Proceedings of the 2018 Winter Simulation Conference
M. Rabe, A. A. Juan, N. Mustafee, A. Skoogh, S. Jain, and B. Johansson, eds.

PARAMETERIZATION AND CALIBRATION OF MICRO-SIMULATION MODEL FOR CERVICAL CANCER AND HIV IN ZAMBIA

Kasey Jones
Joey Morris
Georgiy Bobashev

Center for Data Science
RTI International
3040 E Cornwallis Rd
Durham, NC 27709, USA

Sujha Subramanian

Public Health Research Division
RTI International
307 Waverley Oaks Road
Waltham, MA 02452, USA

ABSTRACT

Zambia has one of the highest rates of cervical cancer in the world. To help policy makers make future intervention decisions regarding cervical cancer, we created a micro-simulation model to simulate cervical cancer and HIV in Zambia. Model calibration faced two major challenges: (1) Much of the available input data was on women in the United States, which do not allow calibration to align to the age-specific targets from Zambia; (2) Significant computational resources were not always available. We addressed issue one by creating age-specific calibration parameters to help better match specific targets. Issue two was addressed by using predictive models before conducting calibration simulations to discard parameter sets that are likely to produce poor results. This paper will demonstrate these two modeling strategies and show the dramatic effect they had on our ability to accurately calibrate to model targets.

1 INTRODUCTION

Cervical cancer lead to almost 266,000 deaths in 2012 (Subramanian et al. 2008). This is particular a problem in countries where resources and screening programs are scarce. Among women in Sub-Saharan Africa, cervical cancer is the most common cause of cancer-related deaths (GLOBOCAN 2014). A high-prevalence of HIV, mixed with the fact that women who have HIV are six times more likely to get cervical cancer, make this region prone to higher cervical cancer rates (Martel et al. 2012).

There are low-cost strategies available to prevent cervical cancer. Testing the impact of these strategies is expensive and in countries with already limited resources may not be possible. Micro-simulations can model the progress of human papillomavirus (HPV) and human immunodeficiency virus (HIV) in Zambia and can access the impact and cost of intervention strategies. For the result of these tests to be trusted, the micro-simulation needs to calibrate well to known rates of cervical cancer, HPV, and HIV in Zambia. Our goal is develop a micro-simulaiton that will allow researches to assess the cost-effectiveness of interventions for fighting cervical cancer in Zambia and build a bases for conducting similar research in other countries.

1.1 Micro-Simulation Basics

On top of modeling HPV and HIV, our micro-simulation models cancer detection, cancer progression, and current life status. Each component is made up of *states*. For example, if a woman is alive, her life state is *alive*. The current state for a woman is advanced through monthly time-steps from age nine until death or age 100. Transitions among the five states (HIV, HPV, cancer detection, cancer progression, and life) are generated using probabilities that account for the woman's other states and current age.

The micro-simulation simulates progression and regression of HPV for four different strains: low-risk (LR), strain 16, strain 18, and other high-risk strains (HR). HPV typically progresses from undetectable HPV (non-HPV) to HPV, HPV to CIN1, HPV to CIN2,3, and finally from CIN2,3 to cancer. As seen in figure 1 however, transitions do not have to follow a linear path, and regression to a lower state of HPV at each stage is possible until cancer is reached.

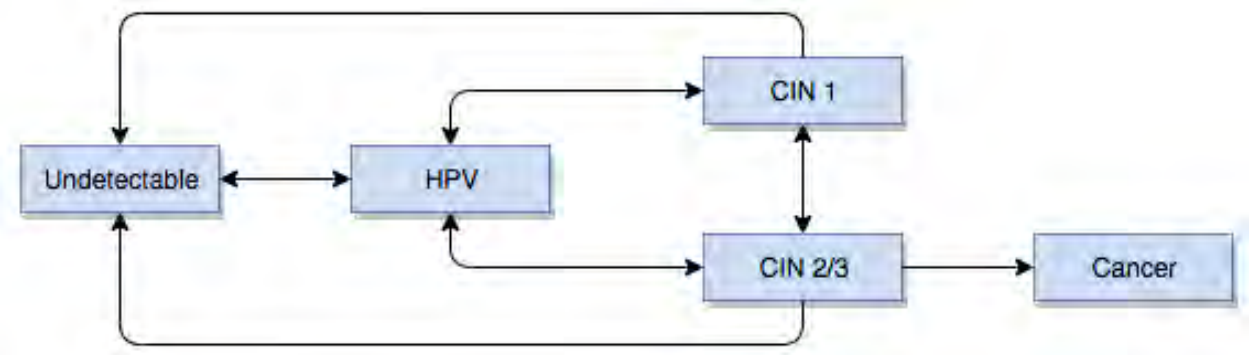

Figure 1: Possible HPV transitions.

Three of the remaining components of our model can only progress forward and are not allowed to revert. Once a woman has transitioned to HIV, she cannot regress to non-HIV. Once a woman's cancer has been detected, it cannot be undetected. And finally, if a woman dies, she remains dead. The last component, cancer status, also progresses linearly but has more states. This is shown in figure 2.

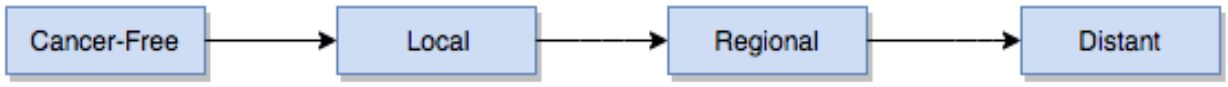

Figure 2: Possible cancer states.

1.2 Calibration

The calibration of our model relied on several targets. We used target data from Zambia whenever possible but supplemented information from South African whenever Zambian specific targets were unavailable. These targets included: HPV prevalence rates by strain and HIV status (Richter et al. 2003), CIN2,3 prevalence rates by strain and HIV status (McDonald et al. 2014), overall HIV prevalence (Kharsany and Karim 2016), and total cancer incidence by strain (Kapambwe et al. 2016). Our last target was cause of cancer. Work by Juckett and Hartman-Adams suggests that 50% of cancer is caused by strain 16, 30% from other high-risk strains, and 20% from strain 18 (2010). The LR strain cannot cause cancer in our model. Table 1 provides a breakdown of the 106 calibration targets.

Table 1: Breakdown of 106 targets.

Type	Strain	HIV Type	Age Groups	Total Targets
HPV Prevalence	LR, HR	HIV, Non-HIV	8	32 (2*2*8)
HPV Prevalence	16, 18	Overall	8	16 (2*1*8)
CIN2,3	16, 18, HR	HIV, Non-HIV	8	48 (3*2*8)
HIV Prevalence	Overall	Overall	2	2
Cancer Incidence	Overall	Overall	5	5
Cause of Cancer	16, 18, HR	Overall	Overall	3

We created a base model that contained the same parameters as Goldhaber-Fiebert et al. (2007) with the addition of a few other parameters, specifically for women with HIV. Table 2 shows 17 of the 42 total parameters we included in our model.

Table 2: Five progression parameter types.

From State	To State	HPV Strains
Non-HPV	HPV	LR, HR, 16, 18
HPV	CIN1	LR, HR, 16, 18
HPV	CIN2,3	HR, 16, 18
CIN1	CIN2,3	HR, 16, 18
CIN2,3	Cancer	HR, 16, 18

In addition to the 17 progression parameters in table 2, our model also has 11 regression parameters, 11 progression parameters specific to women with HIV, and 3 progression parameters for women with some form of HPV immunity. This resulted in 42 total parameters that need to be calibrated.

These parameters effect the base transition rate for progression or regression of HPV. For calibration, a randomly selected number from a distribution is selected for each of the 42 parameters. Distributions that Goldhaber-Fiebert (2007) provided were altered based on comparing Zambian targets to known U.S. values. We raised the mean, lower, and upper limit of several parameters based on how different the targets were. Distributions for parameters not provided were derived using several runs of calibration to find potential upper and lower bounds.

1.3 Data Limitations

There have been several models developed for simulating HPV and cervical cancer in the U.S. It is difficult to replicate these models for Zambia because of a lack of input transition data and possible parameter distributions. To create a micro-simulation , progression and regression rates were taken from studies on U.S. women. Specifically, we used rates from Goldhaber-Fiebert et al. (2007) and Kim et al. (2007).

We used U.S. parameters and transitions because we did not expect the natural history of the disease to substantially differ between countries. It is not that we are using U.S. specific transition rates at all, we are using parameters from the United States to start the calibration of the natural history model. All of our input rates are locally relevant for Zambia. Whenever possible, values for Zambian women were used as inputs. Mortality rates were taken from the Institute for Disease Modeling *EMOD Software* (IDMOD 2017) and transition rates from non-HIV to HIV were produced for Zambia using Avenir Health's *Spectrum* (Avenir-Health 2017).

Unfortunately, most targets for the United States and Zambia do not follow similar shapes (i.e. prevalence and incidence rates do not decrease or increase at the same right across age groups, see figure 3). If we are using transition rates and parameters that went into a model built for the United States, we should not expect the output of our model to match a mix of prevalence and incidence rates from Zambia and South Africa. We can help solve this problem by using age-specific parameterization.

1.4 Computational Limitations

Each simulation is completed in three steps. 42 parameters are randomly drawn from their respective distributions and transition probabilities are created based on these values, the women's age, HIV status, immunity status, and HPV strain. After preliminary setup is complete, a simulation is ran and the results are analyzed. Each simulation from start to finish takes approximately 15 minutes. Running 10,000 scenarios would take almost a month at four simulations at a time. If we could run 100 simulations at a time, it would still take over one day to finish. As we demonstrate in the section 4, even 10,000 scenarios may not be enough for the final calibration.

The majority of these simulations would not be worth analyzing. Consider that several targets are influenced heavily by at least four parameters. All parameters for all targets need to align for the calibration to be a success. Furthermore, since we don't know what the right parameters are, we should not restrict their possible values too heavily. With 42 parameters, if we ran a complete grid search using just 10 steps between each parameters lower and upper limit, we would need 10^{42} scenarios. To help solve this resource constraint, we can use predictive models to eliminate poor performing parameter sets before we even run the scenario.

2 ACCOUNTING FOR AVAILABLE INPUT DATA: AGE-SPECIFIC PARAMETERIZATION

The basic structure of our cervical cancer micro-simulation was built to replicate the results of Goldhaber-Fiebert et al. (2007). This includes using all of the parameters that Goldhaber-Fiebert used for the transition of HPV by strain. We will call these parameters base parameters for the rest of this paper to help distinguish them from other parameter types. We will also refer to the value that is selected for a base parameter as a base multiplier.

Let us consider what applying a base multiplier will do to a target in our model. Since base parameters effect all women, regardless of their age, the current target at every age group will be shifted up or down. Although there are interactions occurring among base parameters, we can generally expect that the shape of the output for a specific target will look similar regardless of the multipliers used in the model. An example of this is seen in figure 3.

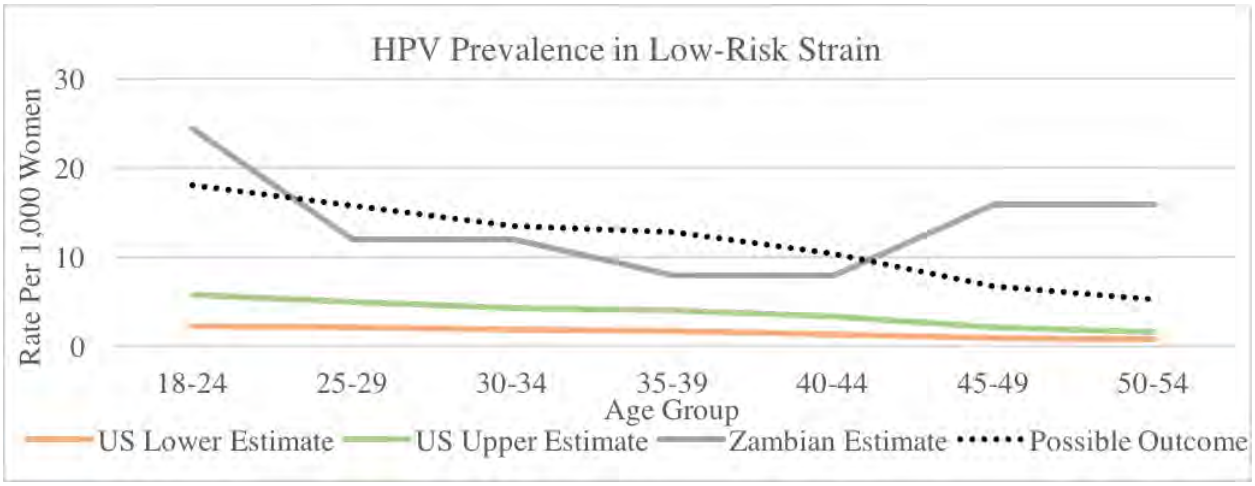

Figure 3: Possible effects of multiplier for one target.

This is problematic. Using the U.S. transition rates may provide a baseline, but women in Zambia do not have the same amount of HPV at each age group that women in the U.S. do. Figure 3 shows a comparison between the targets for HPV LR prevalence for the U.S. and Zambia. The model from Goldhaber-Fiebert et al. (2007) was calibrated to fall within the two U.S. estimates. We can expect that our output will have a similar shape to those lines and although we can raise and lower the output, we cannot change its shape.

Using a single parameter for transition from non-HPV to HPV in the low-risk strain will *likely* never produce an outcome curve shaped like the Zambian estimate in figure 3. The probable outcome is that we achieve a line close to the black dotted line.. This dotted line is closer to our target but it still does not match. The dotted line has a mean average error from the Zambian target of 5.58.

2.1 Age-Specific Parameterization: Concept

We fixed most of the error caused by using transition data and parameters from the United States by implementing age-specific parameterization. We created parameters for each age group and target so that we could manipulate individual sections of the black dotted line. We then calibrated the age-specific parameters using an iterative algorithm before running scenarios to calibrate the base parameters. This allows the black dotted line to have a similar shape to the Zambian estimate that is our target, as seen in figure 4. This dotted line has a mean average error from the Zambian target of only 1.23.

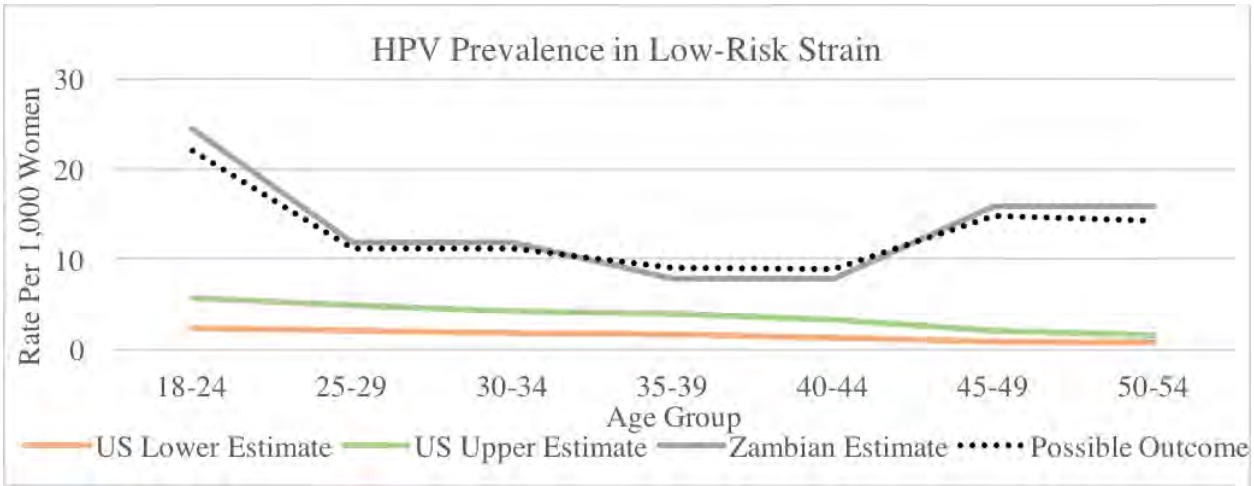

Figure 4: Possible effect of using age-specific calibration.

Since our computational resources are already strained, simply adding an age-specific parameter for each age and calibrating all parameters at once is not possible. We would end up with 140+ total parameters and even several million calibration runs would likely not produce results. This is the reason age-specific parameterization is completed before the final calibration.

2.2 Age-Specific Parameterization: Walk-through

To begin, identify base multipliers that will produce scenarios that match only the first age group of each target to within 10%. This is done iteratively; simply raise or lower base multipliers until the first age group of each target is matched. Table 3 shows how this might work for one target and one parameter. The goal is to get the output target value to be within 10% of the actual target value (in table 3 this value is 24.4) Using a multiplier of 4.5 produced an output of 24.1. We continue tweaking parameters until all of the first age groups are within 10%.

Table 3: Matching first age group by finding starting base multipliers.

Target	Multiplier: 1	Multiplier: 2	Multiplier: 4.5
24.4	8.0	13.5	24.1

After all base multipliers have been set we complete the age-specific parameterization in a specific order. One of the targets for the overall calibration is which strain caused cervical cancer. Strain 16 needs to cause 50% of cancer incidence, so we start the calibration on strain 16, age group 25-29. When we find a value within our bounds, we move to strain high-risk, age group 25-29. For each age group, we used the following order for calibration: 16-Overall, HR-Non-HIV, HR-HIV, 18-Overall, LR-Non-HIV, and LR-HIV.

This approach is taken for three reasons. (1) Although strains are independent, once a person moves into the cancer state for any strain, they cannot get cancer from another strain. Multiple strains cannot cause a person to reach the cancer state. Strain 16 needs to cause 50% of all cancers in our model, so its age parameters are calculated first. (2) We cannot calculate all age groups at the same time. If we find an appropriate age-specific parameter for age group 30-34 of the high-risk strain, the 25-29 parameter may not be locked. When all parameters have been calibrated, the 30-34 parameter would no longer produce the desired change because the 25-29 value would have changed. (3) We tried to calibrate the parameters without any specified type or age group order. After 20-25 parameters become locked, no additional parameters could calibrate (i.e. setting a parameter to 0 was still not capable of producing the desired drop in prevalence).

Note that we completed this calibration process three times: once for HPV prevalence (48 parameters), once for CIN2,3 prevalence (48 parameters), and once for HIV prevalence (2 parameters).

2.3 Age-Specific Parameterization: Calibration

Age group 18-24 for each target is already set because of the static base multipliers we use during this calibration. Therefore, the age-specific parameters for age group 18-24 are set to 1. When calibrating the parameter for an age group, we want to see an appropriate change from the previous age group. For low-risk HPV prevalence, the target decreases by 51% from the first age group to the second. We begin a searching algorithm to find a multiplier for the non-HPV to HPV parameter that will drop HPV prevalence by 51%. This is achieved through the following steps:

1. Calculate the effect of using multipliers .2, .7, 1.5, and 2.5 (an arbitrary net casted around the starting value of 1.0) for the first parameter
2. If the desired change and actual change produced by one of these multipliers are within 2% of each other, save the multiplier value and move to the next parameter
3. Otherwise, find which multiplier worked best (.2, .7, 1, 1.5 or 2.5)
4. Cast another net based on how close the desired and actual change are. The net consists of 4 new multipliers based off the previous rounds best multiplier:
   (a) multiplier +/- 3 * best multiplier * difference between desired and actual change
   (b) multiplier +/- best multiplier * difference between desired and actual change
5. If after recasting the net, the best multiplier remained constant, reduce how far the net is cast
   (a) multiplier +/- 3 * (desired - actual change) / number of rounds for current multiplier
   (b) multiplier +/- (desired - actual change) / number of rounds for current multiplier
6. If the best multiplier is not within 2% of the desired reult and has not changed after 5 rounds, recast the net of varying random values until a new best multiplier is found
7. Repeat steps two through six until an appropriate multiplier is identified

Let us give a concrete example. When testing .2, .7, 1.5, and 2.5 in hopes of producing the 51% decrease mentioned earlier, multiplier .2 produced a 53% decrease, the closest value to 51%. Therefore, .2 becomes our new best multiplier. Since $|.53 - .51| \not< .02$, we cast our net around .2 with values:

$$.2 \pm 3 * .2 * (.13 - .11) \tag{1}$$

$$.2 \pm 1 * .2 * (.13 - .11) \tag{2}$$

Scenarios are ran using .188, .196, .20, .204, and .212. This time, .212 produced a decrease of 51.5%. As $|.515 - .51| < .02$, we finish the calibration for this parameter and move on. This process is completed for each age-specific parameter. Table 4 shows the results of age-specific parameterization for one target.

Table 4: Results of age-specific parameterization for HPV prevalence in the low-risk strain.

Age Group	Target	Change	Parameter
18-24	24.4	N/A	1
25-29	11.9	.51	.2
30-34	11.9	0	.7
35-39	7.9	.336	.206
40-44	7.9	0	.519
45-49	15.9	-1.01	1.822
50-54	15.9	0	1.970
55+	15.9	0	2.711

2.4 Issues with Age-Specific Parameters

Age-specific parameterization worked well for HPV prevalence targets. All 48 age-specific parameters were able to produce the desired drop or rise in the next age group to within 2%. It should be noted, however, that this was calibrated using only 5,000 women and the base multipliers were fixed. An age-specific multiplier that caused a 51% reduction from one age group to the next, will likely not produce a 51% reduction for a different scenario with different base multipliers.

CIN2,3 is not as prevalent as HPV, especially for women who do not have HIV. In some cases, this value can be as low as .15%. Calibrating a reduction from .15% to to .14% is difficult when the number of women in this category is low. Although most age-specific parameters for CIN2,3 calibrated well, there were a few parameters that did not.

2.5 Recapping the Steps

The following is a recap of the generic steps taking for age-specific parameterization.

- Find base multipliers that create a scenario matching all model target's first age group to within 10% (this process is done iteratively)
- Lock these base multipliers for the entire age-specific parameterization and set age-specific parameters to 1 for the first age group. The first age group parameters will always be 1
- Select an order in-which to complete the calibration. This order is used for each age group (i.e. age group 1: item 1, item 2, item 3; age group 2: item 1, item 2, item 3; etc.) and should be based on which targets are most important to the model
- Create an algorithm that will help find age-specific parameters that generate the appropriate increase or decrease in target value to the next age group
- Run the algorithm until all age-specific parameters have been calibrated to provide the desired increase or decrease across all age groups

3 LIMITED RESOURCES

Micro-simulations are notoriously difficult to calibrate, especially when data may be unreliable and the prior parameter distributions are simple borrowed from another model. Rutter et al. noted these difficulties and provided a Bayesian approach as a solution (Rutter, Miglioretti, and Savarino 2009). We provide an alternative solution that makes use of simple predictive models to discard poor performing parameter sets before running the calibration.

3.1 Selecting Multipliers

We now have a dataset that includes the random base multipliers that were selected and how well each scenario matched individual targets. We also know which parameters effect outcome targets the most. For example, low-risk HPV prevalence is effected by the low-risk parameters: non-HPV to HPV, HPV to CIN1, CIN1 to non-HPV, and HPV to non-HPV.

Using a subset of this dataset, we created a statistical prediction model for predicting error for low-risk HPV prevalence (non-HIV), based on the four parameters that were just mentioned. Table 5 shows an example of the target values and possible outcome values from a scenario.

Table 5: Possible outcome for low-risk HPV prevalence for a random scenario.

Age Group	Target	Scenario J
18-24	24.4	23
25-29	11.9	10
30-34	11.9	11
35-39	7.9	7
...	...	...

The mean absolute error for any target with age groups is calculated as:

$$MAE_{tj} = \frac{\sum_{i=i}^{n} |t_i - s_{ji}|}{n}$$

where n is the number of age groups for the target, j is the current scenario, t is the current target, and s is the output from the scenario j.

We split the 10,000 records using an 80/20 training-testing split and built a simple random forest model. Figure 5 shows plots of the predicted error vs actual error for HPV prevalence in the low-risk strain. In this case, the error is the mean absolute error for the eight age groups.

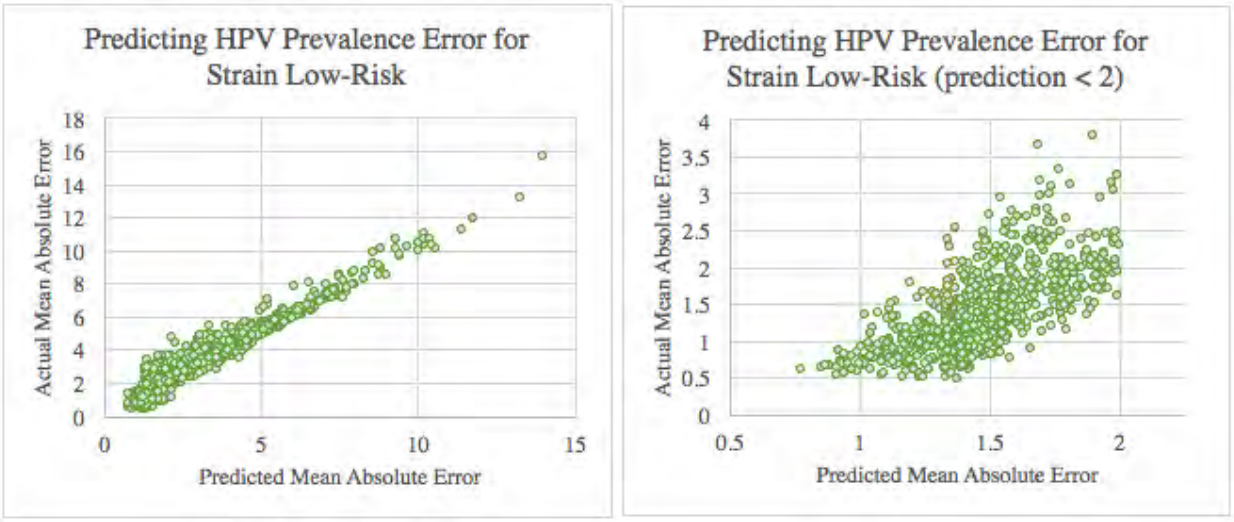

Figure 5: Predicting low-risk HPV prevalence error using selected multipliers.

Looking at parameter sets that had a predicted error of less than 2, we see that the best actual results occurred here. No scenario with a predicted error above 2 had an actual error below 1.5. Before running any future scenarios, we can run the randomly selected base multipliers through our prediction model, and discard any set of selected multipliers that had a predicted error greater than some selected baseline value.

We repeated this process for fourteen of the fifteen targets. Each prediction model was a simple random forest consisting of 100 trees and a max depth of 10. Fine-tuning of individual models was not needed as each model performed well with the base settings.

Baseline cutoff values were set to the highest predicted error for the best 10% of the actual errors (using the test dataset). In the example above, the lowest 10% of actual errors had a predicted error less than 1.586. Only 34% of the randomly selected base multiplier sets have a predicted error less than 1.586. The final cutoffs and the reduction in scenarios are shown in the table 6. We changed the cutoff of some models depending on the accuracy of the individual model. Note that the HIV prevalence target did not need to be tested because the target was already matching for all calibrations.

Table 6: Percent of parameter sets dropped by predictive model.

Test	Type	Strain	HIV Status	Cutoff	Tests Passing
1	HPV Prevalence	Low-Risk	Non-HIV	1.586	33.9%
2	HPV Prevalence	Low-Risk	HIV	5	49.9%
3	HPV Prevalence	High-Risk	Non-HIV	2.563	47.8%
4	HPV Prevalence	High-Risk	HIV	6	49.8%
...	...	...	...	...	...
13	HIV Prevalence	Overall	Overall	n/a	n/a
14	Cancer Incidence	Overall	Overall	45	71.1%
15	Cause of Cancer	16, 18, or HR	Overall	0.15	77.0%

w Not all of the tests were independent. If a set of parameters fails one test, it is likely that the set would have failed another test. It took 66,743 randomly generated multiplier sets to find 25 that passed all 14 tests. This is approximately 1 in 2,700.

4 RESULTS

Assessing the results of these calibrations is not straight forward as we have 106 targets. Most values are prevalence values for eight age groups. However, some targets are for different age ranges (HIV prevalence), some are incidence values, and the cause of cancer target is just a number that describes which strain caused a percentage of cancer. Furthermore, some prevalence values are as large as 64% and as low as .07%. Being off by +/- 30% for such small numbers could be considered successful, but being off by 30% for the larger prevalence values would be a terrible calibration.

We used the average percent difference across all targets as the our accuracy measure, with the sole stipulation that strain 16 caused between 45-55% of cancer. We calculated absolute percent difference as:

$$\frac{|target - output|}{target}$$

4.1 Calibration 1: Calibrating with Base Parameters Only (1,000 scenarios)

For calibration 1, we randomly selected values for the 42 base parameters. We **did not** use age-specific multipliers from section 2 or the models from section 3. The lower and upper bounds for the base parameters were based on parameters provided by Goldhaber-Fiebert (2007) and the Zambian target values. The best scenario produced an average absolute percent difference of 85.5. This is a terrible result.

4.2 Calibration 2: Age-Specific Parameterization (1,000 & 10,000 scenarios)

Calibration 2 included the addition of the age-specific parameterization from section 2 and this dramatically improved the results of the calibration. If we only randomly selected 1,000 base multiplier sets for calibration, the best scenario produced an average absolute percent difference of 24.3.

When 10,000 scenarios were completed, the best scenario produced an average absolute percent difference of 19.5. 51 targets were within 10% of their actual target, and 82 targets were within 30%. We have provided results at both the 1,000 and 10,000 scenario level to show that the addition of age-specific parameterization is the reason for the massive increase in accuracy and not just running the model 10,000 times.

4.3 Calibration 3: Age-Specific Parameterization + Base Parameters with Selection (10,000 scenarios)

Our final calibration included using the output of the 10,000 scenarios in calibration 2 to help model which set of multipliers would likely produce good results. The best scenario in our final calibration had a mean absolute percent difference of only 17.3% It had a mean absolute error of just 0.89. Although only 50 targets were within 10% of the actual target, 88 were within 30%. The only targets above 30% were for CIN2,3 prevalence. These targets are generally the smallest targets we had to calibrate to, with almost all of them being less than 1%.

4.4 Final Results

We conclude the results section with two tables. Table 7 shows the overall comparison metrics for the three calibrations. The numbers represent the best single scenario that resulted from each calibration. We see that the biggest gain in accuracy was due to the implementation of age-specific parameters. The calibration process received an additional boost by testing multiplier sets before actually running scenarios. Table 8 shows the average absolute percent differences across all ages for the best scenario of each calibration run.

Table 7: Overall metrics for the three calibrations.

	Calibration			
Metric	1	2 (1k)	2 (10k)	3
Percent Difference	84.49	24.25	19.45	17.25
Mean Absolute Error	5.75	1.44	1.16	0.89
Targets within 10%	12	33	51	50
Targets within 30%	27	77	82	88
Total Computation Time (hours)	3	3	25	28

Table 8: Comparing calibrations for individual targets: mean absolute percent difference.

	Calibration			
Target	1	2 (1k)	2 (10k)	3
HPV - LR - Non-HIV	48.44	12.77	6.31	4.44
HPV - LR - HIV	17.56	11.90	3.88	3.32
HPV - HR - Non-HIV	28.29	12.66	8.00	11.28
HPV - HR - HIV	69.66	8.23	12.31	8.37
HPV - 16 - Overall	71.45	12.83	4.57	3.40
HPV - 18 - Overall	60.87	12.80	10.18	12.05
...	...	...	...	...
HIV Prevalence	2.78	4.70	4.59	1.51
Cancer Incidence	81.26	52.31	28.50	20.32
Cause of Cancer	14.43	38.72	6.34	8.18

5 CONCLUSION

Calibrating our micro-simulation for Zambia faced many challenges. We overcame using imperfect data by implementing age-specific parameterization to help calibrate age-specific targets appropriately. We overcame a lack of time and resources by using predictive models to help determine which sets of randomly selected multipliers should be used in calibration runs. Our final calibration efforts increased the original accuracy from an average absolute percent difference of 85.5 to only 17.3. Researchers now have a well calibrated model to test cost and effectiveness of interventions for cervical cancer and HIV in Zambia.

REFERENCES

Avenir-Health 2017. "Spectrum". http://www.avenirhealth.org/software-spectrum.php.

GLOBOCAN 2014. "Cancer Incidence and Mortality WorldWide". http://globocan.iarc.fr.

Goldhaber-Fiebert, J., N. Stout, J. Ortendahl, K. Kuntz, S. Goldie, and J. Salomon. 2007. "Modeling Human Papillomavirus and Cervical Cancer in the United States for Analyses of Screening and Vaccination". *Population Health Metrics* 5:11.

IDMOD 2017. "Epidemiological Modeling Software". http://www.idmod.org/documentation.

Juckett, G., and H. Hartman-Adams. 2010. "Human Papillomavirus: Clinical Manifestations and Prevention". *American Family Physician* 82(10):1209–1214.

Kapambwe, S., V. Sahasrabuddhe, M. Blevins, M. Mwanahamuntu, V. Mudenda, B. Shepherd, C. Chibwesha, K. Pfaendler, M. Hicks, S. Vermund, J. Stringer, and G. Parham. 2016. "Implementation and Operational Research: Age Distribution and Determinants of Invasive Cervical Cancer in a "Screen-and-Treat" Program Integrated with HIV/AIDS Care in Zambia". *Journal of Acquired Immune Deficiency Syndrome* 70(1):20–26.

Kharsany, A., and Q. Karim. 2016. "HIV Infection and AIDS in Sub-Saharan Africa: Current Status, Challenges and Opportunities". *The Open Aids Journal* 10:34–48.

Kim, J., K. Kuntz, N. Stout, S. Mahmud, L. Villa, E. Franco, and S. Goldie. 2007. "Multiparameter Calibration of a Natural History Model of Cervical Cancer". *American Journal of Epidemiology* 166(2):137–150.

Martel, C., J. Ferlay, S. Franceschi, J. Vignat, F. Bray, D. Forman, and M. Plummer. 2012. "Global Burden of Cancers Attributable to Infections in 2008: A Review and Synthetic Analysis". *Lancet Oncol* 13(6):607–615.

McDonald, A., A. Tergas, L. Kuhn, L. Denny, and T. Wright. 2014. "Distribution of Human Papillomavirus Genotypes among HIV-Positive and HIV-Negative Women in Cape Town, South Africa". *Frontiers in Oncology* 4(48).

Richter, K., P. Becker, and G. Dreyer. 2003. "Age-Specific Prevalence of Cervical Human Papillomavirus Infection and Cytological Abnormalities in Women in Gauteng Province, South Africa". *The South African Medical Journal* 103(5):313–317.

Rutter, C., D. Miglioretti, and J. Savarino. 2009. "Bayesian Calibration of Microsimulation Models". *American Statistical Association* 104(488):1338–1350.

Subramanian, S., D. Ekwueme, J. Gardner, C. Kramer, B. Bapat, and F. Tangka. 2008. "Identifying and Controlling for Program Level Differences in Cost-Effectiveness Comparisons: Lessons from the Evaluation of the National Breast and Cervical Cancer Screening Program (NBCCEDP)". *Evaluation and Program Planning* 31(2):134–144.

AUTHOR BIOGRAPHIES

KASEY R. JONES is a data scientist in the Center for Data Science at RTI International. Kasey holds an M.S in Analytics from the Institute for Advanced Analytics at North Carolina State University and holds both a B.S. and M.S. in Mathematics from Western Carolina University. While working at RTI, Kasey has worked on several simulation and predictive modeling projects including modeling hospital acquired infections in North Carolina (U.S.) hospitals, as well as predicting underage alcohol and marijuana use in

the District of Columbia (U.S.). Kasey hopes to continue developing models that benefit society and make an impact in this world. His e-mail address is krjones@rti.org.

JOEY MORRIS is a data scientist in the Center for Data Science at RTI International. He holds a B.S. in Statistics and a B.S. in Applied Mathematics, both from North Carolina State University. He has helped develop and publish several agent-based models and micro-simulations focusing primarily on public health issues, including colorectal cancer screening, HIV transmission, influenza transmission, and heroin market dynamics. His e-mail address is rjmorris@rti.org.

GEORGIY BOBASHEV is an RTI Fellow in the Center for Data Science at RTI International with over 20 years of experience in health research. He holds a PhD in Biomathematics from the North Carolina State University. His current research interests follow two major areas: predictive modeling and substance use/risky behaviors studies. Predictive methods often combine mechanistic (e.g., agent-based and system dynamics) and machine learning techniques. Dr. Bobashev has applied modeling, statistical analysis and experimental design to a variety of health- and policy-related areas, including cancer, substance use, HIV, child/maternal health, influenza, diabetes, and violent behavior. His information is available at https://www.rti.org/expert/georgiy-bobashev and his contact email is bobashev@rti.org.

SUJHA SUBRAMANIAN, a Fellow in Health Economics and Policy Research at RTI International, has extensive experience performing economic burden assessments and evaluations of noncommunicable disease (NCD) screening programs both in the United States and in international settings. Over the past two decades, she has directed several program evaluations, including the assessment of the cost and effectiveness of the National Breast and Cervical Cancer Early Detection Program (NBCCEDP) and the Colorectal Cancer Control Program (CRCCP). In collaboration with the World Health Organizations' International Agency for Research on Cancer, she is assessing the implementation and scale-up of cervical cancer screening in India. Dr. Subramanian is also the principal investigator for an R01 grant that is evaluation best practices to scale-up cervical cancer prevention and screening in Sub-Saharan Africa. . Her email address is ssubramanian@rti.org and more information is available at https://www.rti.org/expert/sujha-subramanian.