Data-driven modelling, optimization and control in areas of vaccination and policy for HPV transmission

S Vishnu Tej
CE19B084
IDDD - Data Science
IIT Madras, Tamil Nadu, India
ce19b084@smail.iitm.ac.in
Guide: Prof. Nirav P Bhatt

June 6, 2024

Contents

1 Abstract **3**

2 Introduction **3**

2.1 HPV - Biological Background 3

2.2 HPV infection in India 4

3 Literature Review **5**

3.1 HPVsim 5

3.1.1 Model Structure 5

3.1.2 Results 6

3.2 HPVadvise - LIMC 7

3.2.1 Model Structure 7

3.2.2 Model Parameterization - Calibration 8

3.2.3 Analysis 9

3.2.4 Results 10

3.3 SICR Compartmental Model 11

3.3.1 Model Formulation 11

3.3.2 Parameter Estimation 12

3.3.3 Simulation results 13

3.4 Updated Compartmental Model 14

3.4.1 Model Structure 14

3.4.2 Simulation results 17

4 Screening and Vaccination Models **18**

4.1 Simplified Screening Model 18

4.1.1 Simulation Results 20

4.1.2 Optimal Control Theory Formulation 21

4.2 Vaccination Model 23

4.2.1 Simulation Results 24

4.2.2 Optimal Control Theory Formulation 25

5 Conclusion: **26**

6 Future Work: **27**

References **28**

1 Abstract

This research paper presents a novel compartmental model for simulating Human papillomavirus (HPV) transmission in India, incorporating interventions such as screening and Vaccination. The model aims to enhance our understanding of HPV dynamics in the country, considering its diverse population and healthcare challenges. A comprehensive literature review is provided, focusing on existing models like HPVsim and HPVadvise, along with a compartmental model that was developed by me as a part of my Systems Biology coursework. By integrating screening and vaccination interventions into the model and performing optimal control theory, this study contributes valuable insights into the potential impact of preventive measures on HPV transmission dynamics in the Indian context, offering a basis for informed public health strategies and policy recommendations.

2 Introduction

2.1 HPV - Biological Background

Human Papilloma Virus (HPV) is a common virus transmitted through skin-to-skin contact, primarily during sexual activity. Once inside the body, it infects the mucosal cells of the cervix, genitals, and oropharynx (mouth and throat). Most infections clear up on their own without causing any problems.

Some HPV types, particularly 16 and 18, are classified as "high-risk" due to their potential to cause cancer. If these types persist for years, they can disrupt the normal cell cycle and trigger abnormal cell growth.

Cervical Intraepithelial Neoplasia (CIN): This term describes precancerous changes in the cervical epithelium (lining). There are three grades of CIN, numbered 1 to 3:

1. CIN 1 (mild dysplasia): Mildly abnormal cells confined to the lower layers of the epithelium. Usually cleared naturally in most cases.

2. CIN 2 (moderate dysplasia): More extensive and deeper involvement of abnormal cells. Requires monitoring or treatment depending on various factors.

3. CIN 3 (severe dysplasia): Precancerous cells involve most of the epithelium thickness. More likely to progress to cancer if left untreated.

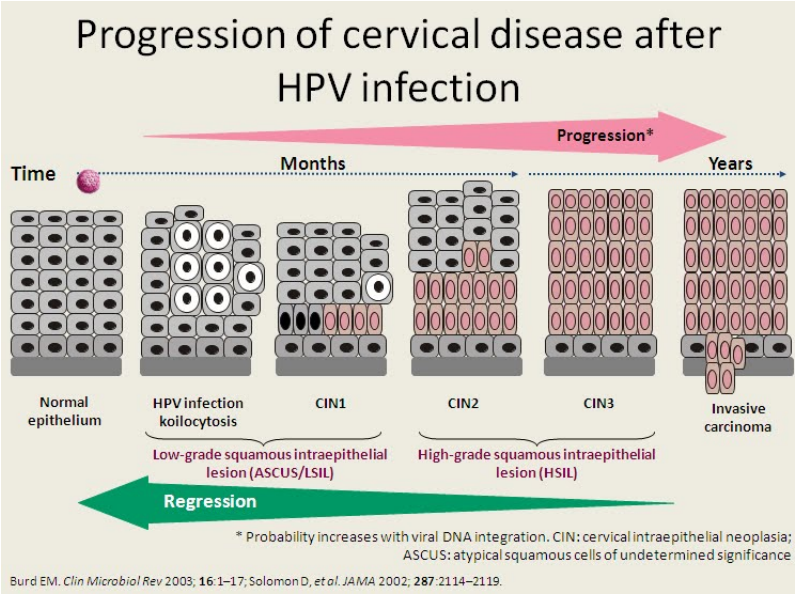


Figure 1: Progression of HPV infection

Several factors can influence the progression of CIN to cancer, including the type of HPV (high-risk vs. low-risk), Duration of infection, Immune system function Smoking and Other co-infections

2.2 HPV infection in India

Human Papilloma Virus (HPV) infection is a global health concern, affecting millions of individuals worldwide, with significant implications for public health. In India, HPV infection is a substantial health issue, with a diverse impact on different regions and demographics. The country bears a considerable burden of cervical cancer, primarily attributed to high-risk HPV types. India's vast and diverse population, coupled with variations in healthcare accessibility, contributes to the complexity of HPV's prevalence and impact. Compared to the world, India faces a disproportionate burden of cervical cancer with about 22 per cent of global cases and deaths, the country's HPV prevalence is double the global average

This alarming situation has prompted the Indian government to take action. In the 2024 Interim Budget, Union Finance Minister Nirmala Sitharaman announced plans to focus on vaccination against cervical cancer for girls aged 9

to 14, aiming to proactively combat HPV infection and prevent future cases of cervical cancer. While HPV vaccination offers a powerful preventive approach, its effectiveness must be evaluated alongside existing screening programs. This project proposes a novel compartmental model to analyze the interplay of the impact of various screening and vaccination strategies in the Indian context.

3 Literature Review

3.1 HPVsim

HPVsim is a robust, open-source tool for simulating the dynamics of HPV transmission and its progression within populations. It operates at the individual level, modelling agents within a sexual network who can contract and transmit various HPV strains. HPV infections in the model can then progress through stages of dysplasia potentially leading to cervical cancer, with the option for spontaneous regression. Furthermore, HPVsim allows the incorporation of interventions like vaccination, screening, and treatment to evaluate their impact on the modelled population. This makes it a valuable tool for public health researchers and policymakers to study and optimize strategies for controlling HPV and its associated diseases.

3.1.1 Model Structure

HPVsim builds a virtual world to study the spread of HPV and cervical cancer. Individuals are represented by agents with characteristics like age, sexual behaviour, and HPV status. These agents interact based on real-world patterns, such as people preferring partners of similar ages. HPVsim can simulate the spread of different HPV strains, some of which can cause cancer. The model tracks how these strains spread through sexual contact, considering factors like a person's immune system and the type of HPV they have. The probability of a person i infected with genotype g transmitting to a susceptible person j within a given time step can be written as

$$\lambda_{i,j,g}(t) = (1 - (1 - \beta_g(1 - eff_{cond} \times cond))^{n(t)} \times (1 - inf\_imm_{j,g}(t)) \quad (1)$$

Where
β_g: per act probability of genotype g
eff_{cond}: efficacy of condoms

cond: probability of condom use within this partnership
n(t): no of sexual acts in time dt
$inf\_imm_{j,g}(t)$: person's protective immunity against infection of 'g' genotype

HPVsim goes beyond simply simulating transmission. It also allows researchers to see how interventions like vaccination, screening, and treatment impact the spread of HPV and cervical cancer. The model can be customized to reflect specific populations and real-world data. By running simulations with different interventions, researchers can identify the most effective strategies for preventing cervical cancer

3.1.2 Results

The number of new infections/CINs/cancers annually, the number of individuals in a particular disease state at a given time, and cumulative aggregates of individuals who have received interventions (e.g., the total number ever screened, vaccinated, or treated) are obtained by running a simulation for the Indian population.

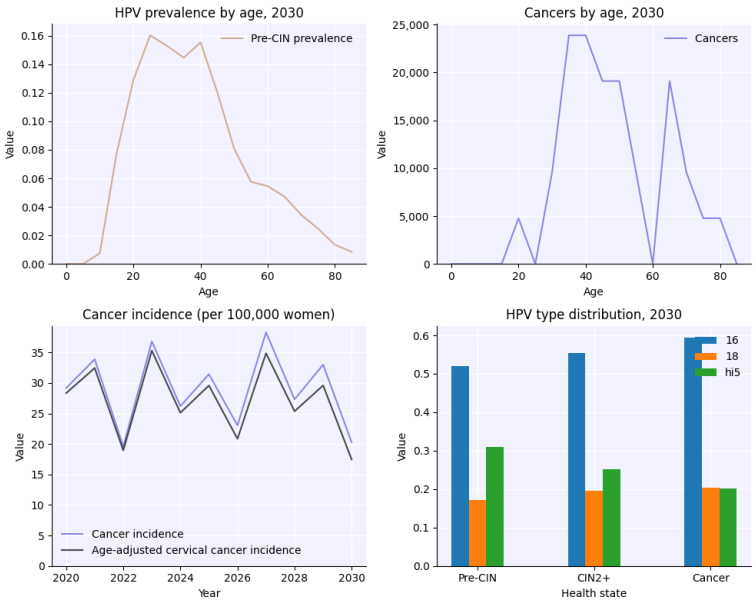


Figure 2: HPVsim - Simulation Results (India)

3.2 HPVadvise - LIMC

HPVadvise is an individual-based transmission-dynamic model of HPV infection and disease and was used to inform previous HPV vaccine policy decisions in the United States and Canada. HPV-ADVISE LMIC has the same basic model structure as HPV-ADVISE, calibrated to India, Vietnam, Uganda, Tanzania, and Benin, using comprehensive demographic, sexual behaviour and epidemiological data available from international databases. This model contains five different integrated modules: Demographic characteristics, Sexual behaviour, Transmission of HPV, Screening and treatment, and Vaccination.

3.2.1 Model Structure

Ten-year-old individuals enter the population at a rate η chosen to balance Indian age-specific death rates. In the simulated population, they are attributed with one of four mutually exclusive levels of sexual activity.

- level 0: has been married to a single partner all of their life
- level 1: has had a partner while they are married.
- level 2: both men and women who choose not to marry.
- level 3: both males who pay for sex and female sex workers.

Based on age and amount of sexual activity, each woman has a certain rate associated with either forming a new partnership if she is single or separating if she is currently in a stable relationship. The model is based on this stochastic process of pair formation and separation. The natural history of infection, sexual behaviour, and the chance of transfer during a sexual act all influence HPV transmission.

Ten-year-old females are assigned a screening behaviour level upon entering the simulated population, which is determined by the time interval between two standard screening tests. The range of screening behaviours is as follows: $S = 0$ is the shortest time between two routine screening tests; $S = 4$ represents never getting screened. Both the visual inspection of the cervix with acetic acid and the Pap test (cytology) are presently the primary screening tests carried out in LMICs. Women are assigned varying probabilities of receiving a diagnosis based on screening technologies and their actual health status.

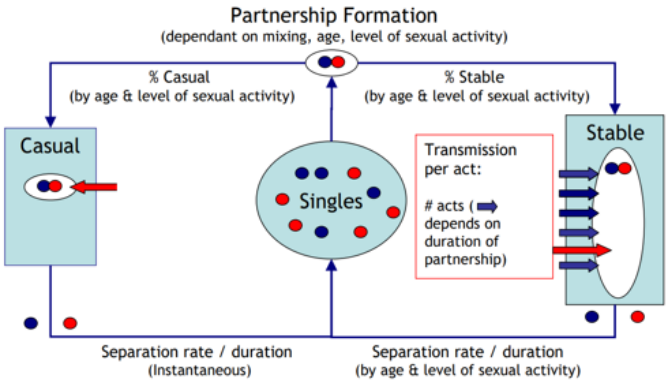


Figure 3: Pair formation in HPVadvise (Ref 6 - fig A1)

3.2.2 Model Parameterization - Calibration

The process of calibration is utilized to find several parameter sets (about fifty) that fit the natural history of HPV data and highly stratified sexual behaviour at the same time. These 50 parameter sets reflect the heterogeneity of sexual behavior and HPV epidemiology within a nation, as well as uncertainty in model parameters. Furthermore, the model is tested to confirm whether the data used for model calibration were not also used for model fit.

CALIBRATION PROCEDURE :

1. Prior distributions are defined for each of the 50 calibrated model parameters

2. Using Latin Hypercube sampling, 1000 distinct combinations of parameter values are extracted from the prior distributions.

3. When parameter sets are included in the posterior parameter sets and the related model predictions fall within pre-established ranges of the natural history, screening data, and observed sexual behaviour, they are considered to be creating a "good fit".

4. These posterior parameter sets are cross-validated by comparing model predictions with observed epidemiological data not used during the fitting procedure

3.2.3 Analysis

The study employed a population-based model to evaluate the effectiveness and cost-effectiveness of various HPV vaccination strategies in four diverse countries: India, Vietnam, Uganda, and Nigeria using Age-standardized cervical cancer incidence and country-specific economic data to analyze strategies differing in the number of vaccinated cohorts, age groups, and populations targeted (girls, boys, or both).

The Effectiveness was measured by calculating the proportion of pre-vaccination cervical cancer cases prevented at equilibrium and estimating the number needed to vaccinate (NNV) to avert one case. The cost-effectiveness analysis compared various strategies against no vaccination, using both incremental cost-effectiveness ratios (ICERs) and country-specific GDP-based thresholds. The analysis adopted a healthcare-payer perspective, considering direct medical costs, and incorporated discounting for future outcomes.

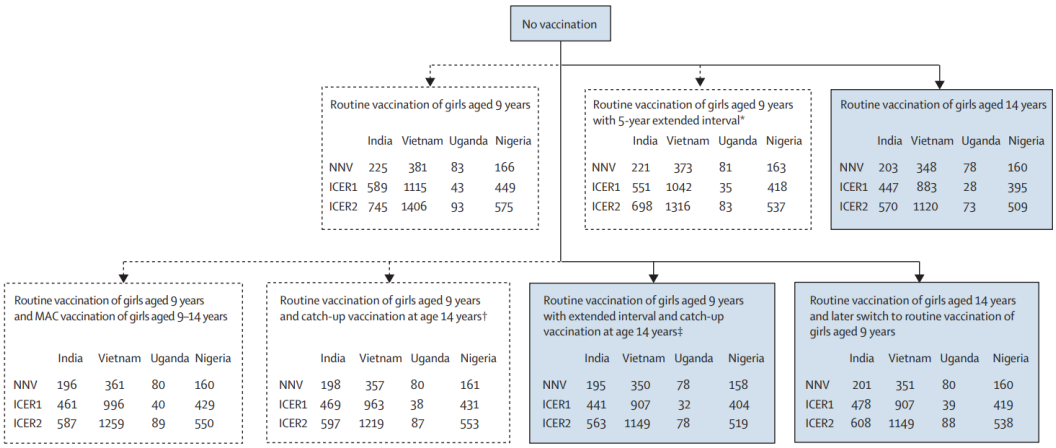


Figure 4: Cost Effectiveness of various Vaccination Strategies (Ref 6 - fig 6)

3.2.4 Results

Regular vaccination of 14-year-olds would result in a slightly lower long-term decrease in age-standardized incidence, but it would accelerate the incidence decrease compared to vaccination of girls at 9 years old because they are closer to becoming sexually active and exposed to HPV. Therefore, as part of a multiple-aged cohort vaccination program, routine vaccination of girls aged 9 years is combined with vaccination of girls aged 9–14 years in the first year of the program to combine the long-term benefits of routine vaccination of girls aged 9 years and the faster incidence decreases observed with routine vaccination of girls aged 14 years.

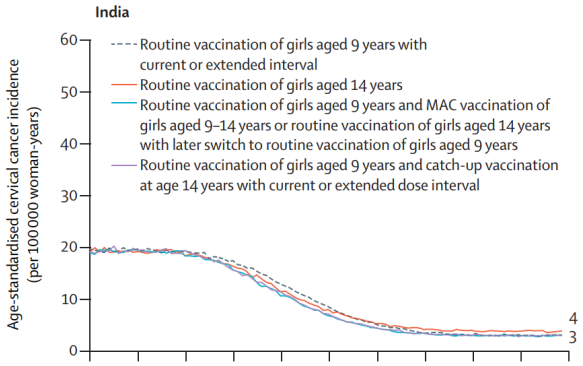


Figure 5: HPVadvise - Simulation results (Ref 6 - fig 3)

Limitations of Agent-Based Models:

Agent-based models depend heavily on accurate and detailed data on individual characteristics, behaviours, and interactions to calibrate and validate the model. For HPV, such data is unavailable and difficult to collect at an individual level, especially in resource-limited settings as in the context of India. While ABMs excel at capturing individual behaviour and interactions, they might also overlook broader epidemiological processes or population-level factors influencing HPV transmission. This can lead to an incomplete understanding of the disease dynamics. Also, ABMs calibrated to specific datasets might overfit those data and fail to generalise to different populations or contexts. This reduces the model's applicability to broader public health planning and policy decisions.

3.3 SICR Compartmental Model

An SICR compartmental model from the research paper " A mathematical model for human papillomavirus and its impact on cervical cancer in India - Praveen Kumar Rajan, Murugesan Kuppusamy, Oluwaseun F. Egbelowo " is considered as a base compartmental model to start with.

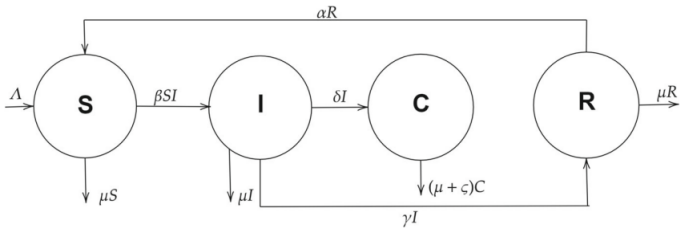


Figure 6: SICR Model (Ref 1 - fig 1)

3.3.1 Model Formulation

The research explores a mathematical model to simulate HPV transmission and cervical cancer development in the Indian female population. Women are categorized into four groups:

- Susceptible (S): Women vulnerable to HPV infection through sexual contact.
- Infected (I): Women carrying HPV but without cervical cancer.
- Cancerous (C): Women diagnosed with cervical cancer due to HPV infection.
- Recovered (R): Women who cleared the HPV infection and are immune.

The total population (N) is the sum of individuals in all compartments (N = S + I + C + R) The following figure illustrates this compartmental model.

The below set of mathematical equations governs the interactions between these groups, considering factors like infection rates, cancer progression, and recovery. This model allows researchers to analyze the dynamics of HPV transmission and cervical cancer spread within the Indian population.

$$\frac{dS}{dT} = \Lambda - \frac{\beta SI}{N} - \mu S + \alpha R \tag{1}$$

$$\frac{dI}{dT} = \frac{\beta SI}{N} - (\alpha + \mu + \gamma)I \tag{2}$$

$$\frac{dC}{dT} = \delta I - (\zeta + \mu)C \tag{3}$$

$$\frac{dN}{dT} = \gamma I - (\alpha + \mu)R \tag{4}$$

Initial conditions: S(0) $\succ$ 0, I(0) $\geq$ 0, C(0) $\geq$ 0, R(0) $\geq$ 0

Λ : recruitment rate of susceptible women in India

α : the rate of lose of people's immunity

β : transmission rate of HPV

γ : recovery rate of HPV infected people without transition to cancer

δ : the rate of cervical cancer development in HPV-infected individuals μ : population's normal death rate

ζ : mortality rate of the population with cervical cancer. All parameters are assumed to be positive.

3.3.2 Parameter Estimation

This mathematical model (SICR) utilizes seven parameters to simulate the dynamics of HPV transmission and cervical cancer development in India. Two parameters, namely the crude mortality rate and recruitment rate, were directly derived from Indian data. The mortality rate was calculated based on the average female lifespan (70.53 years), and the recruitment rate was obtained by multiplying the susceptible population with the inverse of the mortality rate.

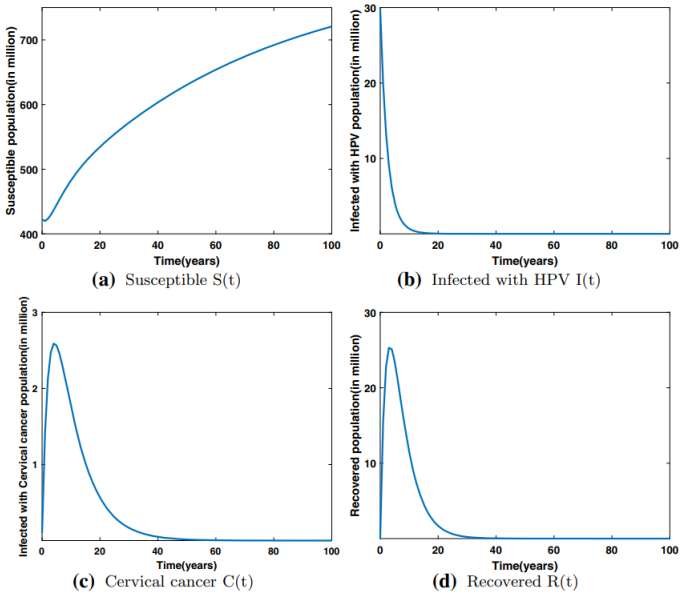
The remaining five parameters were estimated by fitting the model to published data on HPV and cervical cancer cases in India from 2016 to 2020. A nonlinear least squares method was employed in MATLAB software to minimize the deviation between the observed and simulated data points. The resulting root mean squared error (RMSE) close to zero suggests a good fit, indicating the model's capability to accurately predict disease dynamics in the Indian female population. The specific values of the estimated parameters are provided in the accompanying table.

Parameter	Values ($\frac{1}{year}$)
Λ	9.016228
μ	0.0141784
ζ	0.108696
β	0.556
δ	0.05714
γ	0.5
α	0.2

Table 1: Estimated parameter values (Ref 1 - tab 2)

3.3.3 Simulation results

The Indian women population aged above 15 years old in 2016 as the total population is approximately 451.815 million. All the simulations were performed by using the in-built function ODE45 in MATLAB with the initial conditions: S(0) = 421.896 million, I(0) = 29.819 million, C(0) = 0.0991 million, R(0) = 0. We observe the following trends in S,C I and R



(a) Susceptible S(t) **(b)** Infected with HPV I(t)

(c) Cervical cancer C(t) **(d)** Recovered R(t)

Figure 7: SICR Model - Simulation Results (Ref 1 - fig 4)

3.4 Updated Compartmental Model

3.4.1 Model Structure

I have developed the following compartmental model as a part of the Systems Biology Course Project (appendix 1). I would like to include it for the continuity of the whole project and to understand the overall thought process of our model development. In this, we updated the Base SICR model by incorporating screening and treatment. The initial susceptible compartment is split into 2 small compartments: Screened and Not Screened based on the screening rate α. Now the Not screened compartment model is assumed to follow a similar path as the base compartmental model containing Infected and Recovered compartments. We have assumed the HPV test as the primary mode of screening and Thermal Ablation as the treatment method as per the WHO recommendations.

In the Screened compartment we have 2 outcomes Positive and Negative. All patients with positive results in screening are assumed to undergo colposcopy. As these results are not entirely accurate we expect a fraction of patients ${\beta^3}_{FP}$, i.e false positive patients to go to the Susceptible compartment and the remaining True HPV positive patients are split into 3 categories CIN1, CIN2/3, Cancer based on the HPV Screening for Cervical Cancer in Rural India - N Engl J Med 2009. The False-negative patients in screening are assumed to follow a path similar to that of an unscreened patient thus a certain proportion of them join into the Infected Compartment.

$$\beta^1 = \frac{Sensitivity \times prevalence}{Sensitivity \times prevalence + (1 - Specificity) \times (1 - prevalence)} = \mathbf{0.4}$$

$$\beta^2 = [1 - \frac{Specificity \times (1 - prevalence)}{(1 - Sensitivity) \times prevalence + (Specificity) \times (1 - prevalence)}] = \mathbf{0.013}$$

$$\beta^3_{FP} = [1 - \frac{Sensitivity \times prevalence}{Sensitivity \times prevalence + (1 - Specificity) \times (1 - prevalence)}] = \mathbf{0.6}$$

$$\beta^3_{TN} = \frac{Specificity \times (1 - prevalence)}{(1 - Sensitivity) \times prevalence + (Specificity) \times (1 - prevalence)} = \mathbf{0.987}$$

Prevalence = 10%

HPV Test – Sensitivity: 0.9, Specificity: 0.85

Here a few proportion of the Infected population from Not screened stem also join into C3 as generally infected patients get diagnosed in the cancer stage. Patients in CIN2/3 and C3 are assumed to undergo thermal ablation treatment as recommended by the WHO guidelines and the progression and recovery of these compartments is based on the efficacy data of thermal ablation.

The below set of mathematical equations from appendix 1 (pp 6) governs

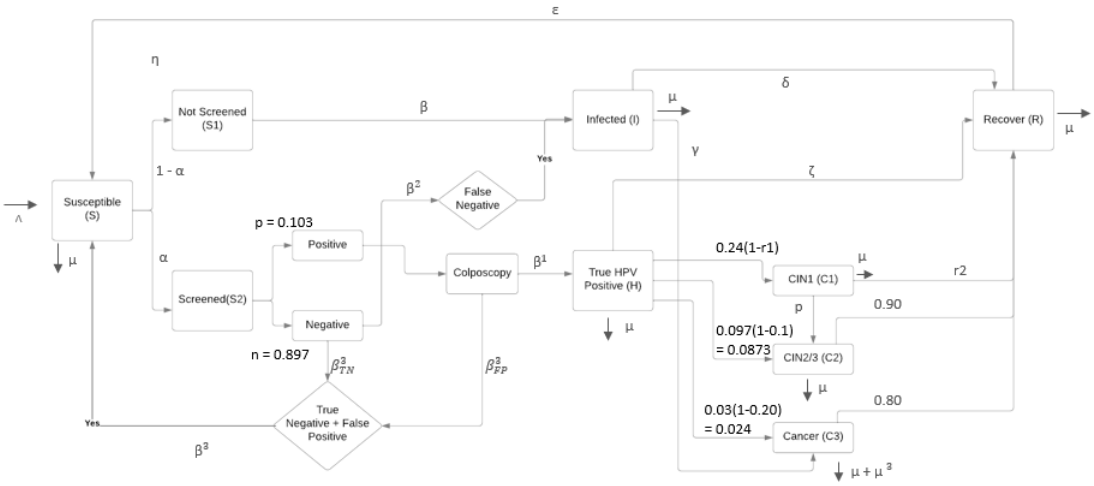


Figure 8: Updated Compartmental Model (appendix 1 fig 4)

the interactions between these groups, considering factors like infection rates, cancer progression, recovery etc

$$\frac{dS}{dT} = \Lambda + \epsilon R - S + \eta S_1 + \beta^3{}_{TN} N + \beta^3{}_{FP} Col - \mu S \tag{1}$$

$$\frac{dS1}{dT} = (1-\alpha)S - \eta S_1 - \frac{\beta S_1 I}{N} \tag{2}$$

$$\frac{dS2}{dT} = \alpha S - S_2 \tag{3}$$

$$\frac{dI}{dT} = \frac{\beta S_1 I}{N} + \beta^2 N - (\delta + \mu + \gamma) I \tag{4}$$

$$\frac{Col}{dT} = pS_2 - (\beta^1 + \beta^3{}_{FP}) Col \tag{5}$$

$$\frac{dN}{dT} = nS_2 - (\beta^2 + \beta^3{}_{TN}) N \tag{6}$$

$$\frac{dH}{dT} = \beta^1 Col - (0.24(1-r_1) + 0.0873 + 0.024 + \zeta + \mu) H \tag{7}$$

$$\frac{dC1}{dT} = (0.24(1 - r_1)H - (r_2 + p + \mu)C1 \tag{8}$$

$$\frac{dC2}{dT} = 0.0873h + pC1 - 0.9C2 - \mu C2 \tag{9}$$

$$\frac{dC3}{dT} = 0.024H - 0.8C3 - (\mu + \mu^3)C3 \tag{10}$$

$$\frac{dR}{dT} = \delta I + \zeta R + r_2 C1 + 0.62C2 + 0.8C3 - (\epsilon + \mu)R \tag{11}$$

We have used the base compartmental model results as the base data available and tried to fine-tune the parameters of the updated model from appendix 1 based on the recovered outcome and obtained the following results using ODEINT for simulation

Parameters	Definition	Value
Λ	recruitment rate of susceptible women	9.016 Mil
α	Screening rate	0.7
β	HPV transmission rate	0.556
γ	rate at HPV progress to develop cervical cancer	0.485484
δ	rate of HPV recovering without cancer	0.5
μ	population's normal death rate	0.014516
μ_3	mortality rate of population after cancer	0.108696
ϵ	rate of becoming susceptible again	0.2
r1	rate of clearance from CIN1 but remaining HPV infected	0.13
r2	rate of clearance from both the CIN1 and infection	0.77
p	rate of progression from CIN1 to CIN2	0.065
β^1	True positive rate from Colposcopy	0.4
β^2	False negative rate during screening	0.013
${\beta^3}_{FP}$	False positive rate from Colposcopy	0.6
${\beta^3}_{TN}$	True negative rate during screening	0.987
η	Rate at which not screened people become susceptible again	0.429484
ζ	Recovery rate of HPV patients	0.6799

Table 2: Initial Parameters (appendix 1 fig 6)

3.4.2 Simulation results

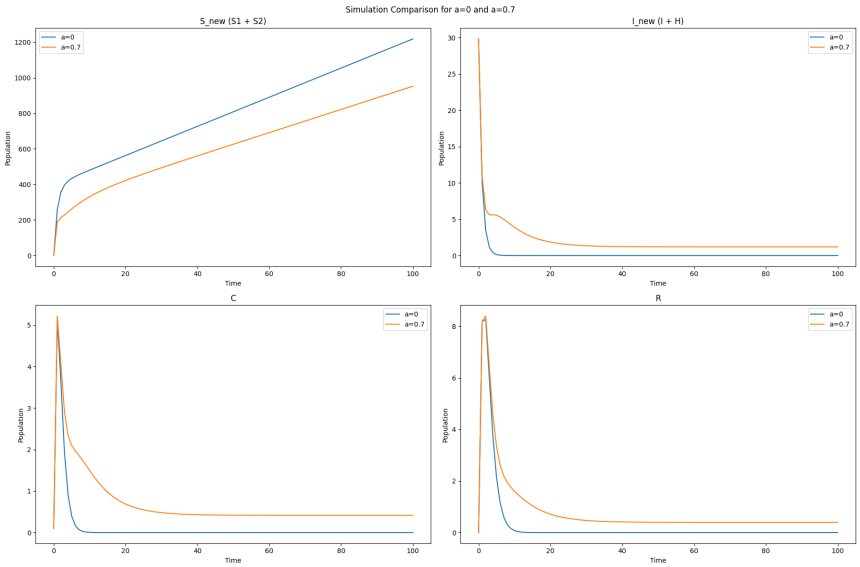


Figure 9: Simulation Results (appendix 1 fig 7)

The model's results present an interesting scenario. The results showed that as the screening percentage (α) increased, the size of the susceptible pool (S) decreased significantly. This indicates that screening effectively identifies and removes individuals from the early stages of HPV infection, thereby limiting the spread of the virus. While screening reduced the susceptible pool (S) as expected, it also predicted a rise in the detected infections (I) initially. This likely reflects effectively identifying existing infections that may have previously gone unnoticed. This initial rise in detected infections reflects the effectiveness of the screening program in uncovering the true burden of HPV in the population.

Surprisingly, we see an increase in the cancer compartment (C) with increasing screening rates. With Screening, Cancers are identified earlier and over time, with continued screening, the number of diagnosed cancers is expected to decrease as the interventions prevent new cancers from developing altogether. So our model might need adjustments to account for lead-time bias. This could involve incorporating factors like cancer progression rates or delays in mortality due to earlier detection. The number of recovered individuals (R) is expected to rise initially with increased screening. This reflects successful treatment or natural clearance of the infection after being identified through screening.

4 Screening and Vaccination Models

4.1 Simplified Screening Model

The current updated model contains 18 parameters and 12 compartments. This complexity presented significant hurdles. Fine-tuning the numerous parameters and ensuring the model's structural identifiability proved challenging. Data collection for such a detailed model also presented difficulties. To address these limitations, we held a meeting with the National Cancer Institute's Division of Cancer Diagnosis and Research (NCDIR). While NCDIR lacked HPV-specific data, they had valuable data on cancer progression, cases, and deaths. This data can be used in fine-tuning parameters related to HPV infection progression. Recognizing the need for a more manageable approach, we opted to simplify the initial screening model. This simplification not only streamlined data collection and parameter estimation but also paved the way for applying optimal control theory to the refined model.

We assumed α per cent of the Susceptible population to be screened (HPV test as the primary mode of screening) and Thermal Ablation as the treatment method as per the WHO recommendations. Now the Not screened population is assumed to follow a similar path as the base compartmental model containing Infected, Cancer and Recovered compartments. Our model classifies the female population at time t denoted by $N(t)$.

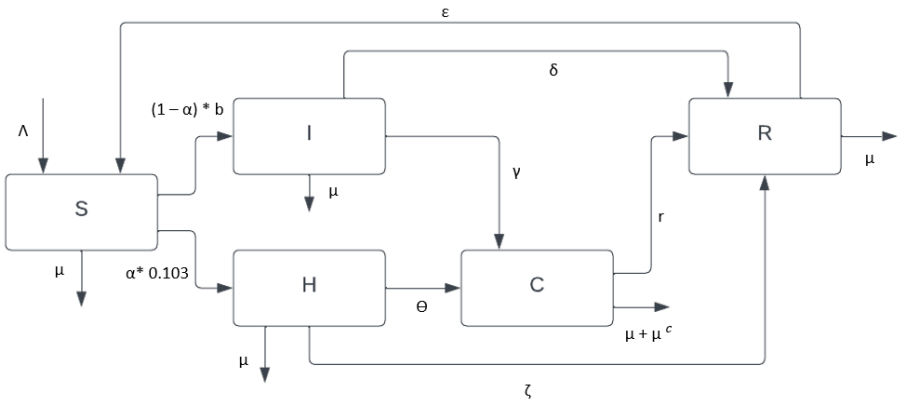


Figure 10: Simplified Screening Compartment Model

The model is governed by the following system of differential equations:

$$N(t) = S(t) + I(t) + H(t) + C(t) + R(t) \tag{1}$$

$$\frac{dS}{dt} = \Lambda + \epsilon R - 0.103\alpha S + \frac{(1-\alpha)\beta SI}{N} - \mu S \tag{2}$$

$$\frac{dI}{dt} = \frac{(1-\alpha)\beta SI}{N} - (\delta + \mu + \gamma)I \tag{3}$$

$$\frac{dH}{dt} = 0.103\alpha S - (\theta + \xi + \mu)H \tag{4}$$

$$\frac{dC}{dt} = \theta H + \gamma I - (r + \mu + \mu^2)C \tag{5}$$

$$\frac{dR}{dt} = \delta I + \xi H + rC - (\epsilon + \mu)R \tag{6}$$

We have used the SICR Model and Updated compartmental model parameters as the source of our parameters

Parameters	Definition	Value
Λ	recruitment rate of susceptible women	9.016 Mil
α	Screening rate	0.8
β	HPV transmission rate	0.556
γ	rate at undetected HPV progress to develop cervical cancer	0.485484
δ	rate of HPV recovering without cancer	0.5
θ	rate of transmission from HPV infected to cancer	0.104
μ	population's normal death rate	0.014516
μ_c	mortality rate of population after cancer	0.108696
ξ	recovery rate of HPV patients	0.68
ϵ	rate of becoming susceptible again	0.2
r	rate of recovery	0.77

Table 3: Parameters

and all the parameters are assumed to be non-negative.

4.1.1 Simulation Results

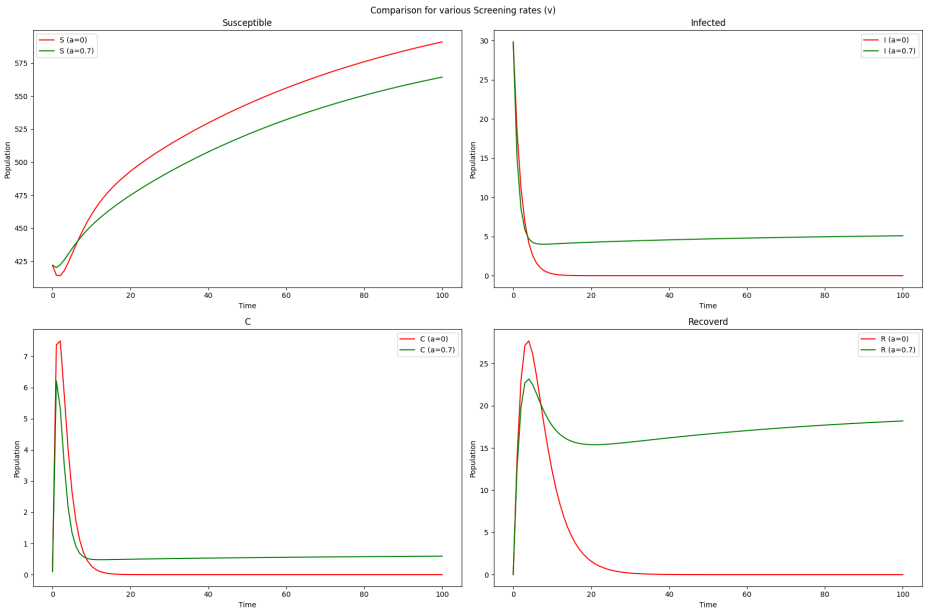


Figure 11: Simplified Compartmental Model - Simulation Results

Our simulations reveal that the model exhibits similar predictive trends to the updated compartmental model. As the screening percentage (α) increases, the susceptible pool (S) diminishes, indicating effective identification and removal of individuals at the early stages of HPV infection. Additionally, a transient rise in detected infections (I) is observed, effectively capturing existing, previously undetected infections.

However, unlike the updated compartmental model, our model predicts a decrease in the cancer compartment (C) at the initial screening stages. This suggests earlier identification and treatment of infections, leading to a reduction in diagnosed cancers. Furthermore, the recovered individuals (R) are expected to rise initially with increased screening, reflecting successful treatment or natural clearance of the infection after detection.

4.1.2 Optimal Control Theory Formulation

Optimal control theory provides a framework for finding the best course of action (control) for a system over time, considering a desired outcome (objective) under any limitations (constraints).

In the context of cervical cancer screening:

- System: The population susceptible to HPV infection
- Controls: The screening strategy, including the rate and type of screening tests offered
- Objective: To minimize the overall cost associated with cancer management. This cost has two components:
  - Treatment cost: The cost of treating diagnosed cancer cases.
  - Screening cost: The cost of implementing the chosen screening program.

Optimal control theory helps to identify the most cost-effective screening strategy. It doesn't simply minimize one cost (e.g., screening) instead it finds the right balance between both screening and treatment costs to achieve the lowest overall cost possible within given constraints.

For our problem we chose a two-pronged method to determine the optimal screening strategy with the assumed compartmental model as the constraint. First, the forward-backward sweep method was utilized. This method involves solving the model's governing ordinary differential equations (ODEs), forward in time. This allows us to predict the future state of the system under various screening scenarios. Subsequently, the method works backwards from the desired endpoint – minimizing the overall cost of treating cancer cases and the cost of the assumed screening schedule that best achieves this goal. Secondly, the fourth-order Runge-Kutta method is used to achieve a higher accuracy. This numerical technique is widely used to solve ODEs. It discretizes the time period into smaller intervals and approximates the solution at each interval. This method's strength lies in its ability to provide highly accurate solutions to the model's equations, which is essential for identifying the optimal screening schedule.

The objective function is

$$min \int_0^T AC(t) + B\alpha^2 S(t)\, dt$$

where

A: Cost corresponding to treatment of Cervical cancer patients

α: Screening rate

B: Cost corresponding to Screening

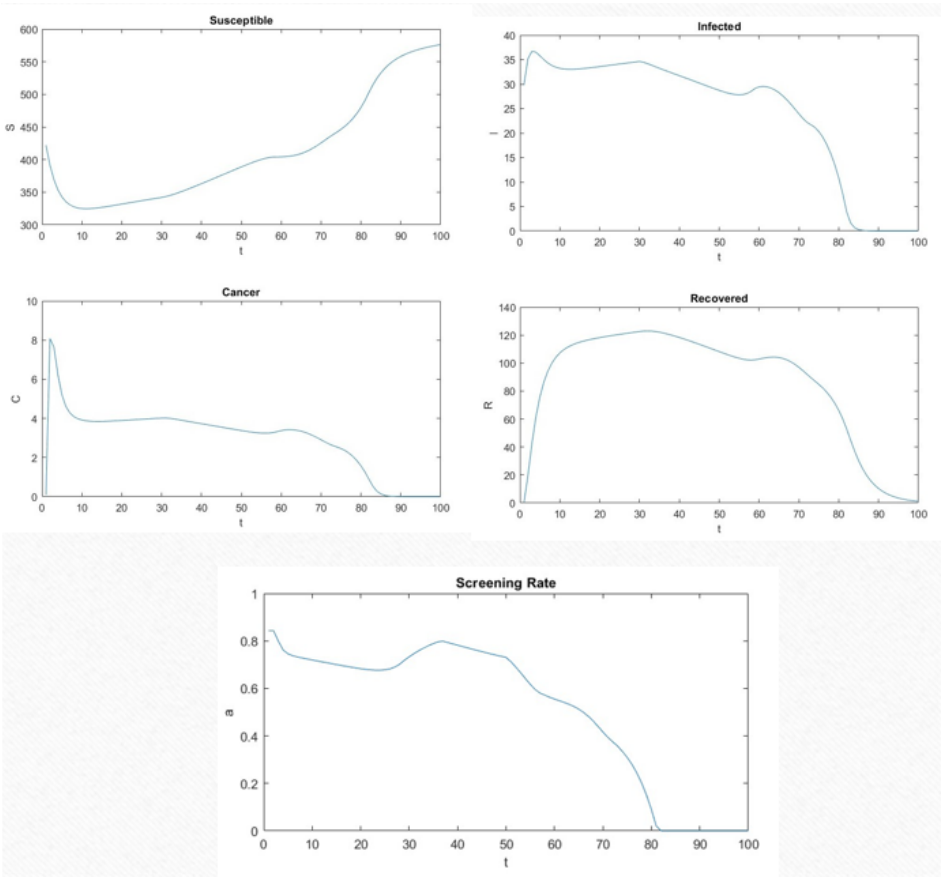


Figure 12: Optimal screening strategy across given time

4.2 Vaccination Model

For vaccination, we considered a simple compartmental model consisting of 4 compartments: Unvaccinated (U), Infected(I), Cancer (C) and Recovered (R). We expect people (15-19 years old) entering the population to be vaccinated at a rate v. Now the not-vaccinated people follow the same trend as base SICR and we expect 95 per cent efficacy[10] for vaccination, so the transmission rate of vaccinated population is calculated accordingly.

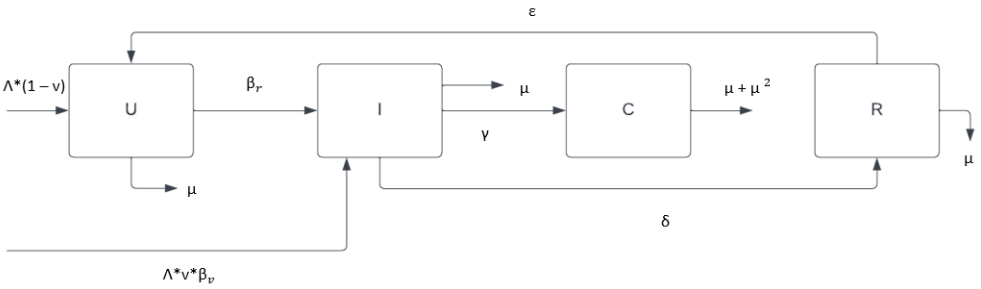


Figure 13: Vaccination Compartmental Model

Parameters	Definition	Value
Λ	recruitment rate of girls aged 15-19 years	0.8554 Mil
v	Vaccination rate	0.8
β_r	Unvaccinated HPV transmission rate	0.556
β_v	Vaccinated HPV transmission rate	0.0278
γ	rate of HPV progression to cervical cancer	0.057
δ	rate of HPV recovering without cancer	0.5
μ	population's normal death rate	0.014516
μ_c	mortality rate of population after cancer	0.108696
ϵ	rate of becoming susceptible again	0.2
r	rate of recovery	0.77

Table 4: Parameters

The model is governed by the following system of differential equations:

$$\frac{dU}{dt} = (1-v)\Lambda + \epsilon R - \frac{\beta UI}{N} - \mu S \tag{1}$$

$$\frac{dI}{dt} = \frac{\beta_r UI}{N} + \frac{\beta_v v\Lambda I}{N} - (\delta + \mu + \gamma)I \tag{2}$$

$$\frac{dC}{dt} = \gamma I - (\mu + \mu^2)C \tag{3}$$

$$\frac{dR}{dt} = \delta I - (\epsilon + \mu)R \tag{4}$$

4.2.1 Simulation Results

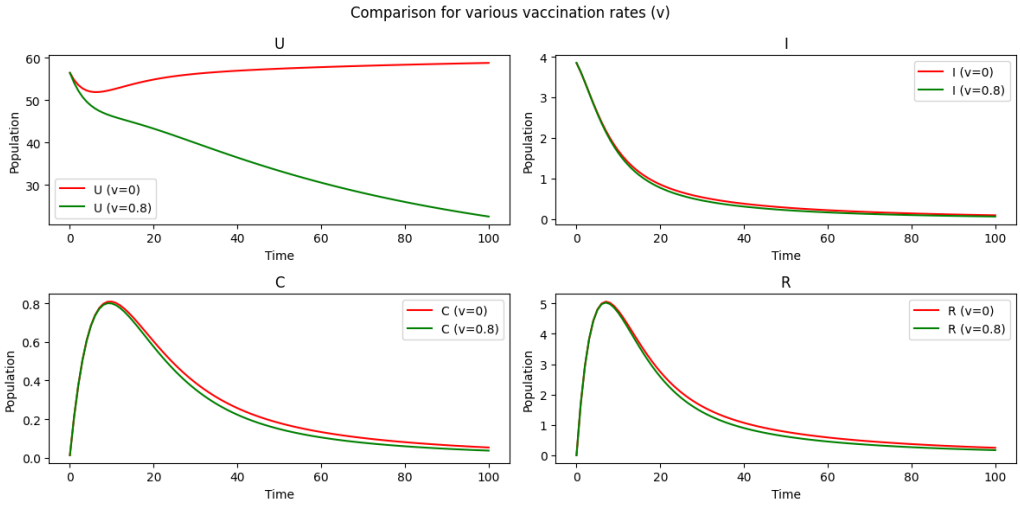


Figure 14: Vaccination Compartmental Model - Simulation results

The simulation results provide a piece of evidence for the effectiveness of the HPV vaccination program in reducing infections, cancer cases, and recovered individuals. Vaccination directly protects the individual. With a 95 per cent efficacy rate, the model reflects a significant decrease in the susceptible pool (U) as more people become vaccinated, leading to a decline in infections (I). However, the benefits extend beyond the vaccinated population. As vaccination rates increase, the overall circulation of the virus within the community drops. This creates a phenomenon known as herd immunity, where even unvaccinated individuals experience a lower risk of infection due to the reduced pool of susceptible individuals. This explains the observed decrease in infections across the entire population, regardless of vaccination status.

With a smaller pool of infected individuals, the number of people who naturally clear the infection or receive successful treatment (R) also declines. In conclusion, the simulation underscores the critical role of HPV vaccination programs in achieving public health goals. By directly protecting individuals and fostering herd immunity, vaccination programs reduce infections, cancer cases, and the overall burden of HPV on the population.

4.2.2 Optimal Control Theory Formulation

In the context of HPV Vaccination:

- System: The population susceptible to HPV infection
- Controls: The rate of Vaccination strategy
- Objective: To minimize the overall cost associated with cancer management (Treatment and Vaccination cost)

The objective function is

$$min \int_0^T AC(t) + Bv^2\Lambda, dt$$

where
A: Cost corresponding to treatment of Cervical cancer patients
v: Vaccination rate
B: Cost corresponding to Vaccination

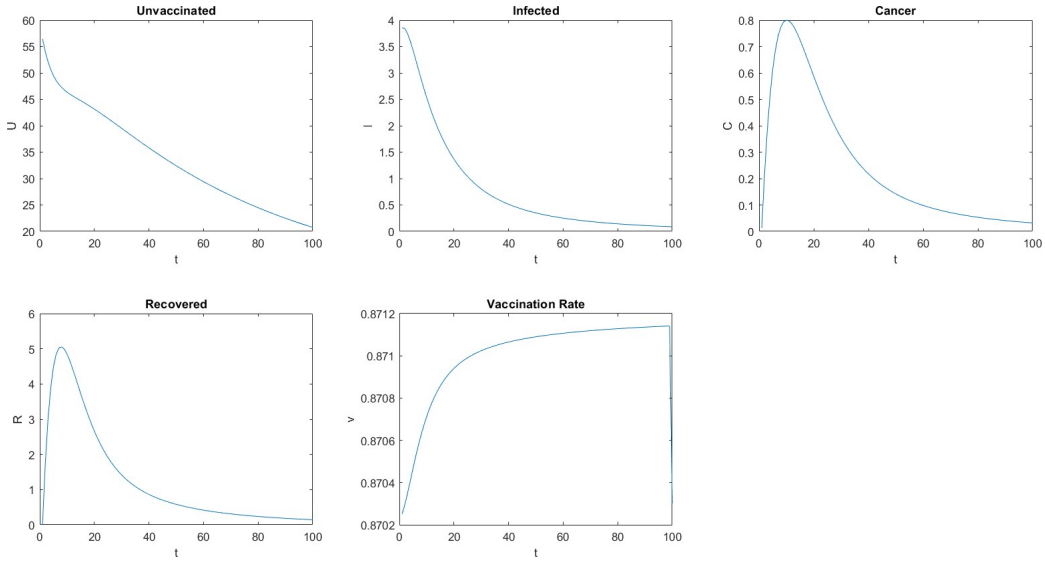


Figure 15: Optimal vaccination strategy across given time

5 Conclusion:

This study explored the potential of compartmental models to evaluate cervical cancer prevention strategies in the context of India. The screening model highlighted the effectiveness of screening in identifying pre-cancerous lesions and reducing the susceptible population. However, the initial rise in detected cancers likely reflects lead-time bias, suggesting the model could benefit from incorporating cancer progression rates. Additionally, the vaccination model demonstrated the overall impact of vaccination in reducing infections, cancer cases, and recovered individuals across the population due to herd immunity.

These findings emphasize the importance of a multi-pronged approach to cervical cancer prevention in India. However, optimizing the allocation of resources for both screening and vaccination programs requires a more nuanced approach. Optimal control theory offers a powerful framework to determine the most effective strategies. By incorporating this theory into a comprehensive model that combines both screening and vaccination, we can identify the optimal screening rates and vaccination schedules to maximize the reduction in cervical cancer cases at minimal cost.

Future research should focus on developing such a comprehensive model, ideally incorporating an age-structured and potentially gender-specific framework. This would allow for a more realistic representation of HPV prevalence and sexual behaviour variations across different demographics. This comprehensive model, coupled with optimal control theory, can guide policymakers in optimizing resource allocation and intervention strategies for maximizing the impact of cervical cancer prevention efforts in India.

6 Future Work:

- Dive deep into the Structural identifiability of various parameters in the compartmental model
- Look into the possibility of developing age-structured compartmental models or gender-specific compartmental models

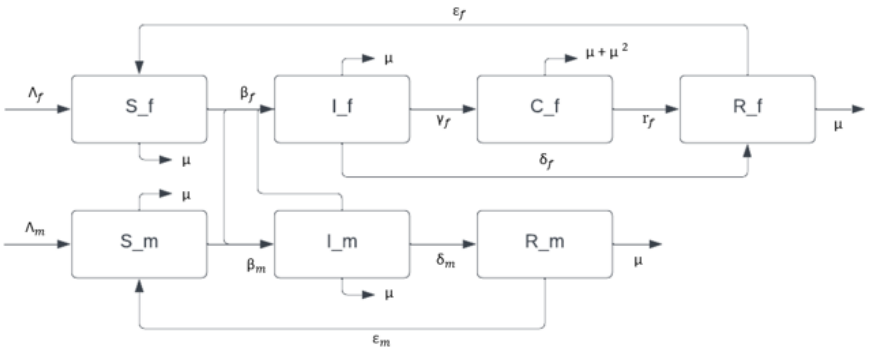


Figure 16: Gender Specific Model - Preliminary Model

- Work on the HPVsim model in the context of India after receiving the required data from the ICMR - NCDIR, Tailoring the model to specific regions or communities within India, allowing for targeted interventions addressing their unique needs and risk factors, ultimately paving the way for localized public health strategies

Data Sharing

Descriptions of the Updated model structure, the parameters included in the model, and the empirical data used for calibration and validation are available in appendix 1 - link

The Supplementary Information, which includes the codes used for the work is given here.

References

1. Rajan, P.K., Kuppusamy, M. & Egbelowo, O.F. A mathematical model for human papillomavirus and its impact on cervical cancer in India. J. Appl. Math. Comput. 69, 753–770 (2023).
https://doi.org/10.1007/s12190-022-01767-2

2. Zhao XL, Liu ZH, Zhao S, Hu SY, Muwonge R, Duan XZ, Du LJ, Su CF, Xiang XE, Zhang X, Pan QJ, Qiao YL, Sankaranarayanan R, Zhao FH, Basu P. Efficacy of point-of-care thermal ablation among high-risk human papillomavirus positive women in China. Int J Cancer. 2021 Mar 15;148(6):1419-1427. doi: 10.1002/ijc.33290. Epub 2020 Sep 16. PMID: 32895912.

3. Sankaranarayanan R, Nene BM, Shastri SS, Jayant K, Muwonge R, Budukh AM, Hingmire S, Malvi SG, Thorat R, Kothari A, Chinoy R, Kelkar R, Kane S, Desai S, Keskar VR, Rajeshwarkar R, Panse N, Dinshaw KA. HPV screening for cervical cancer in rural India. N Engl J Med. 2009 Apr 2;360(14):1385-94. DOI: 10.1056/NEJMoa0808516. PMID: 19339719.

4. HPVsim: Robyn M. Stuart, Jamie A. Cohen, Cliff C. Kerr, Romesh G. Abeysuriya, Marita Zimmermann, Darcy W. Rao, Mariah C. Boudreau, Daniel J. Klein. An agent-based model of HPV transmission and cervical disease.
https://doi.org/10.1101/2023.02.01.23285356

5. An, G., Fitzpatrick, B.G., Christley, S. et al. Optimization and Control of Agent-Based Models in Biology: A Perspective. Bull Math Biol 79, 63–87 (2017).
https://doi.org/10.1007/s11538-016-0225-6

6. lanie Drolet, Jean-François Laprise, Dave Martin, Prof Mark Jit, Élodie Bénard, Guillaume Gingras. Optimal human papillomavirus vaccination strategies to prevent cervical cancer in low-income and middle-income countries in the context of limited resources: a mathematical modelling analysis.
DOI: https://doi.org/10.1016/S1473-3099(20)30860-4

7. Gender and Age Structured population of India -
https://www.populationpyramid.net/india/2016/

8. Human Papilloma Virus and Related Diseases Report - India
https://hpvcentre.net/

9. Jarynowski, Andrzej. (2015). HPV and cervical cancer in Moldova, epidemiological model with intervention cost vs benefit and effectiveness analysis - https://arxiv.org/abs/1504.05383

10. Basu P, Malvi SG, Joshi S, Bhatla N, Muwonge R, Lucas E, Verma Y, Esmy PO, Poli URR, Shah A, Zomawia E, Pimple S, Jayant K, Hingmire S, Chiwate A, Divate U, Vashist S, Mishra G, Jadhav R, Siddiqi M, Sankaran S, Prabhu PR, Kannan TPRA, Varghese R, Shastri SS, Anantharaman D, Gheit T, Tommasino M, Sauvaget C, Pillai MR, Sankaranarayanan R. Vaccine efficacy against persistent human papillomavirus (HPV) 16/18 infection at 10 years after one, two, and three doses of quadrivalent HPV vaccine in girls in India: a multicentre, prospective, cohort study. Lancet Oncol. 2021 Nov;22(11):1518-1529. doi: 10.1016/S1470-2045(21)00453-8. Epub 2021 Oct 8. Erratum in: Lancet Oncol. 2022 Jan;23(1):e16. PMID: 34634254; PMCID: PMC8560643.
https://pubmed.ncbi.nlm.nih.gov/34634254/