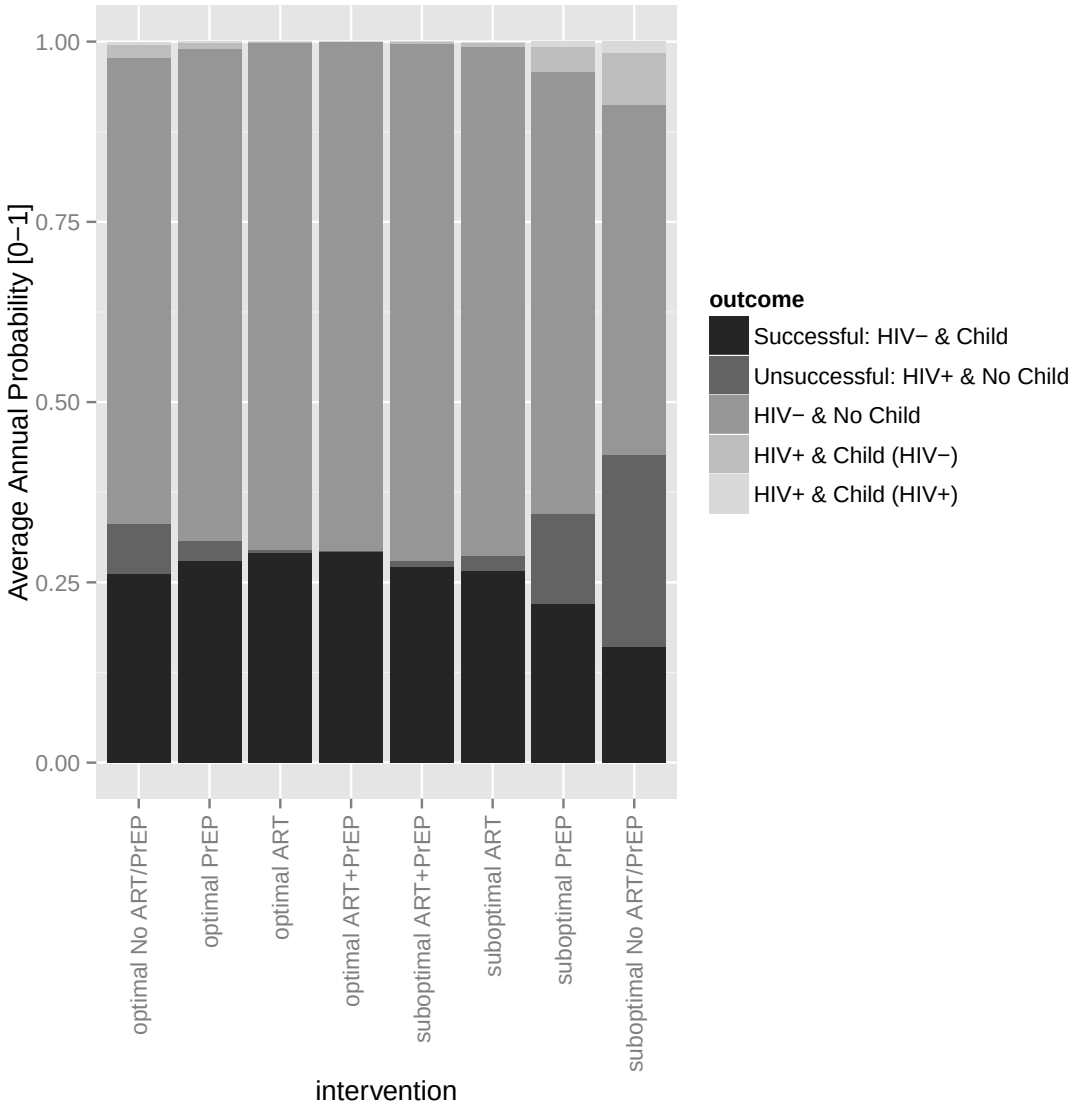

Figure 3: Average annual outcomes by intervention and scenario

The average annual probability across all samples for all ages and for those 30 years and younger are depicted in Figures 4 and 5, respectively. Figures 4 and 5 show the three most likely outcomes at an aggregate level. The timing and low number of condomless sex acts in the optimal scenario keeps the probability of becoming HIV-infected very low, while in the suboptimal scenarios the chance of the female becoming HIV-infected increases dramatically. At younger ages, we see this outcome offset by an increased chance of becoming pregnant and delivering a child. In both scenarios, ART produces the largest reduction in HIV-infection.

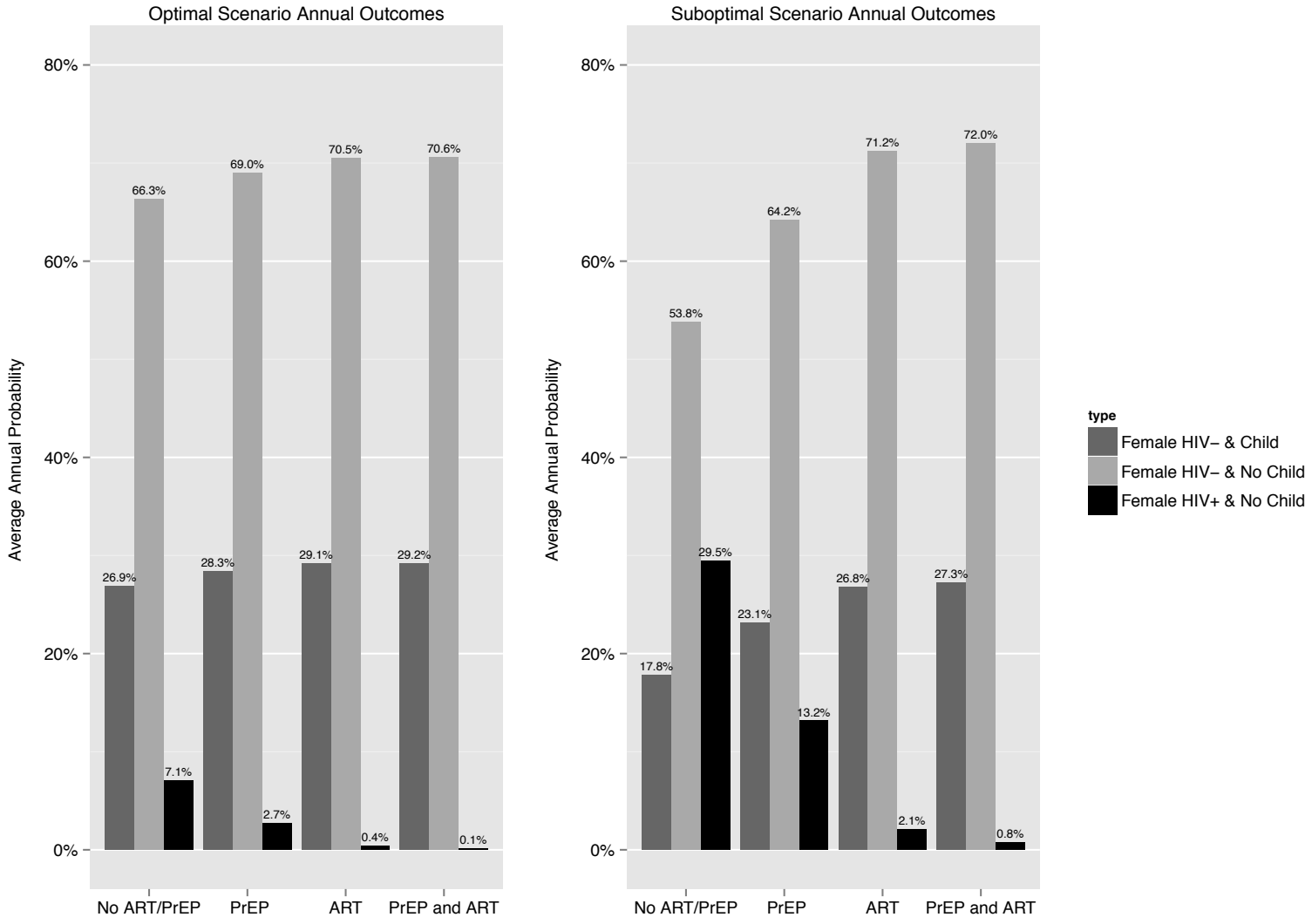

Figure 4: Aggregated annual results across all ages

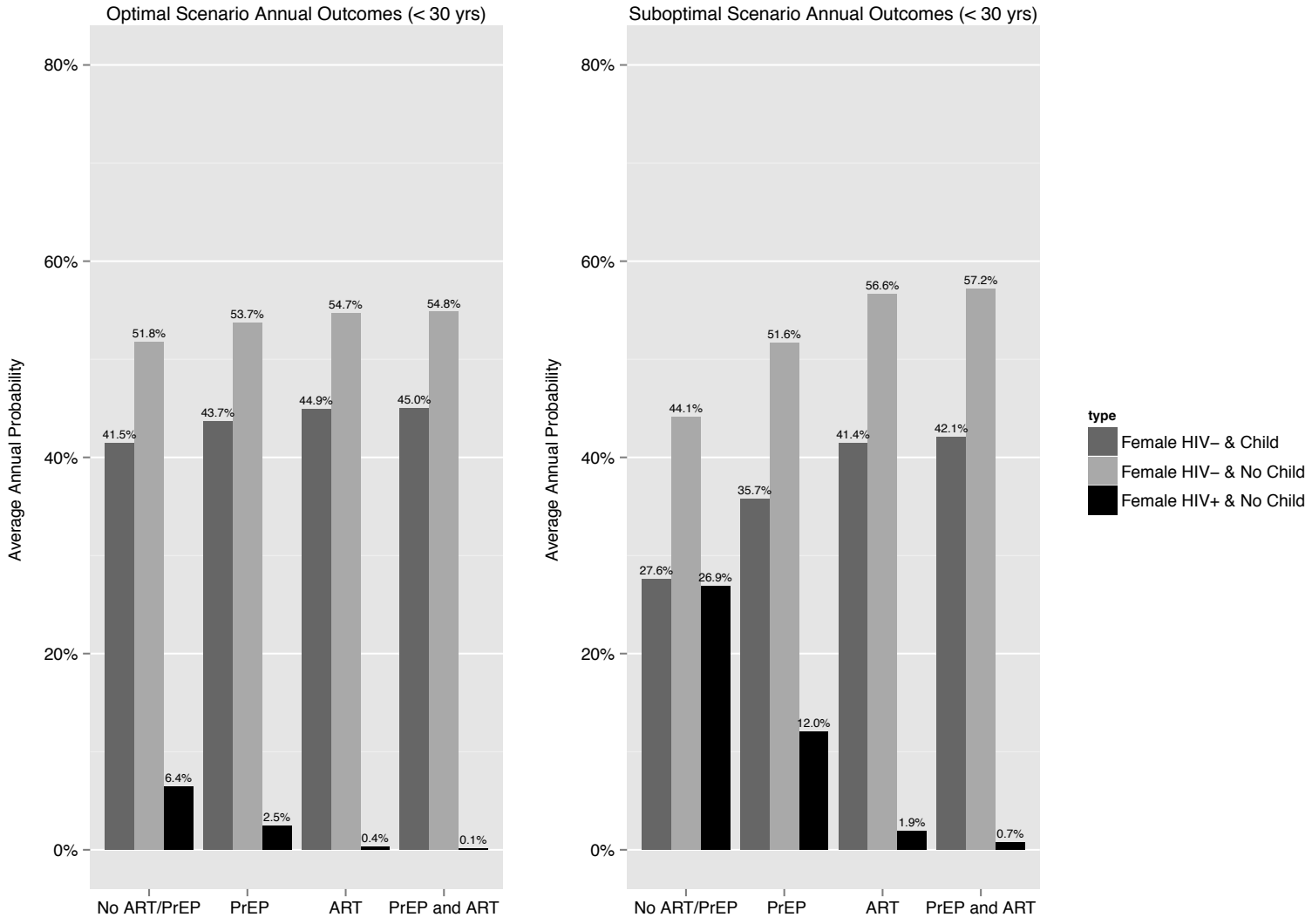

Figure 5: Aggregated annual results for 30 years or younger

We know that age is an important variable for pregnancy and less influential in HIV transmission and this is depicted in Figure 6. We see that conception is not affected across all interventions conception, but the interventions do affect HIV transmission. Conception is influenced by age and the number of condomless sex acts.

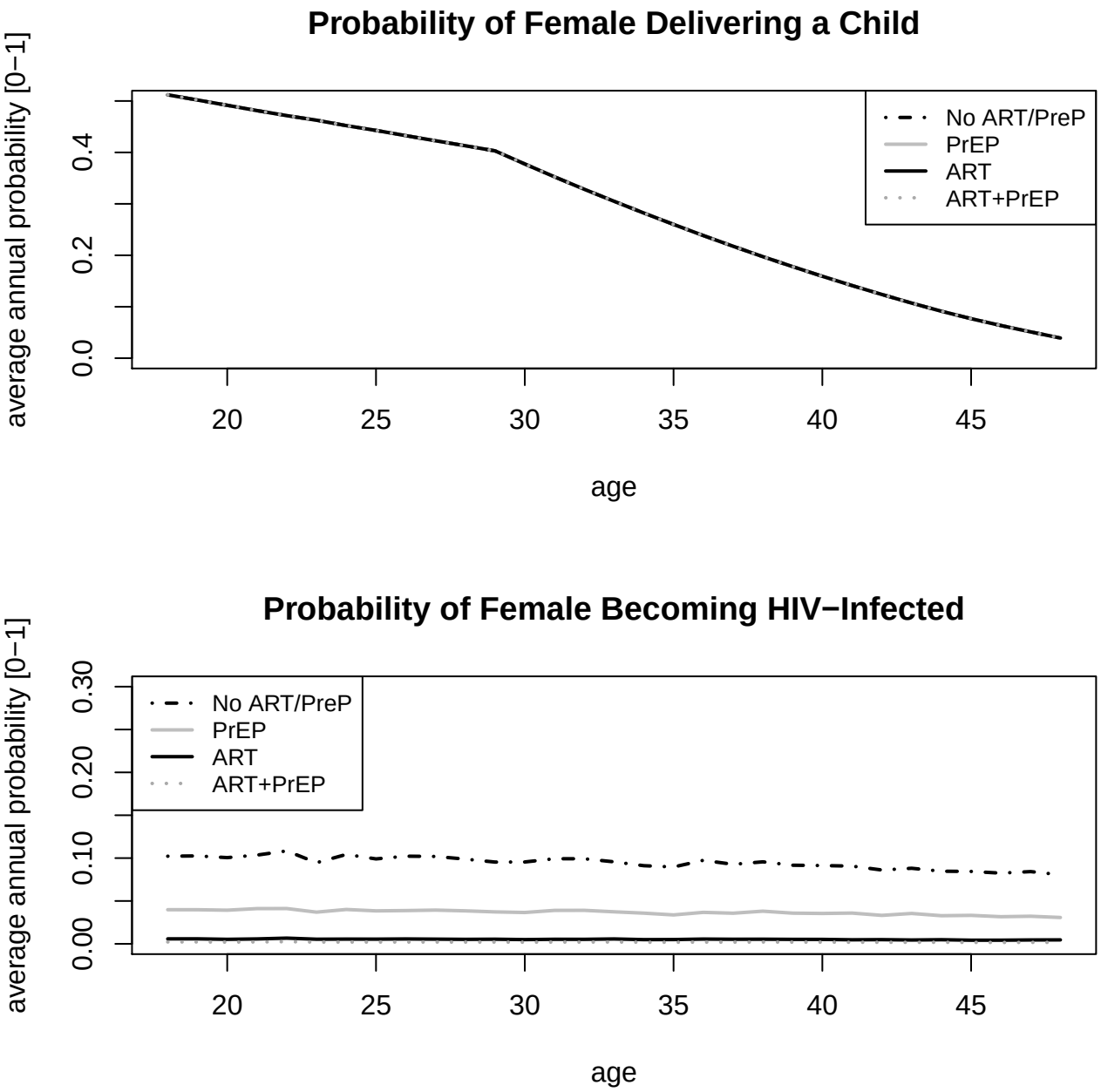

Figure 6: Average annual probabilities by intervention for the optimal scenario

5.2 Impact of Variable Ranges, Uncertainty Analysis

We previously note that age is an important factor in conception and this simple model confirms this fact as can be seen in Figure 7. We vary only age in the model and assume all other levers (i.e., binary variables that can be in an on or off stage, such as the male on ART and the women using PrEP) are all off and all other variables are set to the reference or median values (i.e., transmissibility to 0.0022, the number of unprotected sex acts N to 3, and mother to child transmissibility to 0.255). Under these conditions, the woman has the greatest chance of remaining HIV-uninfected and having a successful pregnancy when she is younger. However, the probability of this outcome steadily declines after age 30.

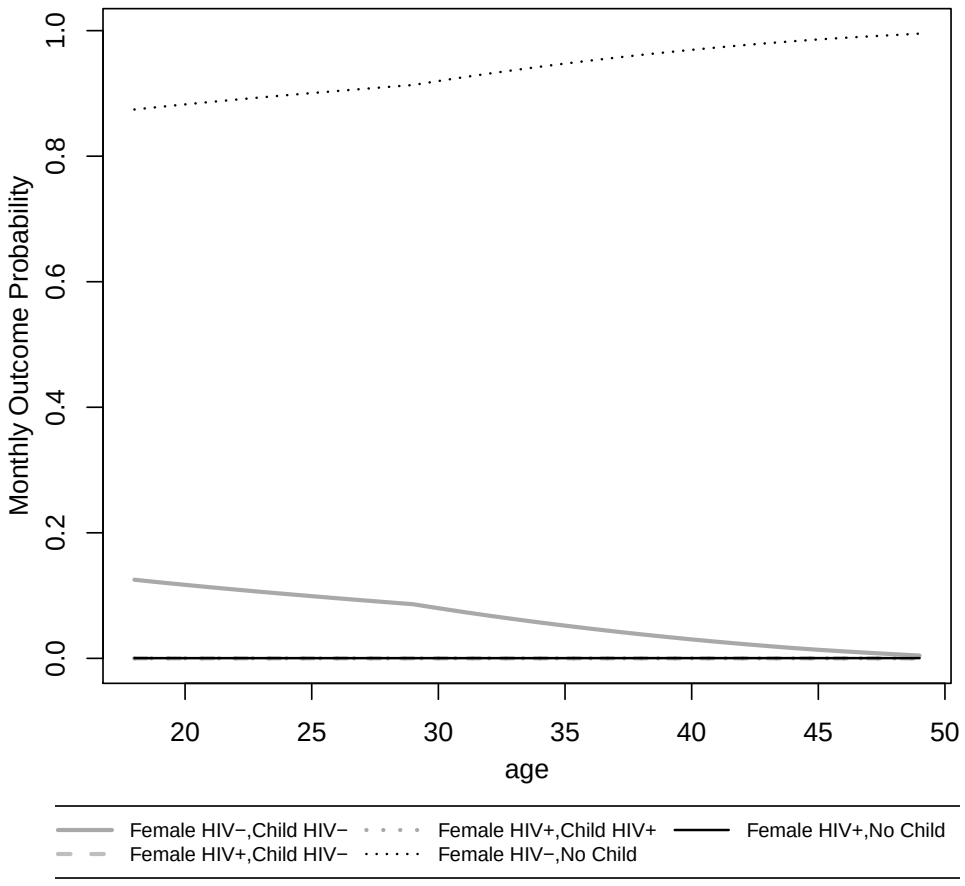

Figure 7: Outcomes by varying age in the optimal scenario, using monthly probabilities

While we know that age influences the chances of conception and delivery, the HIV status of both the mother and child are not dependent upon age as much as the number of condomless sexual acts and the interventions. To better understand how the range of each variable affects the probability of an outcome, we use only one binary variable at a time and then vary that parameter over its entire range. We use the reference values for the transmissibility (α) and mother to child transmission (p_{MTCT}). We examine the outcomes for ages 20 and 35 years in the optimal scenario for a base number of sex acts. In these two instances, age affects both the probability of delivery and probability of conception. This results in the most likely outcome being where the female remains HIV-uninfected and without a child (mostly due to the low number of condomless sexual acts). At both ages, when the multiplicative factor for STIs (h_{STIs}) is very high the transmissibility is multiplied by a rather large factor causing the female to become HIV positive. Interestingly, as the female ages the most likely outcome is an HIV positive status at lower and lower values of the STI factor. Another observation can be seen at age 35 where for all parameters the probability of the successful outcome (Female HIV-, Child HIV-) is significantly reduced as compared to the results at age 20.

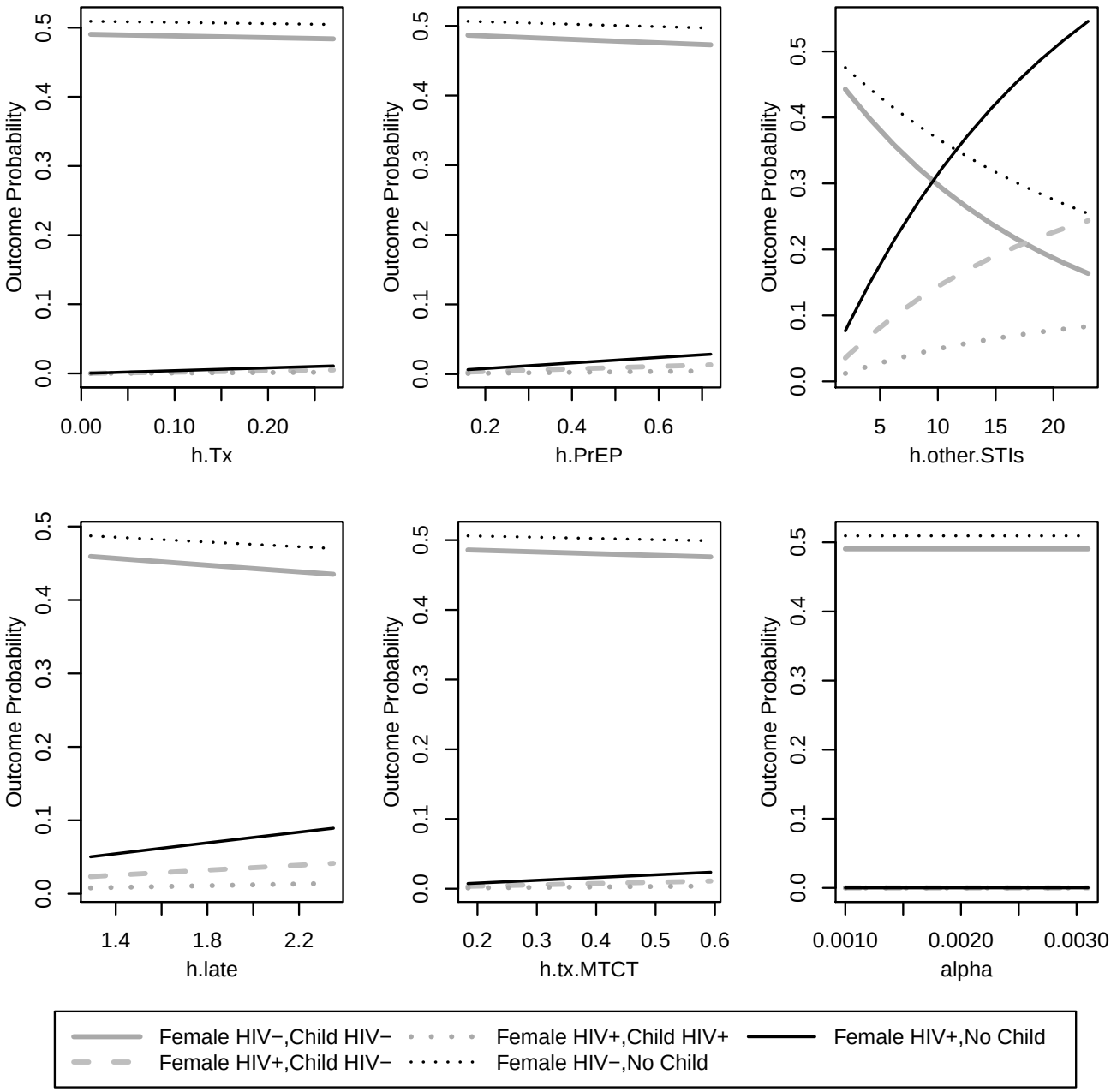

Figure 8: Outcomes by varying parameters at age 20, using monthly probabilities

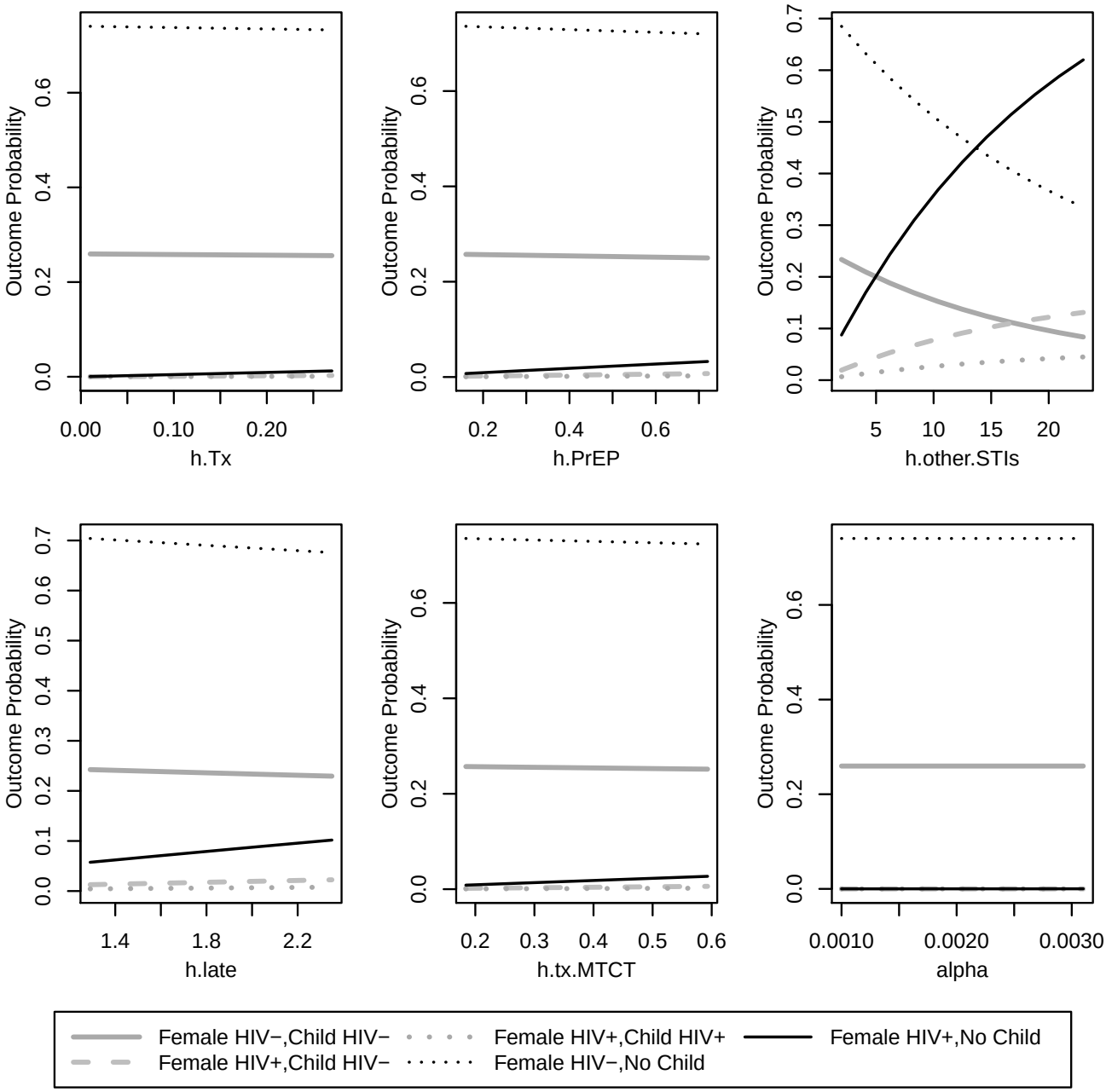

Figure 9: Outcomes by varying parameters at age 35, using monthly probabilities

We know that the magnitude of the STI factor is the largest of the base transmissibility multiplicative factors and has the widest range of all the variables used to compute the overall transmissibility. To demonstrate the relative impact of the range on the predicted probabilities we regress the multiplicative factors on each outcome probability. The multiplicative factor is set to zero when it is "turned off" in a parameter set. We control for age, number of sex acts, base transmissibility, and mother to child transmission (when there is a successful pregnancy). Then to obtain the impact of each variable range we set all other variables to their reference values or median values and obtain the predicted probability for each outcome for each variable's minimum and maximum value. The values displayed in Table 4 are the absolute differences in the predicted probabilities using the minimum and maximum values. For example, for the outcome where a female is HIV negative and has a successful pregnancy, the predicted difference between using the minimum and maximum value for h_{STIs} results in about a 22% change where as the male being in the later stages of HIV is only about a 2.6% difference, controlling for all other variables. What we see is that the uncertainty in the h_{STIs} variable can have the greatest impact on the

outcome in all five instances.

	Female HIV-, Pregnancy Success, Child HIV-	Female HIV+, Pregnancy Success, Child HIV-	Female HIV+, Pregnancy Success, Child HIV+	Female HIV-, Pregnancy Unsuccessful	Female HIV+, Pregnancy Unsuccessful
Other STIs	22.40%	18.45%	3.94%	31.87%	31.88%
Treatment	18.39%	15.04%	3.22%	24.76%	24.90%
Late Stage	2.63%	2.14%	0.46%	3.78%	3.82%
PrEP	2.40%	2.04%	0.44%	3.62%	3.54%

Table 4: Absolute change in predicted probabilities from the minimum and maximum parameter ranges in the optimal scenario, using monthly probabilities

5.3 Influential Variables

While the previous section explores the impact of each parameter's range, this section identifies influential variables; that is regardless of the value, which binary options have the largest impact on the outcome probabilities (e.g., does having other STIs or the male being on treatment change the probability of an outcome more or less?). We employ linear regression models for each of the five outcomes. First we regress the binary variables (whether the male is in the late stage of HIV, if the male is on treatment, if either partner has STIs, if the female is taking PrEP, and on HAART during her pregnancy) on the outcome probability. The mother on HAART is only applicable in the outcomes where she is HIV positive and has a child. We control for age, transmissibility, the number of sex acts and the mother to child transmissibility (in cases where she is HIV positive and has a child). In all five of the outcomes the coefficient for treatment is largest, meaning when a parameter combination has the male on treatment the largest change in an outcome is observed. Table 5 shows the relative strength of each option; the coefficients from the regressions are divided by the male on treatment coefficient to obtain the relative strength in changing the probability of the particular outcome.

	Female HIV-, Pregnancy Success, Child HIV-	Female HIV+, Pregnancy Success, Child HIV-	Female HIV+, Pregnancy Success, Child HIV+	Female HIV-, Pregnancy Unsuccessful	Female HIV+, Pregnancy Unsuccessful
ART	100.00%	100.00%	100.00%	100.00%	100.00%
Other STIs	67.82%	67.87%	67.63%	69.00%	69.00%
PrEP	25.13%	25.06%	24.86%	25.87%	25.87%
Late Stage	20.84%	20.72%	20.81%	22.16%	22.16%
HAART	-	11.17%	52.02%	-	-

Table 5: Relative importance of binary parameters in the optimal scenario, using monthly probabilities

We have noted that the range for the STI multiplicative factor affects the outcomes the most, yet if all variables have the same range what would be the impact? To answer this question we convert all the multiplicative variable ranges to a zero-to-one scale (if the option is "turned off" we set the value to zero). To understand the relative impact of each of these variables, we regress the parameters, converted to the zero-to-one, on the outcomes. Again controlling for age, transmissibility, mother to child transmission, and sex acts. What are reported in Table 6 are the regression coefficients which can be interpreted as the effect of a one percent increase in the value of the multiplicative factor on the outcome probability. While STIs still impact those with successful pregnancies who are HIV-positive the most, treatment is more influential for HIV-negative females who have successful pregnancies.

	Female HIV-, Pregnancy Success, Child HIV-	Female HIV+, Pregnancy Success, Child HIV-	Female HIV+, Pregnancy Success, Child HIV+	Female HIV-, Pregnancy Unsuccessful	Female HIV+, Pregnancy Unsuccessful
PrEP	0.0262	-0.0278	-0.0049	0.0410	-0.0396
ART	0.1682	-0.1373	-0.0294	0.2259	-0.2274
Other STIs	-0.0240	0.1979	0.0422	-0.3437	0.3438
Late Stage	-0.0260	0.0211	0.0045	-0.0372	0.0375

Table 6: Relative importance of one percent increase of parameter on monthly probability of outcomes in the optimal scenario

5.4 Regression Trees and Random Forests

In the previous section, we identified which of the binary variables are of more or less importance in altering the probability within an outcome. To gain a broader perspective that is still inline with the results from the previous sections, we use regression trees that recursively dichotomize variables to produce predicted outcomes based upon interacting variables. While trees can sometimes be inaccurate and overfit for predictive models, we are using them to understand variable interactions and influence. The successful outcome in the optimal scenario Figure 10 shows age driving the outcome; the top split to the left ("$Age > 33$") includes ages 34 and older and the right ("$Age < 34$") includes ages 33 and younger. In the suboptimal scenario, age is still the primary driver but condomless sex acts and being on ART or PrEP exert influence as well. On the contrary, for the unsuccessful outcome, for both the optimal (Figure 12) and the suboptimal (Figure 13), the male being on ART and the female PrEP play primary roles in determining whether the woman becomes HIV-infected. These trees are interpreted as follows: in Figure 12 when the male is on ART the woman has a very low chance of becoming HIV-infected (0%); on the right side, when the male is not on ART, the woman not on PrEP, the couple engage in more than five condomless sex acts over a month, and the male is not in late stage HIV there is about a 7% chance the woman will become HIV-infected and not have a child.

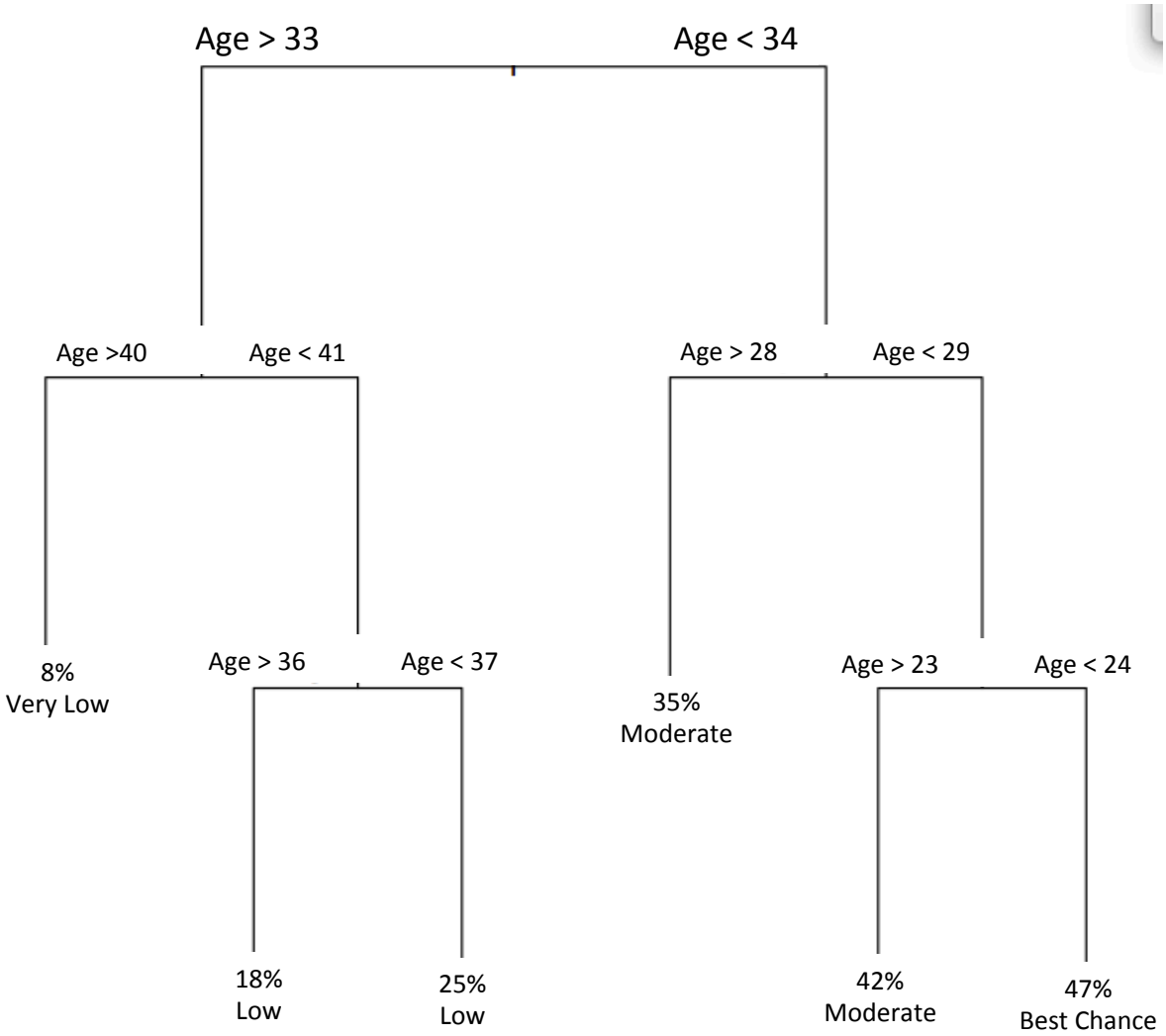

Figure 10: Regression tree for optimal scenario, successful outcome (Female HIV- and Child)

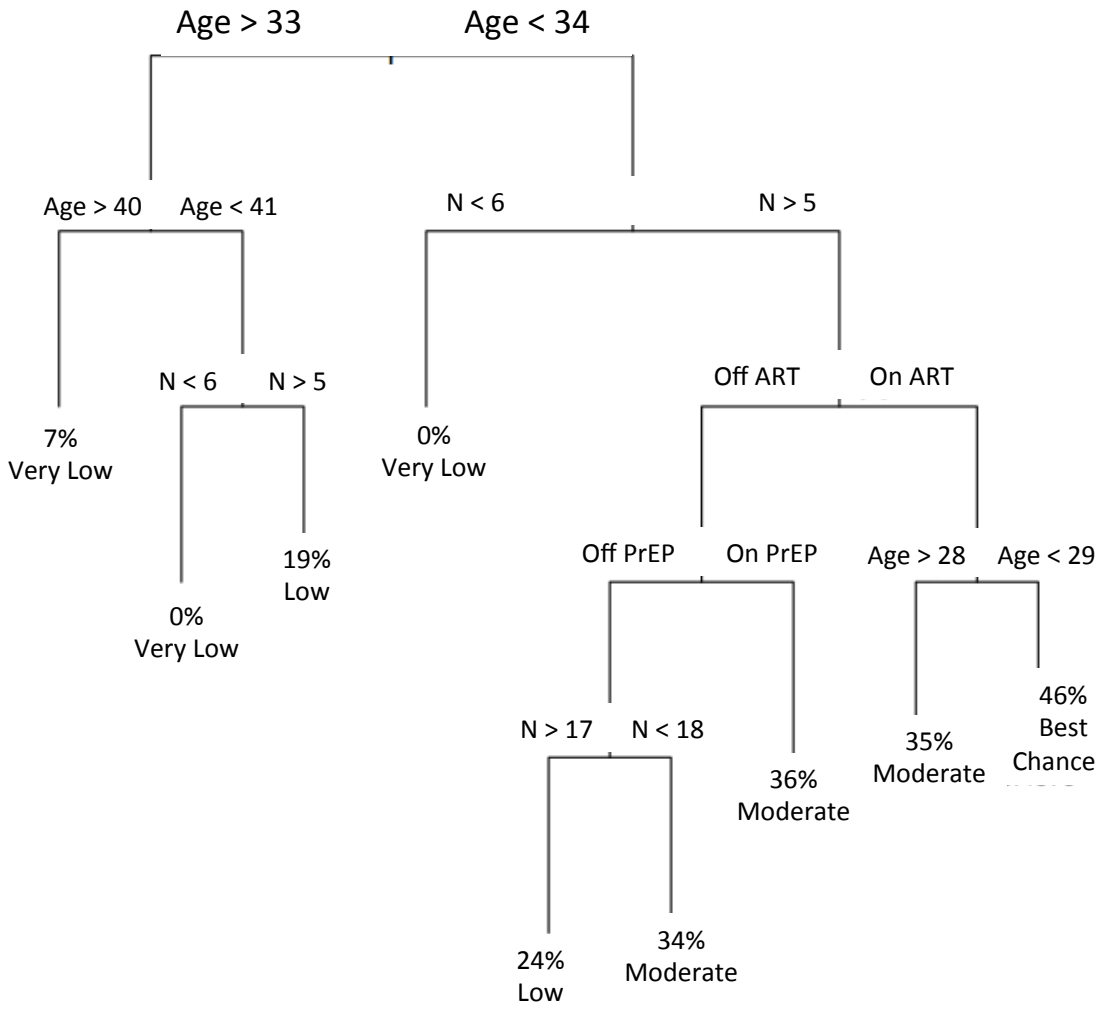

Figure 11: Regression tree for suboptimal scenario, successful outcome (Female HIV- and Child)

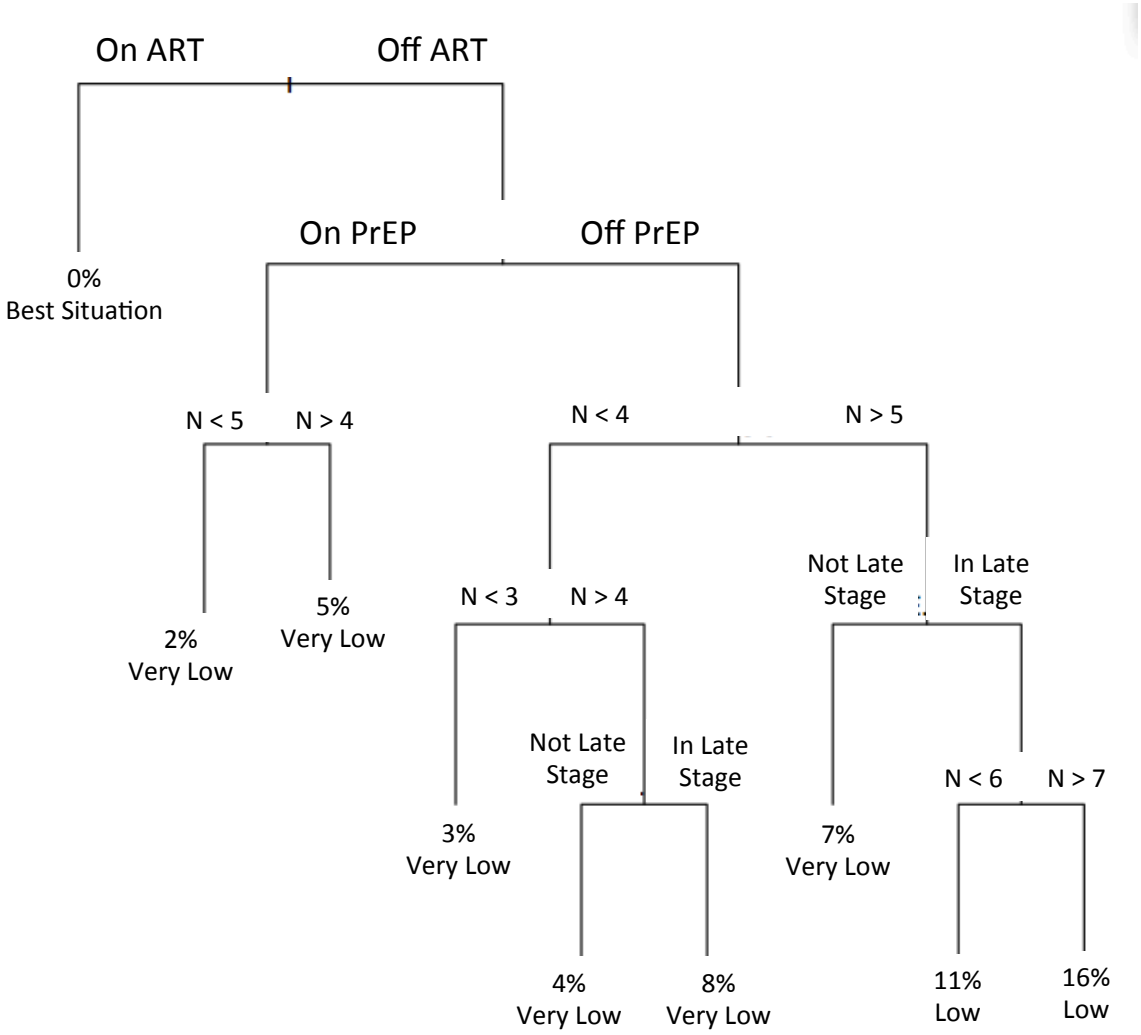

Figure 12: Regression tree for optimal scenario, unsuccessful outcome (Female HIV-infected and no child)

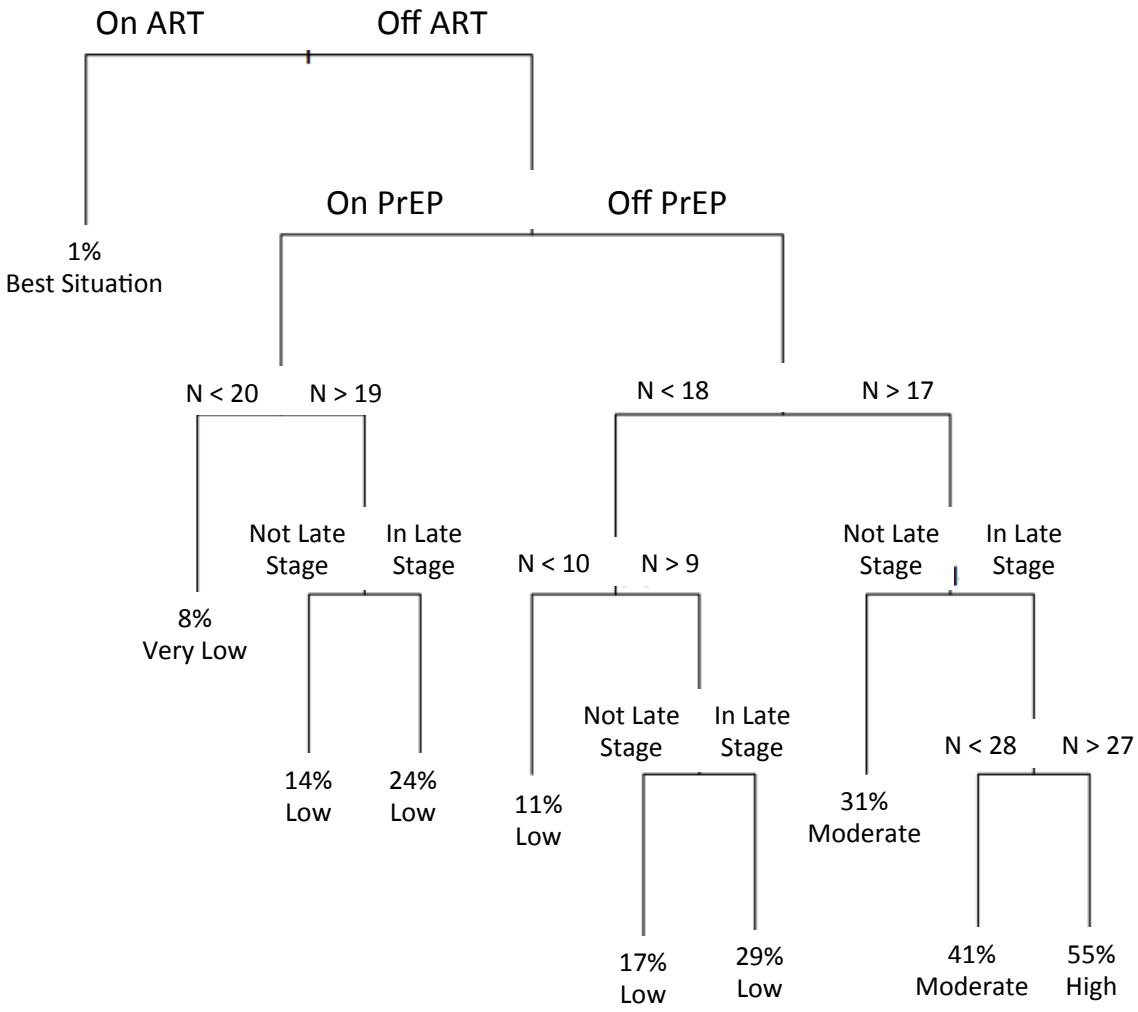

Figure 13: Regression tree for suboptimal scenario, unsuccessful outcome (Female HIV-infected and no child)

Additionally, we examine random forests (an ensemble method that constructs many trees) to determine the ranking of importance for the variables. In this analysis, we exclude the STI's (assume they are under treatment). For the successful outcome (female remains HIV-uninfected and has a child), both the optimal and suboptimal scenarios are primarily influenced by the female's age, yet when the male is using ART the node impurity is reduced about three-fold as compared to just using PrEP. In the unsuccessful outcome (the female becomes HIV-infected and does not have a child), all interventions (ART, PrEP, and Late-stage HIV status) are more influential than age; in this outcome there is a little more than a three-fold increase in node purity from ART as compared to PrEP for the optimal scenario, and in the suboptimal scenario there is a four-fold increase. This indicates ART is the most influential variable for this outcome.

A The beta PERT distribution

We run many independent simulations using different combinations of the sampled parameters. For each parameter, we specify a probability distribution that we assume when sampling within its uncertainty range of values. We assume two different probability distributions: uniform and beta PERT (Program Evaluation and Review Technique). For the case that we have large uncertainty in the value of the parameter and any of the values specified in the uncertainty range seems equally likely, we use a uniform distribution. Thus, in this case the specification of a most likely value within this range plays no role in the sampling. A beta PERT distribution instead is used as a continuous approximation to what is normally used, namely a triangular distribution between the minimum and maximum of the uncertainty range and peaking at its most likely value. Typically, sampling from the beta distribution requires minimum and maximum values (x_{min} and x_{max}) and two shape parameters, v and w. The beta PERT distribution uses the mode or most likely parameter (x_{mode}) to generate the shape parameters v and w of a beta distribution. An additional scale parameter λ scales the height of the distribution; the default value for this parameter is 4. In the PERT distribution, the mean μ is calculated

$$\mu = \frac{x_{min} + x_{max} + \lambda x_{mode}}{\lambda + 2}, \tag{19}$$

and is used to calculate the v and w shape parameters

$$v = \frac{(\mu - x_{min})(2x_{mode} - x_{min} - x_{max})}{(x_{mode} - \mu)(x_{max} - x_{min}}, \tag{20}$$

$$w = \frac{(x_{max} - \mu)(2x_{mode} - x_{min} - x_{max})}{(x_{mode} - \mu)(x_{max} - x_{min}}. \tag{21}$$

Literature Cited

[1] J.B. Stanford, G.L White, and H. Hatasaka. Timing intercourse to achieve pregnancy: current evidence. *Obstet Gynecol*, 2002.

[2] A. J. Wilcox, C. R. Weinberg, and D. D. Baird. Timing of sexual intercourse in relation to ovulation. effects on the probability of conception, survival of the pregnancy, and sex of the baby. *N Engl J Med*, 333(23):1517–1521, Dec 1995. doi: 10.1056/NEJM199512073332301. URL http://dx.doi.org/10.1056/NEJM199512073332301.

[3] R.J. Smith, J.T. Okano, J.S. Kahn, E.N. Bodine, and S. Blower. Evolutionary dynamics of complex networks of hiv drug-resistant strains: the case of san francisco. *Science*, 327(5966):697–701, Feb 2010. doi: 10.1126/science.1180556. URL http://dx.doi.org/10.1126/science.1180556.

[4] R. H. Gray, N. Kiwanuka, T. C. Quinn, N. K. Sewankambo, D. Serwadda, F. W. Mangen, T. Lutalo, F. Nalugoda, R. Kelly, M. Meehan, M. Z. Chen, C. Li, and M. J. Wawer. Male circumcision and hiv acquisition and transmission: cohort studies in rakai, uganda. rakai project team. *AIDS*, 14(15):2371–2381, Oct 2000.

[5] T. C. Quinn, M. J. Wawer, N. Sewankambo, D. Serwadda, C. Li, F. Wabwire-Mangen, M. O. Meehan, T. Lutalo, and R. H. Gray. Viral load and heterosexual transmission of human immunodeficiency virus type 1. rakai project study group. *N Engl J Med*, 342(13):921–929, Mar 2000.

[6] R.H. Gray and M.J. Wawer. Probability of heterosexual hiv-1 transmission per coital act in sub-saharan africa. *J Infectious Disease*, 205:351–352, Feb 2012. URL http://www.natap.org/2012/HIV/011312_02.htm.

[7] J.P. Hughes, J.M. Baeten, J.R. Ligappa, A.S. Magaret, A Wald, G deBruyn, J Kiarie, M Inambao adn W Kilembe, C Farquhar, C Celum, and the Partners in Prevention HSV/HIV Transmission Study Team. Determinants of per-coital-act hiv-1 infectivity among african hiv-1–serodiscordant couples. *J Infectious Diseases*, 205(3):358–365, Feb 2012.

[8] 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(2):118–129, Feb 2009. doi: 10.1016/S1473-3099(09)70021-0. URL http://dx.doi.org/10.1016/S1473-3099(09)70021-0.

[9] 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(9):553–563, Sep 2008. doi: 10.1016/S1473-3099(08)70156-7. URL http://dx.doi.org/10.1016/S1473-3099(08)70156-7.

[10] M.J. Wawer, R.H. Gray, N.K. Sewankambo, D. Serwadda, X. Li, O. Laeyendecker, N. Kiwanuka, G. Kigozi, M.Kiddugavu, T.s Lutalo, F. Nalugoda, F. Wabwire-Mangen, M.P. Meehan, and T.C. Quinn. Rates of hiv-1 transmission per coital act, by stage of hiv-1 infection, in rakai, uganda. *J Infect Dis*, 191(9):1403–1409, May 2005. doi: 10.1086/429411. URL http://dx.doi.org/10.1086/429411.

[11] T.D Hollingsworth, R.M. Anderson, and C. Fraser. Hiv-1 transmission, by stage of infection. *J Infect Dis*, 198 (5):687–693, Sep 2008. doi: 10.1086/590501. URL http://dx.doi.org/10.1086/590501.

[12] M.S. Cohen, Y.Q. Chen, Marybeth M. McCauley, Theresa Gamble, Mina C. Hosseinipour, Nagalingeswaran Kumarasamy, James G. Hakim, Johnstone Kumwenda, Beatriz Grinsztejn, Jose H. S. Pilotto, Sheela V. Godbole, Sanjay Mehendale, Suwat Chariyalertsak, Breno R. Santos, Kenneth H. Mayer, Irving F. Hoffman, Susan H. Eshleman, Estelle Piwowar-Manning, Lei Wang, Joseph Makhema, Lisa A. Mills, Guy de Bruyn, Ian Sanne, Joseph Eron, Joel Gallant, Diane Havlir, Susan Swindells, Heather Ribaudo, Vanessa Elharrar, David Burns, Taha E. Taha, Karin Nielsen-Saines, David Celentano, Max Essex, and Thomas R. Fleming. Prevention of HIV-1 Infection with Early Antiretroviral Therapy. *N Engl J Med*, 365(6):493–505, July 2011. ISSN 1533-4406. doi: 10.1056/NEJMoa1105243. URL http://dx.doi.org/10.1056/NEJMoa1105243.

[13] D. T. Fleming and J. N. Wasserheit. From epidemiological synergy to public health policy and practice: the contribution of other sexually transmitted diseases to sexual transmission of hiv infection. *Sex Transm Infect*, 75(1):3–17, Feb 1999.

[14] H. Grosskurth, F. Mosha, J. Todd, E. Mwijarubi, A. Klokke, K. Senkoro, P. Mayaud, J. Changalucha, A. Nicoll, and G. ka Gina. Impact of improved treatment of sexually transmitted diseases on hiv infection in rural tanzania: randomised controlled trial. *Lancet*, 346(8974):530–536, Aug 1995.

[15] Q.A. Abdool Karim, S.S. Abdool Karim, J.A. Frohlich, Anneke C Grobler, Cheryl Baxter, Leila E Mansoor, Ayesha B M Kharsany, Sengeziwe Sibeko, Koleka P Mlisana, Zaheen Omar, Tanuja N Gengiah, Silvia Maarschalk, Natasha Arulappan, Mukelisiwe Mlotshwa, Lynn Morris, Douglas Taylor, and C. A. P. R. I. S. A. 004 Trial Group. Effectiveness and safety of tenofovir gel, an antiretroviral microbicide, for the prevention of hiv infection in women. *Science*, 329(5996):1168–1174, Sep 2010. doi: 10.1126/science.1193748. URL http://dx.doi.org/10.1126/science.1193748.

[16] G.M. Myers and K.H. Mayer. Oral preexposure anti-hiv prophylaxis for high-risk u.s. populations: current considerations in light of new findings. *AIDS Patient Care STDS*, 25(2):63–71, Feb 2011. doi: 10.1089/apc.2010.0222. URL http://dx.doi.org/10.1089/apc.2010.0222.

[17] E. M. Connor, R. S. Sperling, R. Gelber, P. Kiselev, G. Scott, M. J. O'Sullivan, R. VanDyke, M. Bey, W. Shearer, and R. L. Jacobson. Reduction of maternal-infant transmission of human immunodeficiency virus type 1 with zidovudine treatment. pediatric aids clinical trials group protocol 076 study group. *N Engl J Med*, 331(18):1173–1180, Nov 1994. doi: 10.1056/NEJM199411033311801. URL http://dx.doi.org/10.1056/NEJM199411033311801.

[18] K. M. De Cock, M. G. Fowler, E. Mercier, I. de Vincenzi, J. Saba, E. Hoff, D. J. Alnwick, M. Rogers, and N. Shaffer. Prevention of mother-to-child hiv transmission in resource-poor countries: translating research into policy and practice. *JAMA*, 283(9):1175–1182, Mar 2000.

[19] L Mofenson IC Hanson J Pitt C Diaz K Hayani E Handelsman V Smeriglio R Hoff W Blattner E.R. Cooper, M. Charurat. Combination antiretroviral strategies for the treatment of pregnant hiv-1-infected women and prevention of perinatal hiv-1 transmission. *J Acquir Immune Defic Syndr*, 29(5):484–494, 15 April 2002.

[20] T. Zutlevics. Should art be offered to hiv-serodiscordant and hiv-seroconcordant couples: an ethical discussion? *Hum Reprod*, 21(8):1956–1960, Aug 2006. doi: 10.1093/humrep/del082. URL http://dx.doi.org/10.1093/humrep/del082.

[21] 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 adn K.H. Fiffe, J. Kiarie, C. Farquhar adn G. John-Stewart, A. Kakia, J. Odoyo, A Macunguzi, E. Nakku-Joloba, R. Twesigye, K. Ngure, C. Apaka, H. Tamooh, F. Gabona, A. Mujugira, D. Panteleef adn K.K. Thomas, L. Kidoguchi, M.Krows adn J. Revall, S. Morrison adn H. Haugen, M. Emmanuel-Ogier adn L. Ondrejcek, R.W. Coombs, L. Frenkel, C. Hendrix adn N.N. Bumpus, D. Bangsberg, J.E. Haberer, W.S. Stevens, L.R. Lingappa, and C. Celum. Antiretroviral prophylaxis for hiv prevention in heterosexual men and women. *The New England Journal of Medicine.*, 367:399–410, July 2012. doi: 10.1056/NEJMoa1108524.

[22] B. M. van Noord-Zaadstra, C. W. Looman, H. Alsbach, J. D. Habbema, E. R. te Velde, and J. Karbaat. Delaying childbearing: effect of age on fecundity and outcome of pregnancy. *BMJ*, 302(6789):1361–1365, Jun 1991.