Quantitative Models to Design and Evaluate Risk-Specific Screening Strategies for Cervical Cancer Prevention

DISSERTATION

Presented in Partial Fulfillment of the Requirements for the Degree Doctor of Philosophy in the Graduate School of The Ohio State University

By

Adriana Ley-Chavez

Graduate Program in Industrial and Systems Engineering

The Ohio State University

2012

Dissertation Committee:

Professor Julia L. Higle, Advisor

Professor Electra Paskett

Professor Ramteen Sioshansi
Copyright by

Adriana Ley-Chavez

2012
Abstract

Cervical cancer is the second leading cause of female cancer mortality worldwide, and it can be prevented effectively with appropriate screening. Infection with an oncogenic strain of the human papillomavirus (HPV) is a necessary cause for cervical cancer and HPV is a very common sexually transmitted infection. Demographic and behavioral risk factors for HPV infection have been identified, but the current screening guidelines for cervical cancer prevention do not make distinctions related to risk factors in their recommendations. Cost-effectiveness analyses have compared the screening strategies approved by current guidelines, but the effect of incorporating each patient’s risk for infection into the selection of recommended screening strategies has not been studied. We explore the possibility of implementing risk-specific screening strategies and evaluate their cost-effectiveness with respect to the screening strategies currently used. We develop a model of risk that allows us to estimate the probability of oncogenic HPV infection based on a patient’s risk characteristics. We use this to implement risk-specific screening strategies in a simulation model of disease progression and screening, which we use to compare our proposed risk-specific strategies to the traditional screening strategies. We evaluate all strategies in terms of total costs and quality-adjusted life years accumulated in the simulated population. We find risk-specific strategies to be more cost-effective and conclude this to be evidence of the potential benefits of risk-specific screening programs for cervical cancer prevention.
To Michael and my parents
Acknowledgments

I am very grateful for the guidance, support and help from Dr. Julie Higle, Dr. Electra Paskett, and Dr. Ramteen Sioshansi. Additional thanks to Dr. Mack Ruffin, Dr. John McLaughlin, Dr. Marc Posner, and my colleagues at ISE. Special thanks to Dr. Seth Bonder for his generosity through the Bonder Fellowship.
Vita

2004 .................................................... B.S. Industrial Engineering, Universidad de las Américas-Puebla

2004-2006 ............................................ Project Consultant, Kratzer Automation

2007 .................................................... M.S. Industrial Engineering, The Ohio State University

2006 to present ..................................... Graduate Associate, Department of Integrated Systems Engineering, The Ohio State University

Publications


Fields of Study

Major Field: Industrial and Systems Engineering

Specialization in Operations Research

Minor Field: Public Health - Epidemiology

Minor Field: Cognitive Systems Engineering
Table of Contents

Abstract .................................................................................................................................................... ii

Acknowledgments ..................................................................................................................................... iv

Vita .............................................................................................................................................................. v

List of Tables ............................................................................................................................................ ix

List of Figures .......................................................................................................................................... xi

Chapter 1: Introduction ............................................................................................................................... 1

1.1 References for Chapter 1 ....................................................................................................................... 4

Chapter 2: Literature Review ...................................................................................................................... 6

2.1 Quantitative Model-Based Tools for Screening and Treatment Decisions ......................... 6

2.2 Applications of Operations Research to Cervical Cancer Prevention ................................ 22

2.3 References for Chapter 2 ................................................................................................................... 31

Chapter 3: Comprehensive Model to Evaluate Cervical Cancer Screening Strategies ....... 34

3.1 Introduction ........................................................................................................................................... 34

3.2 Methods ............................................................................................................................................... 37

3.2.1 Natural History of Disease ........................................................................................................ 38
3.2.2 Screening Strategies ................................................................. 42
3.2.3 Input Data .................................................................................. 44
3.2.4 Description of the Experiment ................................................... 56
3.3 Results ............................................................................................ 57
3.3.1 Validation ....................................................................................... 57
3.3.2 Results of Homogeneous Populations ......................................... 59
3.4 Extension to Mixed-Risk Population ................................................ 62
3.4.1 Results of Mixed-Risk Population ................................................ 63
3.5 Sensitivity Analysis .......................................................................... 64
3.5.1 Model Including Vaccine .............................................................. 64
3.5.2 Model Including Vaccine Results ................................................ 67
3.5.3 Model Including Adherence .......................................................... 69
3.5.4 Model Including Adherence Results ............................................. 70
3.6 Discussion ......................................................................................... 72
3.7 References for Chapter 3 ................................................................. 77

Chapter 4: Logistic Regression Analysis to Predict Oncogenic HPV Infection Probability
.............................................................................................................. 82
4.1 Introduction ....................................................................................... 82
4.2 Methods ............................................................................................ 83
List of Tables

Table 3.1  Prevalence of Oncogenic HPV Infection (Kahn et al. 2007).........................44
Table 3.2  Estimated Age-Specific 6-Month Transition Probabilities from Normal to
Infected State (P_a(I))........................................................................................................45
Table 3.3  Colposcopy Result Conditioned on Lesion State (Kulasingam et al. 2006).....47
Table 3.4  Parameters Related to Pap Results Found in the Literature ......................48
Table 3.5  Probability of Positive Oncogenic HPV DNA Result..................................49
Table 3.6  Screening Costs, USD (Melnikow et al. 2010)..............................................50
Table 3.7  Parameters Used to Estimate the Cost of an Invasive Cervical Cancer Case ..50
Table 3.8  Input Data Selected with NLP ........................................................................54
Table 3.9  Data on Probabilities of Lesion Derived Using Probability Computations and
Input Data Selected with NLP ..........................................................................................55
Table 3.10  Accuracy of Pap Results ...............................................................................56
Table 3.11  Cost-Effectiveness Analysis for Populations with a Constant P(I)=0.06 .....61
Table 3.12  Preferred Screening Strategies for each Risk Level .................................61
Table 3.13  Cost-Effectiveness Analysis for a Mixed-Risk Population .......................63
Table 3.14  Screening Recommendations, by Age .......................................................64
Table 3.15  Cost-Effectiveness Analysis for a Population with 100% Vaccination........67
Table 3.16  Screening Recommendations for a 100% Vaccinated Population, by Age....67
Table 3.17  Cost-Effectiveness Analysis for a Population with 33% Vaccination .........68
Table 3.18  Estimated Probability Distributions of Adherence Conditioned on $P_a(I)$......70

Table 3.19  Preferred Screening Strategies for Homogeneous Populations with Imperfect Adherence ..............................................................................................................71

Table 3.20  Cost-Effectiveness Analysis for a Population with Imperfect Adherence ....71

Table 4.1  Univariate Association Between HPV and Selected Demographic and Behavioral Variables ........................................................................................................87

Table 4.2  Predictor Variables Included in the Final Multivariate Model .....................88

Table 4.3  Stata Output from a Hosmer-Lemeshow Goodness-of-fit Test .........................94

Table 4.4  Groups Constructed According to the Description of the Archer & Lemeshow F-adjusted Mean Residual Goodness-of-fit Test ...........................................95

Table 5.1  Probability Distribution of Age of First Marriage ........................................104

Table 5.2  Probability Distribution of First Marriage Duration......................................105

Table 5.3  Probability Distribution of Duration of Divorce ...........................................106

Table 5.4  Probability Distribution of Second Marriage Duration ................................106

Table 5.5  Yearly Probability Distribution of New Partners in Unmarried Women ......108

Table 5.6  Yearly Probability of Entering and Exiting Poverty.......................................108

Table 5.7  Screening Strategies Used for Each Risk Level in Risk-Specific Strategies..111

Table 5.8  Cost-Effectiveness Analysis of Screening Strategies ....................................112

Table 5.9  Cost-Effectiveness Analysis for a Population with 100% Vaccination.............114

Table 5.10 Cost-Effectiveness Analysis for a Population with 33% Vaccination.........115

Table 5.11 Cost-Effectiveness Analysis for a Population with Imperfect Adherence ....116

Table 5.12 Summary of Results ..................................................................................116
List of Figures

Figure 3.1 Model of Disease Progression

Figure 3.2 Distribution of Diagnoses by Age, Comparison of Our Results to SEER Data

Figure 3.3 Performance of Screening Strategies in Terms of Total Costs and QALYs, for a Population with a Constant P(I)=0.06

Figure 5.1 Distribution of Diagnoses by Age, Comparison of Our Results to SEER Data
Chapter 1: Introduction

Cervical cancer is the second leading cause of female cancer mortality worldwide with 250,000 deaths each year (World Health Organization 2011). The National Cancer Institute (2011) estimates that 12,710 cases of invasive cervical cancer were diagnosed in the US in 2011. Precancerous cervical lesions progress very slowly, can regress, and are generally asymptomatic. Therefore cervical cancer can be prevented effectively with appropriate screening, and making decisions about preventive screening is important and challenging.

Infection with an oncogenic strain of the human papillomavirus (HPV) is a necessary cause for cervical cancer1 (Schiffman and Castle 2003, Muñoz 2000), and HPV is the most common STI in the US (Centers for Disease Control and Prevention 2011). There are a number of known demographic and behavioral risk factors for HPV infection, such as the number of lifetime and recent sex partners, marital status, and a past history of STIs (Kahn et al. 2007, Dunne et al. 2007, Sasagawa et al. 2005, Wright et al. 2004, Sellors et al. 2003, Winer et al. 2003, Moscicki et al. 2001). This, and research on adherence to risk-appropriate cervical cancer screening (Paskett et al. 2010) suggest that risk-differentiated screening strategies for cervical cancer prevention may be superior alternatives to the current invariant strategies.

Current screening guidelines (Smith et al. 2010, American College of Obstetricians and Gynecologists 2009, Wright et al. 2007) make some distinctions in their recommendations based on women’s age, but none based on their behavioral and demographic risk factors. The effect of

---

1 One known exception is mother’s exposure to Diethylstilbestrol (DES) during pregnancy. DES was prescribed to pregnant women between 1940 and 1971, and DES-exposed daughters are estimated to be 2.2 times more likely to have abnormal pre-cancerous cell changes in the cervix than unexposed women.
incorporating each patient’s risk for infection (based on her self-reported behavior) into the selection of recommended screening strategies has not been studied. In the following chapters we explore the possibility of implementing risk-specific screening strategies and evaluate their cost-effectiveness with respect to the screening strategies currently used. The chapters that follow include a literature review (Chapter 2) and a set of three related papers.

In the first paper (Chapter 3) we develop a simulation model of disease progression from infection with an oncogenic HPV strain to the development of precancerous lesions and invasive cervical cancer. We incorporate screening programs which consist of different screening tests and account for the imperfect accuracy of the tests. We find some discrepancies in the input data and data that are unavailable, and we use mathematical modeling to make reasonable estimations of these data. For this initial exploration of the potential of risk-specific screening, we evaluate the performance of screening strategies in homogeneous populations with different levels of risk. We find that the screening strategies that are most cost-effective vary according to the level of risk of the population, and consider this to be evidence that risk-specific screening can be beneficial. We design risk-specific strategies that are found to be cost-effective when implemented in a population in which each woman’s risk varies over time as a function of her age.

Having concluded that risk-specific screening strategies can be beneficial for the prevention of cervical cancer, the second paper (Chapter 4) reports an analysis of data from the National Health and Nutrition Examination Survey (NHANES) and a logistic regression model that relates a woman’s demographic and behavioral profile to her risk for acquiring an oncogenic HPV infection. As a result, we have a logistic regression equation that can be used to calculate the probability of oncogenic HPV infection, given the levels of four risk variables.
The third paper (Chapter 5) brings the first two papers together and incorporates US census data to study the potential cost-effectiveness of risk-specific strategies for a spectrum of women across the population as they progress through life. We simulate changes throughout each woman’s life in age, marital status, partner acquisition and poverty status and calculate her probability of infection based on these changing risk factors. We evaluate risk-specific strategies in this more realistic representation of the US population of women, and find them to be more cost-effective than the traditional screening strategies.

We conclude with some comments on the implications of our findings.
1.1 References for Chapter 1


Chapter 2: Literature Review

We use operations research (OR) methods to design and evaluate the performance of risk-specific screening strategies to prevent cervical cancer (CC). OR methods have been applied to public health problems such as the modeling of disease progression and epidemics, cost-effectiveness studies to compare screening and treatment alternatives, and calculating patients’ utility for specific alternatives. Here we review research done to solve these types of public health problems with the help of OR tools. We focus on the use of Markov models, simulation, influence diagrams and other decision analysis tools. We first review applications of OR to screening and treatment decisions of various diseases, followed by a review of applications of OR to cervical cancer prevention.

2.1 Quantitative Model-Based Tools for Screening and Treatment Decisions

The following are examples of quantitative decision models for screening and treatment decisions, primarily applied to cancer but also to other diseases. In general, the process begins by modeling the natural progression of the disease. Alternative screening or treatment strategies are evaluated by including their effects in the model and comparing their outcomes.

The effects of the different strategies often consist of different transition probabilities into health states such as death and disease survival, and can also result in the possibility of transitioning into
new states such as treatment and surveillance (a state of being monitored more closely than with regular screening).

Common performance measures are the costs and the number of quality-adjusted life years (QALYs) accumulated with each alternative being evaluated. QALYs were developed as a measure of health effectiveness for cost-effectiveness analysis (Weinstein et al. 2009) and consist of weights that are assigned to health states that range from zero to one, with one representing a state of perfect health. As an example, a person who lives 80 years in perfect health would accumulate 80 QALYs in her lifetime. If instead she spent her final 10 years in a health state considered to have a QALY weight of 0.50, she would accumulate 75 QALYs in her lifetime. In using QALYs, an assumption is made that a health state that is more desirable is more valuable (Weinstein et al. 2009). To determine how people value possible health states, preferences are elicited from a representative sample of the affected population (Weinstein et al. 2009). The Cost-Effectiveness Analysis Registry (2011) has collected data on QALY weights elicited and used which correspond to health states associated with many diseases and treatments.

Costs and QALYs are commonly used to calculate incremental cost-effectiveness ratios (ICER), which serve to compare the performance of strategies in terms of both of these measures. The ICER between two strategies is the cost per each additional QALY gained by using the more expensive and more effective strategy of the two. To calculate the ICER between the two strategies, strategies are ordered in order of increasing costs and the additional costs of a strategy are divided by the additional QALYs compared with the previous, less costly strategy. Strategies that are both more costly and less effective in terms of QALYs than another strategy are considered dominated. In the US, more costly but more effective strategies costing more than $50,000 per QALY gained (having an ICER of more than $50,000) are commonly not considered
cost-effective (Eichler et al. 2004, Kulasingam and Myers 2003). Others (Goldhaber-Fiebert et al. 2008) also consider strategies that fall within the range of $50,000-$100,000 per QALY gained to be cost-effective. Following this approach, the strategy considered most cost-effective is the one that results in the highest number of QALYs, while paying no more per QALY than the threshold established (e.g. $50,000).

The data commonly required for quantitative decision models for screening or treatment are: disease prevalence, obtained from cross-sectional studies; disease incidence and progression rates, obtained from cohort (longitudinal) studies; possible outcomes and their probabilities; societal and patient costs, and QALYs. Cancer is a primary health care area that has been studied in this fashion.

Chen et al. (1998) develop Markov models to evaluate routine breast cancer screening programs based on different screening frequencies, and compare the results with those of a randomized trial. Screening programs are carried out in real-life practice settings such as clinics. Randomized (controlled) trials are the gold standard for evaluating interventions such as treatments and screening tests. Therefore, comparing the results of a model of a screening program with those of a randomized trial can give an indication of the quality of the model. The states modeled include: no detectable disease, various preclinical and clinical phases of cancer, and various tumor sizes. The application of the model to a screening program resulted in similar estimates of tumor progression as has been found in a randomized trial, suggesting that the model can accurately represent the effects of screening.

More recently, Lee et al. (2008) consider three annual breast cancer screening strategies, mammography, magnetic resonance imaging (MRI), and a combination of mammography and
MRI, and compare them with traditional clinical surveillance (close monitoring without mammography or MRI as primary screening tests). A simulated population consisting of initially asymptomatic women with high-risk gene mutations is considered. A Markov model was developed and analyzed as a Monte Carlo simulation. The model consists of three modules: breast cancer development and detection, treatment and follow-up, and screening. The health states modeled are: Healthy, four progressive stages of Undiagnosed Breast Cancer, Treated Breast Cancer, and death by breast cancer or other cause. In the model of the natural history of the disease, women enter the Treated Breast Cancer state after presenting symptoms. When considering screening in the model, women in the Undiagnosed Breast Cancer state that do not present symptoms can also enter the Treated Breast Cancer state if the cancer is detected by screening. The probability of detection by screening depends on the patient’s cancer stage (more developed tumors are more likely to be detected). The model is populated with a hypothetical cohort and run for a lifetime horizon. The primary outcome measured is absolute life expectancy gained by each screening strategy. The results indicate that an annual strategy of combined mammography and MRI increases the patients' life expectancy the most, though the average difference in life years between this alternative and the ‘worst’ one (clinical surveillance) is extremely small, approximately two years. A comparison of the strategies by monetary cost was not carried out in this study, and would probably be useful in discriminating between strategies.

Delco and Sonnenberg (2000) evaluate the use of colonoscopy to screen for colorectal cancer. Two Markov processes are developed, one without screening and one with biannual colonoscopy. The model without screening considers only the health states: No Cancer, Colorectal Cancer, Proctocolectomy (surgical removal of the rectum and colon) and Death (these last two being absorbing states). The model with screening considers the additional health state Dysplasia (pre-cancerous cell formation), which if identified with colonoscopy leads to a transition into the state
Dysplasia Surveillance (close monitoring) and to a greater probability of Proctocolectomy (instead of Death). In reality a patient can be in a health state of dysplasia independently of the existence of a screening program, but in this case the Dysplasia state is only included in the screening model, in which it can be observed through colonoscopy. However, not considering a state of Dysplasia in the model of the natural history of the disease is inaccurate, since patients do not transition from a state of no cancer to a state of cancer directly. Additionally, it is unlikely that such a process would have the Markov property. A hypothetical cohort of 50 year old patients with ulcerative colitis is considered. The benefit of life years saved and the costs saved from the terminal care of patients dying from colorectal cancer is compared to the costs of prevention with colonoscopy screening. The results do not indicate that a biannual colonoscopy screening strategy is more cost-effective than not screening patients.

Obuchowski and Lieber (2008) use a Markov model to study the effects of misclassification in randomized clinical trials (RCTs) of screening tests. Two common types of misclassification occur in RCTs: over-diagnosis, when the subject’s event is wrongfully attributed to the target disease, and under-diagnosis, when a subject experiences an event caused by the target disease, but an alternative diagnosis is given. A Markov model is used to simulate the natural history of competing diseases and the effect of screening. The model distinguishes between disease-free, symptomatic disease, and undetectable disease states before and after a critical point at which detection through screening would be beneficial. In the model, if the screening test detects an abnormality, early treatment is begun. The outcomes measured are disease-specific mortality and disease-specific symptoms. In the baseline model, in which the RCT is assumed to have perfect classification, screening results in a reduction in symptoms and mortality, as compared to a situation without screening. When different amounts of misclassification are introduced and their effects on the study’s results are examined, a set of conclusions are drawn regarding the
conditions under which misclassification can significantly alter a clinical trial’s results. These findings are relevant for investigators designing and conducting RCTs.

We propose to evaluate risk-specific screening strategies for CC screening, which has not yet been studied. The following are examples of studies that evaluate screening programs for other diseases, considering each patient’s specific risk level.

Brandeau et al. (1991) develop a dynamic control model to evaluate human immunodeficiency virus (HIV) screening policies. The model includes disease progression, screening and other interventions (i.e. educational), and behavioral changes in response to the interventions. The population in the model is divided into subgroups with differing risks for infection, distinguished by factors such as age, gender, and risky behavior practice. The transitions in the model represent the members of the population transitioning over time through disease states and risk levels. Various types of interventions impact the transition rates between states. The model evaluates screening programs by quantifying the following outcomes: costs, number of HIV-infected newborns, HIV incidence, number of deaths from HIV, HIV prevalence, fraction of people infected with HIV that were identified by screening while asymptomatic, and number of false positive screening results. The screening programs evaluated are: no screening, one-time screening of high-risk women, ongoing screening of all pregnant women, and ongoing screening of high-risk pregnant women. The model is validated by comparing its predictions of acquired immune deficiency syndrome (AIDS) incidence with the existing figures and projections of the California Department of Health Services. The results suggest that no screening and ongoing screening of all pregnant women do not make economic sense. However, screening high-risk women (both one-time and ongoing) results in overall savings; the costs of screening are much lower than the costs of medical care for HIV-infected infants.
Ragnarson et al. (2001) use a Markov model to evaluate the cost-effectiveness of intensified prevention programs in diabetic patients with different risks for foot ulcers and amputations. Four patient-specific risk categories and three age groups are considered. A Markov model is used to simulate hypothetical cohorts for each of the risk and age groups. The model consists of eight health states: no ulcer, three foot ulcer states, three outcome states and death. In the simulations, the current level of prevention in Sweden for each risk group is compared with an intensified strategy defined as optimal patient prevention according to recently published international guidelines. The pathway of possible transitions between health states is identical for present and optimal prevention, but with different transition probabilities. The outcomes measured are costs and QALYs. The results indicate that the intensified prevention strategy is cost effective for patients in the three highest risk groups, not so for the lowest risk patients.

Col et al. (1997) analyze the effect of hormone replacement therapy (HRT) on the life expectancy of postmenopausal women with different risk profiles for breast cancer, coronary heart disease (CHD) and hip fracture. These diseases are relevant because the risk (probability) of breast cancer is reported to be increased by HRT (although the association between hormone replacement therapy and breast cancer is controversial, most well-designed studies examining long-term estrogen use have found a small increase in risk, including the Nurses' Health Study), while the risks of CHD and hip fracture are reported to be decreased by HRT. A Markov model is developed to project the risks and benefits of HRT. The natural history of two hypothetical cohorts of postmenopausal women is simulated, with one of the cohorts receiving HRT. The model includes disease states for CHD, breast cancer, hip fracture, endometrial cancer, and combinations of these. Once a woman transitions into a particular disease state, she is then at risk of dying from that disease and from other demographic-related causes. Cohorts with different risk
profiles for the diseases are simulated, and the outcomes of life expectancy and disease incidence are quantified. The results suggest that the only women not expected to live longer with the use of HRT than without the use of it are those at low to lowest risk for CHD and who also have an increased risk of breast cancer.

Perreault et al. (2005) construct a model similar to that of Col et al. (1997), which also considers that different patients are at different levels of risk for the various diseases believed to be influenced by HRT. A multi-disease model based on patient-specific risk profiles is used to assess the benefits and risks of HRT for post-menopausal women. A Markov model is constructed to estimate the life expectancy and years of life saved. In the model, women can develop a number of different diseases over time, with the transition probabilities depending on their specific risk factors and age. After developing a disease, women transition to a new health state in which they may die from that disease and are still at risk for developing the other diseases considered in the model. The years of life saved by HRT are calculated as the difference in life expectancy between women being treated and women not being treated. The predictions from the model regarding the effects of HRT on disease incidence were cross-checked with non-source data, and concur within a 95% confidence interval with results from clinical trials which have studied disease incidence in women using HRT and in women not using it. This indicates that the model is a valid predictor of the consequences of HRT use, considering the age and risk level of the patient.

Brothers et al. (2004) perform a randomized trial comparing recommendations obtained applying a Markov decision analysis model to surgeons’ recommendations, with the purpose of finding whether knowledge of the results of the decision analysis would affect the surgeons’ final decision regarding what treatment to use. The trial is performed in 206 patients with symptomatic lower extremity arterial disease. The treatments considered are: bypass surgery, angioplasty,
amputation and medical management. A decision tree with four Markov nodes (Markov models within a decision tree), one for each of the treatment alternatives, is constructed. In each Markov model, the patient can transition to post-treatment states (i.e. treatment success, failure, healing, improvement, death) specific to each treatment alternative. Transition probabilities for the various outcomes are estimated for each patient by the attending surgeon. Patient-specific utility values are collected from the patients upon enrollment. The outcome measure used by the decision tree to evaluate the alternatives is QALYs. Patients are randomized for their attending surgeon to be made aware (Surgeon Aware, SA) or remain unaware (Surgeon Unaware, SU) of the results of the decision model. The surgeons’ initial plan differed significantly from the treatment recommended by the model, only agreeing 55% of the times. Whether the surgeons’ final decision was influenced by knowing the model’s recommendation was measured by comparing the level of agreement between the surgeon’s initial and final plans in both groups (SA and SU). There was no difference in the level of agreement found; implying that in this case the surgeons’ awareness of the decision model’s recommendation did not impact their final decision. However, there was a tendency to use bypass surgery less and medical therapy more among the patients for whom the surgeon was aware of the model’s results.

In addition to Markov models, other quantitative modeling tools that have been applied to screening and treatment decisions are influence diagrams. Influence diagrams are compact graphical and mathematical representations useful for modeling decision problems with uncertainty and probabilistic dependence among variables. An influence diagram is composed of decision nodes, uncertain event nodes, a value node and arcs joining the nodes. Within the decision nodes are the alternatives available to choose from. Uncertain event nodes represent probability distributions, conditioned on the probability distributions of any other uncertain event nodes that have an arc going into them; and on the alternative chosen in any decision nodes that
have an arc going into them. The value node contains a function of the costs and benefits accumulated in the nodes that have an arc going into it, and their probabilities of occurrence. The following are examples of health applications that use influence diagrams to model treatment decisions.

Gomez et al. (2007) develop an influence diagram-based decision support tool to make decisions about the treatment of neonatal jaundice. Neonatal jaundice is a common disease in newborn babies that consists of having pathologically high levels of bilirubin, which can cause irreversible brain damage. Since most babies have higher levels of bilirubin than adults, it is challenging to distinguish between the normally high levels of bilirubin and the pathologically high levels. Current jaundice guidelines are based on a few risk factors, but there is no consensus about treatment strategies and the timing of treatments is not clearly stated, leaving physicians to make decisions based on their own judgment. The influence diagram model build by Gomez et al. (2007) is based on the input of expert physicians, the parents’ opinions, and costs. A decision support software tool that allows physicians to modify and run the model was developed. The decisions modeled are whether to admit the baby to the hospital and begin treatment, and what treatment to use. The model considers variables that describe the mother’s demographic and social characteristics, the newborn’s characteristics such as birth weight and age in hours, the preliminary test results and delivery data. There are also variables that monitor the patient evolution after treatment has begun. All of these variables affect the treatment decisions, which is represented by arcs going into the decision nodes in the influence diagram. All the arcs among chance nodes reflect conditional probabilities that were elicited from three expert physicians. The value node represents the utility function that models the decision maker’s preferences. The variables that compose the utility function in this case refer to costs and injuries. Preferences were elicited in the form of utilities from physicians and parents by a combination of the probability
equivalent and the certainty equivalent methods. In order to make the model reusable by physicians, the decision support tool IctNeo was developed. Its use has been implemented at a public general hospital in Madrid, Spain and has resulted in a decrease over the past 3 years of the most aggressive type of treatment (exchange blood transfusions) from 4 or 5 per year to zero. The system has also been found (by comparing system recommendations to real cases) to recommend earlier discharge from the hospital than was previously practiced by non-expert doctors; expert doctors tend to administer earlier discharge, as does the system. In this way the use of this system has helped neonatologists in situations in which their lack of experience may lead to unnecessary treatments.

Meyer et al. (2004) use an influence diagram to rank alternative intensity-modulated radiation therapy plans for the treatment of prostate cancer, considering relevant patient characteristics and data from clinical trials. Each alternative treatment plan has probabilities of cure and complication associated with it which are modeled as uncertain events. Other nodes in the influence diagram are grade of tumor, spread of disease and patient age. Some of the conditional probabilities among nodes are obtained from results of clinical trials, others from experts’ judgments. The value node is a function of outcomes such as whether the patient survives disease-free and the quality of life after treatment, the probability of occurrence of each outcome, and the preferences expressed by an expert about the outcomes (it is not specified how these preferences are elicited and measured). The influence diagram is used to compare six plans by calculating their values. The alternative plans can then be compared by the decision maker, who can view and compare not only their overall value, but also their specific values for each of the attributes considered, such as control of disease, complications, life quality and cost.
Lee et al. (2006) use an influence diagram to estimate the probability of incorrect breast cancer staging (estimation of the extent of the cancer or cancer stage) and treatment prescription. The primary treatment for breast cancer is surgery, and is often followed by radiation therapy (RT). RT is most beneficial when the level of RT used is adequately matched to the actual cancer stage of the patient. A simulation model is created that follows a hypothetical cohort of patients. For each patient, the simulation model assigns a “true” cancer stage and results for the different diagnostic tests (based on the estimated accuracy of the diagnostic tests and on the existing guidelines about which tests to perform). The combination of evidence from the different diagnostic procedures is modeled using an influence diagram, which is used to make a decision regarding the cancer stage of the patient. An RT treatment is prescribed based on the estimated cancer stage. To calculate the probability of incorrectly staging a patient, the outcome of the model is compared to data from a Canadian cancer treatment facility. No patient-specific data are used, all data used for variable definitions are taken from the literature or taken from experts’ opinions. The expected value of the information obtained from diagnostic tests is calculated by comparing an influence diagram without any information to one with the information provided by the tests. This analysis identifies four tests, the use of which results in the highest expected value of information. The model finds a probability of incorrect staging of 18%, and a probability of incorrect treatment prescription of 8%. Since this number is small and the result is found to be sensitive to a number of variables, no definitive conclusions relevant to breast cancer staging and treatment prescription can be reached from this work.

Many congenital disorders can be diagnosed before birth through prenatal testing, but the tests are costly and often dangerous to the developing baby. Norman et al. (1998) construct an influence diagram to recommend prenatal testing strategies, considering the patient’s risk factors (i.e. maternal age, maternal diabetes) and preferences (i.e. willingness to abort). The model considers
five diseases and four tests, and consists of a sequence of decision nodes about performing tests and terminating the pregnancy. Uncertain event nodes represent risk factors and test results, and influence future decisions in the model. The health outcomes considered are: birth of a healthy baby, birth of a baby with each of the diseases, and several kinds of pregnancy loss. Utilities based on the patient’s (mother) preferences are associated with each health outcome (it is not specified how these utilities are elicited). Customized software was created to analyze the decision model. Sample analyses of decision strategies for hypothetical patients are presented, and strategies that maximize their expected utility are generated using the model. Use of the software is expected to result in prenatal testing decisions more appropriate for the patients considering their risk and preferences.

Dynamic influence diagrams and their relevance for health applications are described in Hazen (2004). Dynamic influence diagrams are influence diagrams that facilitate the representation of stochastic processes. Nodes in a dynamic influence diagram may represent variables that change state over time, such as health states; these are called stochastic or Markov nodes. Transition rates among the states that form part of stochastic nodes can be made dependent on uncertain event or decision nodes. In the health care context, rates of disease progression can be dependent on events such as the outcome of screening, treatment, or the patient’s actions.

An analysis of joint replacement decisions and their long-term consequences using a dynamic influence diagram is described in Hazen (2004). The joint replacements that are studied are mostly performed on the geriatric population, are generally optional, and the procedures are very expensive in the short-term, which has caused their cost-effectiveness to be questioned. Therefore, estimating their long-term cost-effectiveness is of importance. The model includes the choice between a joint replacement and a conservative management approach. The stochastic
nodes are: background mortality, progression under conservative management, and prosthesis status after joint replacement. Each of these consists of several health states among which transitions occur with specified transition rates. Some of the uncertain event nodes used in the model are: initial joint replacement outcome, number of infections, and number of medical revisions. The model is used to perform a cost-effectiveness analysis, being solved first quantifying QALYs in the value (outcome) node, and then quantifying costs. It is also possible to include both QALYs and costs in a value function, as well as any other outcome attributes. The results indicate that, considering improvements on quality of life (measured in QALYs), the joint replacements studied are cost-effective procedures, despite representing very high costs in the short-term.

Van Gerven et al. (2007) use a dynamic influence diagram model to define treatment strategies for high-grade carcinoid tumors, an aggressive type of neuroendocrine tumor. For this type of tumor an important decision is when, and for how long, to administer chemotherapy, since starting treatment too early and for longer time may cause unnecessary deterioration to the patient’s overall health, while treating too late and for a shorter period of time may fail to stop the tumor progression. The model was constructed with the collaboration of an expert physician and models the progression of the tumors and the effect of chemotherapy strategies over time, allowing for partially observable variables to be represented as uncertain events with probability distributions. The patient’s general health status is influenced by variables such as: age, gender, tumor mass and treatment strategy. The decision modeled is what level of chemotherapy to use at each point in time: none, reduced or standard. The reductions in tumor mass due to chemotherapy are modeled, according to the World Health Organization’s criteria, as a tumor response variable with the levels: complete remission, partial remission (more than 50% decrease in tumor mass), progressive disease (more than 25% increase in lesions) and stable disease. The level of the tumor
response variable affects tumor mass in the future. The outcomes measured are quality of life and costs. The strategy found upon solving the influence diagram is to treat once if the patient’s health status is good enough (at or below a specified level), then to let the patient recover and decide whether treat again, depending on the health status. The expert physician confirms that the strategy found by the model agrees with the treatment protocol used in clinical practice, which indicates that reasonable strategies can be found using a dynamic influence diagram model.

Another modeling tool useful for modeling stochastic processes with decision situations are Markov decision processes (MDPs) and partially observable Markov decision processes (POMDPs). The following are examples of these types of models applied to treatment decisions.

Magni et al. (2000) describe MDPs and apply one to find the optimal time to treat mild hereditary spherocytosis (HS), and compare the results to those obtained using a static decision modeling approach. MDPs are based on decision theory and discrete time Markov process theory. It is assumed that at each time point, the current state of the Markov process can be observed and an action is selected from a set of possible actions, considering the current state. The transition matrix of the Markov process is dependent at each time point on the last action selected, which leads to different outcomes that depend on the actions taken. The possible sequences of actions (strategies) are compared to determine the optimal one by scoring each strategy with an outcome function. The outcome function is the sum of costs and benefits accumulated in each step of the Markov process by performing actions and being in states.

HS consists of a chronic destruction of red blood cells. To find the optimal time to treat mild HS, Magni et al. (2000) consider four alternative treatments. The health states modeled are: without gallstones, asymptomatic gallstones, recurrent colics, occasional colics, gallbladder removed and
death. Patient age and sex are variables that affect the transition to the ‘death’ state, and the variable ‘presence of spleen’ increases the probability of gallstone formation. The problem is modeled as an MDP and as a sequence of decision trees solved at specified time points (static approach). The optimal treatment times found by the two models differ in many cases (for different patient characteristics). The optimal strategy found by the MDP model was slightly higher in life expectancy than the one found by the static approach model. Costs are not considered in the model.

Hauskrecht and Fraser (2000) describe POMDPs as MDPs that also model uncertainty associated with the Markov process being partially observable by the decision maker, that is, the current state can only be observed indirectly via imperfect observations. A POMDP model makes a distinction between states defining the stochastic process (i.e. disease states) and observations. A POMDP model consists of a set of process states (disease states), a set of actions (diagnostic and treatment interventions), a set of observations (results of diagnostic tests), a set of transition probabilities between states, and probabilities that describe the relationships among observations, states and actions.

Hauskrecht and Fraser (2000) use a POMDP to model the problem of therapy planning for patients with ischemic heart disease (IHD), which is caused by an imbalance between the supply and demand of oxygen to the heart. The state of the patient is modeled using a set of state variables and their values. State variables can represent partially observable (hidden) attributes, such as ‘status of coronary artery disease’ and ‘ischemia level’, or perfectly observable attributes, such as ‘chest pain’ and ‘stress test result’. Four treatment actions and two investigative actions are considered. An influence diagram is used to model dependencies among state variables and among variables and actions. The cost for a transition between states is a function of a cost
associated with a patient state and a cost associated with an action; these costs include both quality of life and economic cost. To test the model, an evaluation on ten patient cases generated by a cardiologist is performed. The recommendations produced by the model are almost all considered appropriate by the cardiologist, and the disagreements helped uncover deficiencies of the model due to oversimplification.

The modeling tools used in these examples include Markov models, which are used often to model the natural progression of a disease. MDPs and decision trees with Markov nodes are used to evaluate the long-term effects of different screening and treatment alternatives. POMDPs, influence diagrams and dynamic influence diagrams with Markov nodes are useful to model situations in which the health states are not known with certainty: a patient has a certain probability of being in a given state which is probabilistically dependent on the accuracy of the tests used and on other variables, such as behavioral risk factors and age. We now review some applications of OR to decisions regarding cervical cancer prevention.

2.2 Applications of Operations Research to Cervical Cancer Prevention

There is an increasing number of studies that use mathematical models and simulation to evaluate the cost-effectiveness of cervical cancer (CC) screening strategies. In order to evaluate screening strategies, it is necessary to model the progression of the disease, and how it is affected by the different strategies. Because HPV infections and pre-cancerous lesions progress and regress independently of past history, Markov models can be used to model the disease progression towards CC. However, because screening decisions depend on past history (i.e., they are not memoryless) simulation models are often used to study alternatives for CC screening. Some
Goldie et al. (1999) evaluate the cost-effectiveness of several CC screening strategies in HIV-infected women. A Markov model is used to represent the natural history of the disease and the effects of screening, diagnosis and treatment. A societal perspective is considered and lifetime screening strategies are evaluated according to their estimated costs, life expectancy, and quality-adjusted life-years (QALYs). The comparative performance of the different strategies is measured using the incremental cost-effectiveness ratio, defined as the additional cost of a specific screening strategy divided by its additional clinical benefit (measured in QALYs), compared with the next least expensive strategy. Six main screening strategies are evaluated, which include the use of either the Pap test or colposcopy and have varying frequencies. The model consists of the cervical health states: normal, low grade squamous intraepithelial lesion (LSIL), high grade squamous intraepithelial lesion (HSIL), CC, and death. Each cervical health state is stratified by CD4 cell count (a measure of the development of HIV) and by history of cervical neoplasia (abnormal and pre-cancerous cell growth in the cervix). This model does not include a health state for HPV infection to account for the difference in the probability of developing cervical lesions if one is infected or not (infection with oncogenic HPV types is now considered a necessary cause for CC). In later models (Goldie et al. 2001, Goldie et al. 2004), this distinction is included. To estimate the effect of screening, the model distinguishes between detected and undetected cervical disease. An analysis of the results found the most cost-effective screening strategy for HIV-infected women to be: perform two Pap tests six months apart, if both negative, continue with annual Pap tests.
Myers et al. (2000) construct a Markov model that simulates the natural history of HPV and CC in a hypothetical cohort of women. The model approximates age-specific HPV incidence rates in unscreened populations and is used to evaluate the cost-effectiveness of screening strategies. The model predicts the age-specific prevalence of HPV infection and of low-grade and high-grade lesions, and the age-specific incidence of CC. The disease states modeled include well (not infected), HPV (infected), LSIL, HSIL, unknown CC stages I through IV, detected CC stages I through IV, cancer survivor, hysterectomy, death from CC and death by other causes. There is no distinction made between different types of HPV, although only some types are oncogenic. The population modeled includes non-sexually active women, although HPV is transmitted through sexual contact.

Kulasingam et al. (2006b) evaluate the cost-effectiveness of extending CC screening intervals among women with prior normal Pap tests, using an extended version of the model developed in Myers et al. (2000). They simulate the natural history of CC in a theoretical cohort of women. Hypothetical women are initially distributed among the different health states based on the known disease prevalence. The number of prior normal Pap tests that each woman has is tracked, and different screening frequencies are evaluated. The results show that as the number of prior normal Pap tests increases, the costs per life-year saved increase (this is because women with more prior normal Pap tests are less likely to develop CC, so screening them is less cost-effective). This has implications for the prioritization of screening resources: women who have never been screened or who have fewer prior normal Pap tests should have priority for screening.

Besides primary screening, another interesting question related to CC prevention is how to manage ambiguous cytology test results, such as atypical squamous cells of undetermined significance (ASCUS). Many high grade lesions are identified in women with an ASCUS result,
but most women with an ASCUS result have either no lesion or low grade lesions. Kim et al. (2002) perform a cost-effectiveness analysis of different ASCUS triage strategies. The model developed in Goldie et al. (1999) and Goldie et al. (2001) is adapted and used to simulate the natural history of HPV and CC. Screening strategies are evaluated using a cost-effectiveness ratio, where effectiveness is measured in QALYs. Three screening strategies are compared: immediate colposcopy, HPV DNA triage, and repeat cytology at 6 and 12 months. Very small differences are found between the three alternatives in terms of cancer incidence reduction, but the cost difference is considerable. Immediate colposcopy is always the most costly, and HPV DNA triage is the least costly. Strategies with less frequent screening are found to be preferable to annual screening strategies.

The ASCUS and LSIL Triage Study (ALTS) is a randomized trial sponsored by the National Cancer Institute that evaluates the three triage management strategies compared in Kim et al. (2002). The findings from the ALTS trial suggest that the HPV DNA testing alternative is as effective at detecting CC as referring all women to immediate colposcopy. Kulasingam et al. (2006a) use a decision tree to analyze the ALTS trial from a cost-effectiveness standpoint. The model includes the triage strategies as decision alternatives, and the true disease states and results of tests as uncertain events. The outcomes are costs and CC detection. The results indicate that HPV DNA testing is a cost-effective alternative.

Johnson et al. (1993) also study the decision of which triage strategy to use. The alternatives considered are: to refer to colposcopy after an abnormal Pap, or repeat the smear after six months. A decision tree is used to analyze the problem, with each of the two alternatives being a branch stemming from the decision node. The possible outcomes of each test are modeled as uncertain events. The outcome branches are evaluated in terms of the percentage of women that would
develop cancer. Costs for each alternative are estimated using the results from the decision tree and cost equations. They find conservative screening with a repeat Pap after six months to be more cost-effective than referral to colposcopy.

Eggington et al. (2006) perform a cost-effectiveness analysis of different policies of referral to colposcopy. A Markov model which includes the referral of patients to colposcopy, their treatment and follow-up is used. The authors take into account reports from practice: the model is developed through interviews with clinicians, from which a large set of possible patient pathways are constructed. Guidelines are collected from many clinics throughout England, and are condensed into three types of service: high intensity, low intensity and typical. The different clinic types are evaluated with the Markov model. The outcome measures are colposcopy workload, patient waiting times, and QALYs. Their results suggest that a policy that refers patients to colposcopy after one, not two, Pap results of mild dyskaryosis is cost-effective, with the benefits achieved through the improved detection of high-grade lesions outweighing the additional costs of referring more women to the colposcopy clinics.

Goldie et al. (2004) evaluate the cost-effectiveness of using the HPV DNA test as primary screening in combination with the Pap smear in women over 30. The model is a simulation of the natural history of HPV infection and CC. A hypothetical cohort of non-infected 13 year old women enter the model. Each month they have an age-dependent probability of acquiring HPV. Health states distinguish women with previously abnormal screening tests, prior treatment, and detected CC. Strategies are compared using a cost-effectiveness ratio, and they vary by the type of cytology test, the use of HPV DNA testing, and the screening frequency. The results suggest that for women over 30, 2-year or 3-year screening strategies that use the HPV DNA test are more effective and less costly than the traditional primary screening strategy of annual cervical

Sanders and Taira (2003) evaluate the cost-effectiveness of a vaccine that protects against infections with thirteen oncogenic HPV strains. Their study was developed before the clinical trials for the HPV vaccine against HPV types 16 and 18 were completed, therefore important variables such as vaccine efficacy were not yet available, and the authors conducted extensive sensitivity analyses. The analyses assume a universal vaccination strategy for adolescent girls. A decision tree is used to evaluate the decision of whether to vaccinate, and each woman’s health is simulated by a Markov model. They track a cohort of girls who are either vaccinated against oncogenic HPV types or who receive the current standard of care. The findings suggest that the vaccine against oncogenic HPV strains is more expensive than current practice but results in greater quality-adjusted life expectancy.

Elbasha et al. (2007) evaluate the cost-effectiveness of alternative vaccination strategies with the quadrivalent HPV vaccine Gardasil. The strategies are: routine HPV vaccination of girls by age 12, routine vaccination of girls and boys by age 12, and strategies that included catch-up vaccination for those ages 12 to 24. They develop a simulation model of a heterosexually mixing population in which people have low, medium, or high sexual activity. The model simulates HPV transmission and the development of precancerous cervical lesions, cervical cancer, and genital warts. Compared with current practice, vaccinating only girls before age 12 is found to be cost-effective. However, the strategy of vaccination of girls and boys by age 12 with a catch up program up to age 24 resulted in the largest reduction of cervical cancer, high grade lesions and genital warts.
McLay et al. (2010) develop a simulation-optimization model to determine the ages at which screening should be performed. They design dynamic, age-based screening policies by varying how many lifetime screenings to perform and at what ages. This is the first exploration of dynamic strategies for cervical cancer screening, and they are found to be more cost-effective than traditional screening strategies. However, the disease model considers the presence of an HPV infection as equivalent to having a low grade cervical lesion, and it considers primary screening only with the Pap test and triage screening only with direct colposcopy.

Mathematical models like the ones described above have also been developed and applied to other settings, such as low-resource settings, other countries’ screening policies, and HPV vaccine decisions. Goldie et al. (2001) develop a state transition model to simulate the natural history of pre-cancerous lesions and CC. This model is used to evaluate the cost-effectiveness of screening strategies designed for low-resource settings. In such settings, the only feasible strategies may be those that require fewer visits and fewer lifetime screenings (because it may be difficult for women to go to the testing site), and that replace the Pap smear with visual screening methods that don’t require a laboratory (because the resources to run a laboratory may not be in place), like direct visual inspection and the HPV DNA test. The health states of the model include states corresponding to HPV infection status, cervical disease status, and HIV infection status. The progression of cervical disease is more likely to occur in HIV-infected women. The findings suggest that a single lifetime screen at 35 years of age, using a 1-visit strategy (i.e., women do not need to return shortly after the screening to retrieve results or for additional testing) with visual inspection is more effective and costs less than no screening at all. This same strategy done twice in a lifetime provides additional life expectancy, but this strategy done 3 times in a lifetime only increases mean life expectancy by 3 days, and is dominated by the strategy of screening every 5
years with visual inspection. At most screening frequencies, a more frequent 1-visit strategy with visual inspection dominates a 2-visit HPV DNA testing strategy.

Siebert et al. (2006) develop a Markov model for the German health care context and policies. The model includes pre-cancerous disease states, undiagnosed CC stages I-IV, diagnosed CC stages I-IV, death by CC, death by other, and Well. States that refer to HPV status are not included. The model predicts lifetime CC risk and mortality consistent with data from the German cancer registries, and the percentage of CC cases and deaths that can be prevented with different screening frequencies.

Favato et al. (2007) perform a cost-effectiveness analysis of HPV vaccine programs in Italy. A Markov model is used to model disease progression. Clinical benefits (cost and events prevented) and the costs associated with the vaccine program are considered. Four different hypothetical cohorts of women enter the vaccination program at different ages ranging from 11 to 25. The entering age is important in the case of vaccination because it is a one-time event that has a great effect on the probability of contracting HPV in the future. It is different than when considering screening, which occurs periodically, often yearly. The results indicate that a vaccination strategy would result in a decrease of CC cases.

Kohli et al. (2007) estimate the long-term impact of an HPV vaccine in the UK. The model consists of three modules. The natural history module models the natural history of HPV and CC without intervention using a Markov model. The parallel screening module adjusts the probabilities corresponding to the proportion of women detected and treated, and is based on a decision tree that evaluates triage screening alternatives. The vaccination module modifies the
rates of infection onset. The findings suggest that a vaccination program would result in a large reduction of CC cases and deaths.

All of the studies summarized above consider the costs as well as effectiveness of different screening strategies. Some use mathematical models such as Markov models and influence diagrams to represent diseases and interventions. Most studies of cervical cancer prevention use simulation to represent the natural history of HPV and CC in a hypothetical cohort of women, and track the effects of screening interventions. Different settings are considered, such as HIV-infected women, low-resource settings, and the policies of different countries. Both primary screening and ASCUS triage screening strategies have been studied. Most analyses consider age-dependent rates of incidence and prevalence, although only some make recommendations for different age groups. None of the analyses of cervical cancer prevention strategies consider specific risk factors of the patient in the evaluation of alternative strategies, which is what we study.
2.3 References for Chapter 2


Chapter 3: Comprehensive Model to Evaluate Cervical Cancer

Screening Strategies

3.1 Introduction

Cervical cancer is the second leading cause of female cancer mortality worldwide with 250,000 deaths each year (World Health Organization 2011). The National Cancer Institute (2011) estimates that 12,710 cases of invasive cervical cancer were diagnosed in the US in 2011, and approximately 4,290 women died from cervical cancer.

Infection with oncogenic strains of human papillomavirus (HPV) has been identified as a necessary cause of cervical cancer (Schiffman and Castle 2003, Muñoz 2000). There are over 100 strains of HPV, approximately 40 strains are sexually transmitted infections (STIs), and approximately 15 of those strains are oncogenic (Schiffman and Kjaer 2003). HPV types 16 and 18 are the two oncogenic strains of HPV that cause most (70%) cervical cancers (Munoz et al. 2004). HPV is the most common STI in the United States (Centers for Disease Control and Prevention 2011), and the prevalence of oncogenic HPV among US women 14-59 years old was found to be 15.6% by Kahn et al. (2007). However, most HPV infections clear without treatment and without causing changes in the cervix. Sellors et al. (2003) found that 51.9% of HPV-positive women cleared an HPV infection within one year; and Moscicki et al. (2006) report that although time to clearance varies between studies, almost all show that 90% of the infections are cleared after two years.
Invasive cervical cancer can be prevented with appropriate screening. In recent decades, screening for pre-cancerous cervical lesions with cervical cytology has greatly reduced the incidence of cervical cancer. Although cervical cytology tests often fail to detect lesions (a review by Nanda et al. (2000) found the Papanicolaou (Pap) test to have an average sensitivity of 58% for high grade lesions), the slow progression of lesions usually allows for them to be eventually detected with a regular screening program. More recently, an HPV DNA test has been developed and is recommended by screening guidelines (Smith et al. 2010, American College of Obstetricians and Gynecologists 2009, Wright et al. 2007). The HPV DNA test can identify if a person is infected with oncogenic strains of HPV.

Several studies have identified relationships between specific demographic and behavioral characteristics and the probability of contracting HPV infections (Kahn et al. 2007, Dunne et al. 2007, Sasagawa et al. 2005, Wright et al. 2004, Sellors et al. 2003, Winer et al. 2003, Moscicki et al. 2001). Some factors that have been found to be associated with a higher probability of HPV infection are: age (e.g., women between 18-30 years old are most likely to contract an HPV infection), higher numbers of lifetime and recent partners, and a past history of sexually transmitted diseases. Some factors that have been found to be associated with a lower probability of HPV infection are: being married, higher levels of education, and having partners with no previous partners. For example, Dunne et al. 2007 estimate the prevalence of HPV among married women to be 17.3%, with a 95% confidence interval of (14.0-21.5), while the prevalence of HPV among women that were widowed, divorced or separated was 41.2%, with a 95% confidence interval of (32.3-52.4). Dunne et al. 2007 also found the prevalence of HPV infection among women that had zero sex partners in the past year to be 18.7% (11.5-30.3), while the prevalence of HPV among women that had three to five partners in the past year was 58.6% (42.7-80.4).
Reiter et al. (2009) have identified some risky sexual behaviors that make women more likely to develop cervical cancer, such as an early age at first intercourse and increasing numbers of sex partners. Sexual behaviors such as these increase the risk of cervical cancer because they increase the risk of HPV infection.

Current screening guidelines are described in section 3.2.2 and do not incorporate the patient’s risky behaviors in the screening decisions (Smith et al. 2010, American College of Obstetricians and Gynecologists 2009, Wright et al. 2007). However, because HPV is transmitted through sexual contact, patient characteristics such as age, marital status and number of recent partners are good indicators of the probability of each patient of contracting an HPV infection and of developing cervical cancer. The effect of incorporating each patient’s risk for infection (based on her self-reported behavior) into the selection of recommended screening strategies has not been studied (although Elbasha et al. (2007) do consider different levels of sexual activity in a population when assessing HPV vaccination strategies, and Paskett et al. (2010) have studied patient adherence to risk-appropriate cervical cancer screening). With many women being on either of the extremes, low risk and high risk, it seems likely that a screening system designed for a woman of “average” risky behaviors is in fact under-screening many women and over-screening many others.

We hypothesize that personalized screening strategies that treat patients according to their specific level of risk would be more cost-effective than the current one-size-fits-all strategies. To explore this hypothesis, we investigate the impact of various screening strategies on cervical cancer prevention. Strategies that are of particular interest are those that incorporate a woman’s probability of being infected with an oncogenic HPV strain. To accomplish this, we develop a
patient-based discrete event simulation model that includes both disease progression and screening for populations of women with different probabilities of HPV infection. To evaluate the performance of different screening strategies, we collect performance metrics such as medical costs, the number of invasive cancer cases that develop, and the number of quality-adjusted life years (QALYS) accumulated throughout the course of the simulation.

### 3.2 Methods

The simulated population of women is initialized in a manner that is consistent with the age distribution of women in the US. Each woman within the simulated population is assigned an initial health state, which is defined in terms of the existence or absence of oncogenic HPV infection and the various stages of pre-cancerous cervical lesions. Our model of the natural history of the disease simulates each woman’s transitions between these health states over time.

Cervical cancer screening guidelines specify a set of tests (“primary” screening) and follow-up procedures (“triage” screening) that are approved, but leave room for some decisions to be made by healthcare providers (Smith et al. 2010, American College of Obstetricians and Gynecologists (ACOG) 2009, Wright et al. 2007). We evaluate several alternatives for primary and triage screening, (e.g., annual Pap smear with an HPV DNA test when the Pap is inconclusive), based on a simulation of the outcome of the screening tests.

Our analysis involves simulations of theoretical cohorts of 100,000 women. Each woman was followed through her screening years until she reached an end state, observing her health state and test results over time. To study the effectiveness of each screening strategy when applied to women with varying levels of risk, we varied the probability of contracting an oncogenic HPV
infection as an input parameter and repeated the simulation of each screening strategy as this probability varied from 0 to 0.50 in increments of 0.01. We did not simulate for probabilities of infection larger than 0.50 because the prevalence of oncogenic HPV infection has not been found to be greater than 0.30 in the general population or in subgroups divided by age and demographic characteristics (Kahn et al. 2007).

By quantifying and comparing total screening costs and the total quality-adjusted life years (QALYs) accumulated, we were able to evaluate the relative performance of the various screening strategies as we vary the patients’ level of risk for oncogenic HPV infection. Our simulation model includes three critical elements: the natural history of the disease, the screening strategies, and the data that populate the model. In the following we describe each of these in turn.

3.2.1 Natural History of Disease

In our model of disease progression and regression, women who are not infected with oncogenic HPV can acquire an infection, which results in a transition to the “infected” state. Infections can clear or progress to lesions, and lesions can clear, progress to more serious lesions, or regress to less serious lesions. Women can remain in the “lesion” states for many years before progressing to invasive cancer, and they may transition to “death by other cause” prior to that. In moving between health states, transition probabilities are age dependent. The natural history model represents the way in which a woman would transition between the health states if left unattended, and is depicted in Figure 3.1.

With i denoting age, the health states are as follows:
\( N_i \) Not infected with oncogenic HPV, no lesion, age \( i \)

\( I_i \) Infected with oncogenic HPV, no lesion, age \( i \)

\( L0_i \) No lesion; \( L0_i = N_i \cup I_i \)

\( L1_i \) Lesion level 1, age \( i \)

\( L2_i \) Lesion level 2, age \( i \)

\( L3_i \) Lesion level 3, age \( i \)

\( C \) Invasive cancer (end state)

\( CP \) Cancer prevented (end state)

\( DO \) Death by other cause (end state)

**Figure 3.1 Model of Disease Progression**

As the patient ages it is possible to observe disease progression (e.g. \( I_i \) to \( L_{i+1} \)) and regression (e.g. \( I_i \) to \( N_{i+1} \)).

39
The lesions level 1, 2, and 3 are commonly referred to as cervical intraepithelial neoplasia (CIN) grade 1, 2, and 3. The distribution of initial health states of women entering the simulation are consistent with reports of the age-specific prevalence of oncogenic HPV infection (Kahn et al. 2007), and with our estimates of the prevalence of lesion states as described in section 3.2.3.2.

Within the model, a woman’s health state is simulated over time until she reaches one of the three end states:

- **C**, defined as transitioning from L3 to C without having had a positive biopsy result while in L3. The stages of invasive cancer that would follow are outside the scope of our model of preventive screening.
- **CP**, defined as reaching the health state L3 and having a positive biopsy result. In practice this leads to treatment and very close surveillance (beyond the scope of preventive screening).
- **DO**, simulated using current age-specific mortality rates or by reaching age 90 prior to reaching either of the other end states.

Studies have found the mean time for an initial lesion detection to be 43.3 (95% CI, 36.4-50.1) and 46.4 (95% CI, 42.0-50.7) months after an initial detection of infection with HPV types 16/18 and other oncogenic HPV types, respectively (Trottier et al. 2009). Additionally, Liaw et al. (1999) indicate that the first cases of lesions level 2 and 3 tend to appear no earlier than after two years of follow-up among women who were HPV negative at enrollment in their study. Moscicki et al. (2006) and Schiffman and Kjaer (2003) explain that women with invasive cancer detected during screening tend to be more than 10 years older, on average, than women with lesions level 3, which suggests a long average sojourn time in the precancerous lesion state. For these reasons, a discrete time model of the natural history in which transitions occur in six month intervals is
used, and this time interval is judged to be sufficiently small, relative to the data available on observed transition times.

From a given health state at a given time, the health state six months later is randomly generated using transition probabilities consistent with available data regarding rates of disease progression and regression. These rates have been inferred from cohort studies (e.g. Sherman et al. 2003, Moscicki et al. 1998, Insinga et al. 2007, Trottier et al. 2009), and have been compiled to be used as transition probabilities in cost-effectiveness analyses (Goldie et al. 1999, Myers et al. 2000, Goldie et al. 2001, Sanders and Taira 2003, Canfell et al. 2004, Kohli et al. 2007, Bergeron et al. 2008).

Data on the rates of progression from lesion level 3 to invasive cervical cancer are limited. In general, cohort studies have lesion level 3 as an endpoint and a biopsy is performed when a participant is identified as being in this disease state. Schiffman and Kjaer (2003) note that starting treatment at the first signs of possible precancer is obligatory in the United States, and that historic literature is used to estimate the proportion of lesions level 3 that develop into invasive cervical cancer, such as Petersen (1956) and Kinlen and Spriggs (1978). Similarly, Widdice and Moscicki (2008) reference a study (Syrjanen et al. 1992) which followed a cohort of women and collected data on the progression of lesions level 3 to cervical cancer, and note that studies such as these cannot be repeated now that lesions level 3 are recognized as precancerous lesions.

As in several other studies (Myers et al. 2000, Sanders and Taira 2003, Goldie et al. 2004, Bergeron et al. 2008), our health state definitions do not distinguish between the presence or absence of oncogenic HPV infections when women are in the L1, L2, and L3 states. In this sense,
they do not reflect an interaction between the existence of lesions and infections, which may be pertinent to the assessment of the efficacy of some screening strategies. Our decision to forego the use of the more specific health states is primarily due to the lack of quantitative information regarding transitions between health states given the simultaneous presence of these lesions and infections. It is well known that oncogenic HPV infections are much less likely to clear spontaneously when lesions are present (Manos et al. 1999, Kulasingam et al. 2006). However, cohort studies that provide data on oncogenic HPV infection clearance often do not examine patients for lesions, and consequently cannot provide insight on clearance data conditioned on lesion states (Ho et al. 1998, Munoz et al. 2004). Some cohort studies that did perform both HPV DNA testing and colposcopies (Moscicki et al. 1998, Insinga et al. 2007, Trottier et al. 2009) do report the appearance of lesions in addition to the incidence and clearance of HPV infections. Some of these also report data referring to the likelihood of lesions appearing conditioned on infection state, and are useful to determine lesion progression rates (Moscicki et al. 1998, Trottier et al. 2009). However, none of these cohort studies report data on the rates of infection clearance conditioned on lesion state.

3.2.2 Screening Strategies

Regular preventive screening practices for cervical cancer are called “primary screening”. Positive primary screening test results are followed by a colposcopy, and negative test results are not given further follow up until the time of the next regular primary screening. When the primary screening test results are inconclusive, follow up tests called “triage screening” are performed. A Pap result of Atypical Squamous Cells of Undetermined Significance (ASCUS) is considered inconclusive. The American Society for Colposcopy and Cervical Pathology (ASCCP) consensus guidelines (Wright et al. 2007) recommend a yearly Pap test for primary
screening. For women over thirty, a combination of the Pap test and the HPV DNA test, every three years, is also approved for primary screening (a more frequent use of the HPV DNA test or its use as a sole primary screening test are not approved). For triage screening, the three alternatives approved are: two repeat Pap tests in six and twelve months, an immediate HPV DNA test, and an immediate colposcopy.

We evaluate primary and triage screening strategies consistent with the current ASCCP consensus guidelines (Wright et al. 2007), with the following variations. We evaluate all screening strategies in women of all ages. We added two primary screening frequencies to our set of evaluated strategies: every 2 and 3 years. This is because recent guidelines from the ACOG (2009) recommend that women under age 30 be screened every 2 years, and that women over age 30 can be screened every 3 years if they have had 3 consecutive negative Pap results. We evaluated a total of 12 distinct combinations of primary and triage screening strategies.

All simulated women are screened until they reach one of the 3 terminal states. Our model accounts for the fact that screening tests have varying levels of precision (Nanda et al 2000, Solomon et al. 2001, ALTS Group 2003, Kulasingam et al. 2006), and the outcomes of these tests are conditioned on the patient’s current health state. We quantify how each of the screening strategies performs in terms of total screening costs, number of invasive cancer cases, and the number of quality-adjusted life years (QALYs) accumulated in the population.

The management of detected lesions is modeled in accordance with the ASCCP consensus guidelines for the management of women with CIN (Wright et al. 2007b). If any primary or triage test is positive for cervical lesions, a colposcopy is simulated. If the outcome of the colposcopy confirms a cervical lesion, we simulate the lesion being treated by biopsy and the patient is
monitored closely (a Pap every 6 months, with an immediate colposcopy following any abnormal result) until she has two consecutive negative Pap results. If the colposcopy is negative, the patient is still placed under close monitoring.

3.2.3 Input Data

Numerous studies and data sources are available for inclusion in a model such as ours. We use prevalence and transition probabilities to model the natural history of the disease, and data regarding the accuracy and cost of screening tests to simulate screening strategies. Some of these data can be found directly in the literature, although in some cases there are conflicting sources. Some data are not available in the literature, so models are required to estimate them. We consider these 2 cases separately in the following subsections.

3.2.3.1 Data from the Literature

The prevalence of oncogenic HPV infections, by age, is necessary for defining initial health states. These data are available from Kahn et al. (2007), and appear in Table 3.1.

<table>
<thead>
<tr>
<th>Age</th>
<th>Prevalence of Oncogenic HPV</th>
</tr>
</thead>
<tbody>
<tr>
<td>14-17</td>
<td>0.133</td>
</tr>
<tr>
<td>18-21</td>
<td>0.269</td>
</tr>
<tr>
<td>22-25</td>
<td>0.290</td>
</tr>
<tr>
<td>26-29</td>
<td>0.106</td>
</tr>
<tr>
<td>30-39</td>
<td>0.172</td>
</tr>
<tr>
<td>40-59</td>
<td>0.112</td>
</tr>
</tbody>
</table>

Table 3.1 Prevalence of Oncogenic HPV Infection (Kahn et al. 2007)

In defining transitions from $N_i$ to $I_{i+1}$, it is necessary to estimate the probability that a patient of age $a$ without an oncogenic HPV infection becomes infected within 6 months. We define this
probability as $P_a(I)$ and note that it corresponds to an age specific six month incidence rate. To estimate $P_a(I)$, we use the data in Table 3.1 and the epidemiologic relationship $Prevalence = Incidence \times Duration$ (Aschengrau and Seage 2003). Ho et al. (1998) found the median duration of oncogenic HPV infection to be 8 months (95% CI:7-10) in a population with a mean age of 20. Richardson et al. (2003) found the mean duration to be 16.3 months (95% CI:13.7–18.9) in a population with a mean age of 23. Munoz et al. (2004) studied a population of women age 15-85, and found the median duration of oncogenic HPV infection to be 14.2 months (95% CI:10.2-17.2) in women under 30 and 16.2 (95% CI:13.0-17.7) in women 30 and over. Based on this information, we assume a duration of oncogenic HPV infection of 8 months for women under age 20, and of 15 months for women over age 20. From this we have the 6-month transition probabilities ($P_a(I)$) in Table 3.2 (the infection probability for women over 60 is taken from Myers et al. (2000)).

<table>
<thead>
<tr>
<th>Age (a)</th>
<th>$P_a(I)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>14-17</td>
<td>0.105</td>
</tr>
<tr>
<td>18-19</td>
<td>0.228</td>
</tr>
<tr>
<td>20-21</td>
<td>0.114</td>
</tr>
<tr>
<td>22-25</td>
<td>0.124</td>
</tr>
<tr>
<td>26-29</td>
<td>0.043</td>
</tr>
<tr>
<td>30-39</td>
<td>0.071</td>
</tr>
<tr>
<td>40-59</td>
<td>0.046</td>
</tr>
<tr>
<td>≥ 60</td>
<td>0.002</td>
</tr>
</tbody>
</table>

Table 3.2 Estimated Age-Specific 6-Month Transition Probabilities from Normal to Infected State ($P_a(I)$)

To populate our model of the natural history of the disease, we used the annual probabilities of infection clearance and lesion progression and regression from Sanders and Taira (2003), in addition to the transition probabilities between lesion states L2 and L3 from Bergeron et al. (2008). Regarding transitions to terminal states, we used the transition probabilities from Sanders.

In order to include screening tests, it is necessary to represent their accuracy – both the sensitivity and the specificity. To do this, it is necessary to distinguish between the “health states” as defined in section 3.2.1, and the “test results”, which we define as follows.

Test results:

\[
\begin{align*}
P_+ & \quad \text{Positive Pap result} \\
P_0 & \quad \text{Pap result of ASCUS} \\
P & \quad \text{Negative Pap result} \\
P_0 & \quad \text{P}_0 \text{JP}_+ \quad \text{(i.e., non-negative Pap result)} \\
H_+ & \quad \text{Positive for oncogenic HPV DNA test result} \\
H & \quad \text{Negative for oncogenic HPV DNA test result} \\
T_+ & \quad \text{Positive colposcopy result} \\
T & \quad \text{Negative colposcopy result}
\end{align*}
\]

We interpret the results of most studies reporting the prevalence of lesions, and the accuracy of the Pap, HPV DNA test, and colposcopy to detect lesions (Nanda et al. 2000, Shlay et al. 2001, Dannecker et al. 2004, Arbyn et al. 2004), as test results indicating lesions, not as lesion health states. However, the Low-Grade Squamous Intraepithelial Lesion Triage Study (ALTS) (Solomon et al. 2001, ALTS Group 2003, Kulasingam et al. 2006) and the study conducted by Manos et al. (1999) were rigorous enough in their testing with colposcopy (including performing repeat
colposcopies and having all samples analyzed by a separate quality control group) that we interpret their final diagnoses as health states $L_i$, that is, as accurate indicators of the existence of lesions.

From ALTS (Kulasingam et al. 2006) we obtain data on the accuracy of colposcopy, conditioned on a patient’s lesion status, shown in Table 3.3. Note that although past analyses of the cost-effectiveness of screening strategies assume that the sensitivity of colposcopy is perfect or almost perfect (Sanders and Taira 2003, Goldie et al. 2004, Elbasha et al. 2007, Bergeron et al. 2008), the data in Table 3.3 suggest rather clearly that this is not the case.

<table>
<thead>
<tr>
<th>Probability of test result</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$P(T_+</td>
<td>L_0)$</td>
</tr>
<tr>
<td>$P(T_+</td>
<td>L_1)$</td>
</tr>
<tr>
<td>$P(T_+</td>
<td>L_2)$</td>
</tr>
<tr>
<td>$P(T_+</td>
<td>L_3)$</td>
</tr>
<tr>
<td>$P(T_+</td>
<td>L_0)$</td>
</tr>
<tr>
<td>$P(T_+</td>
<td>L_1)$</td>
</tr>
<tr>
<td>$P(T_+</td>
<td>L_2)$</td>
</tr>
<tr>
<td>$P(T_+</td>
<td>L_3)$</td>
</tr>
</tbody>
</table>

Table 3.3 Colposcopy Result Conditioned on Lesion State (Kulasingam et al. 2006)

Data on the probability of a positive Pap result conditioned on lesion state, a non-negative Pap result conditioned on lesion state, and the prevalence of ASCUS Pap results are available in the literature and are shown in Table 3.4. These data are used to estimate the probabilities of an ASCUS Pap result and a negative Pap result conditioned on lesion state as described in section 3.2.3.2. The lower and upper bounds on these estimated probabilities, which appear as columns LB and UB in Table 3.4, correspond to 95% confidence intervals. Table 3.4 also includes data on the probability of lesions conditioned on an ASCUS Pap result.
Table 3.4 Parameters Related to Pap Results Found in the Literature

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
<th>LB</th>
<th>UB</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>$P(P_0</td>
<td>L_0)$</td>
<td>0.020</td>
<td>0.008</td>
<td>0.032</td>
</tr>
<tr>
<td>$P(P_1</td>
<td>L_1)$</td>
<td>0.149</td>
<td>0.095</td>
<td>0.203</td>
</tr>
<tr>
<td>$P(P_2</td>
<td>L_2)$</td>
<td>0.218</td>
<td>0.123</td>
<td>0.313</td>
</tr>
<tr>
<td>$P(P_3</td>
<td>L_3)$</td>
<td>0.407</td>
<td>0.314</td>
<td>0.506</td>
</tr>
<tr>
<td>$P(P_0</td>
<td>L_0)$</td>
<td>0.318</td>
<td>0.285</td>
<td>0.351</td>
</tr>
<tr>
<td>$P(P_0</td>
<td>L_1)$</td>
<td>0.637</td>
<td>0.552</td>
<td>0.722</td>
</tr>
<tr>
<td>$P(P_0</td>
<td>L_2)$</td>
<td>0.834</td>
<td>0.787</td>
<td>0.875</td>
</tr>
<tr>
<td>$P(P_0</td>
<td>L_3)$</td>
<td>0.834</td>
<td>0.787</td>
<td>0.875</td>
</tr>
<tr>
<td>$P(P_0)$</td>
<td>0.100</td>
<td>0.090</td>
<td>0.110</td>
<td>Kulasingam et al. 2002</td>
</tr>
<tr>
<td>$P(P_0)$</td>
<td>0.040</td>
<td>0.030</td>
<td>0.050</td>
<td>Manos et al. 1999</td>
</tr>
<tr>
<td>$P(L_0</td>
<td>P_0)$</td>
<td>0.669</td>
<td>0.642</td>
<td>0.696</td>
</tr>
<tr>
<td>$P(L_1</td>
<td>P_0)$</td>
<td>0.167</td>
<td>0.146</td>
<td>0.188</td>
</tr>
<tr>
<td>$P(L_2</td>
<td>P_0)$</td>
<td>0.080</td>
<td>0.064</td>
<td>0.096</td>
</tr>
<tr>
<td>$P(L_3</td>
<td>P_0)$</td>
<td>0.084</td>
<td>0.068</td>
<td>0.100</td>
</tr>
</tbody>
</table>

* Kulasingam et al. (2006) and the ALTS Group (2003) do not report this value, and Manos et al. (1999) report $P(P_0|L_2)$. In the absence of available data, we used the same value here as the value of $P(P_0|L_3)$ from the ALTS Group (2003).

Data for the probability of a positive oncogenic HPV DNA result, conditioned on the lesion state, appear in Table 3.5. These are used to define the efficacy of the HPV DNA test in predicting the existence of lesions. The data in Table 3.5 are also used in section 3.2.3.2 to infer probabilities that are not directly available in the literature.
<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
<th>LB</th>
<th>UB</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>P(H₀</td>
<td>L₀)</td>
<td>0.379</td>
<td>0.338</td>
<td>0.420</td>
</tr>
<tr>
<td>P(H₀</td>
<td>L₀)</td>
<td>0.305</td>
<td>0.273</td>
<td>0.337</td>
</tr>
<tr>
<td>P(H₁</td>
<td>L₁)</td>
<td>0.876</td>
<td>0.826</td>
<td>0.926</td>
</tr>
<tr>
<td>P(H₁</td>
<td>L₁)</td>
<td>0.696</td>
<td>0.615</td>
<td>0.777</td>
</tr>
<tr>
<td>P(H₁</td>
<td>L₂)</td>
<td>0.953</td>
<td>0.908</td>
<td>0.998</td>
</tr>
<tr>
<td>P(H₂</td>
<td>L₂)</td>
<td>0.948</td>
<td>0.883</td>
<td>0.983</td>
</tr>
<tr>
<td>P(H₂</td>
<td>L₃)</td>
<td>0.924</td>
<td>0.887</td>
<td>0.952</td>
</tr>
<tr>
<td>P(H₃)</td>
<td>0.284</td>
<td>0.258</td>
<td>0.310</td>
<td>Kulasingam et al. 2002</td>
</tr>
<tr>
<td>P(H₃)</td>
<td>0.156</td>
<td>0.140</td>
<td>0.172</td>
<td>Kahn et al. 2007</td>
</tr>
</tbody>
</table>

Table 3.5 Probability of Positive Oncogenic HPV DNA Result

Table 3.5 shows results from Manos et al. (1999) and from ALTS (ALTS Group 2003, Kulasingam et al. 2006). Note that even between these two carefully conducted studies, the data conflict. We chose to use the results from Kulasingam et al. (2006) because they are derived from more recently collected data, and are based on a larger sample size than the study by Manos et al. (1999). There is also a discrepancy between Kulasingam et al. (2002) and Kahn et al. (2007). We use the values from Kahn et al. (2007) because they are based on more recent data from the National Health and Nutrition Examination Survey (NHANES).

For the evaluation of the cost of the various screening strategies we use data on the costs of screening tests (Table 3.6) and an estimate of the costs of each cancer case (Table 3.7). The screening costs in Table 3.6 are derived from average Medicare charges (Melnikow et al. 2010).
Table 3.6 Screening Costs, USD (Melnikow et al. 2010)

<table>
<thead>
<tr>
<th>Test</th>
<th>Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>Liquid-based Pap smear</td>
<td>$30</td>
</tr>
<tr>
<td>Office visit</td>
<td>$72</td>
</tr>
<tr>
<td>Physician review (for abnormal results only)</td>
<td>$15</td>
</tr>
<tr>
<td>Human papillomavirus test</td>
<td></td>
</tr>
<tr>
<td>Test cost</td>
<td>$50</td>
</tr>
<tr>
<td>Office visit</td>
<td>$72</td>
</tr>
<tr>
<td>Colposcopy (no biopsy) + office visit</td>
<td>$143</td>
</tr>
<tr>
<td>Colposcopy (with biopsy) + office visit</td>
<td>$271</td>
</tr>
</tbody>
</table>

Table 3.6 Screening Costs, USD (Melnikow et al. 2010)

The total cost of an invasive cervical cancer case was estimated using the parameters shown in Table 3.7, from the following sources: Goldhaber-Fiebert et al. (2008), Surveillance, Epidemiology and End Results cancer statistics (survival), Cancer Trends Progress Report (2009/2010 update), Elwood and Sutcliffe (2010). We multiply the economic burden per life lost by the years of life lost (YLL) for each person, which we calculate by subtracting the patient’s age when she reaches the invasive cancer state, from the average life expectancy (80 years, Arias 2010).

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cost of treatment</td>
<td>$41,959</td>
</tr>
<tr>
<td>Mortality given incidence</td>
<td>33 %</td>
</tr>
<tr>
<td>Economic burden per life lost</td>
<td>$28,000/yr * YLL</td>
</tr>
</tbody>
</table>

Table 3.7 Parameters Used to Estimate the Cost of an Invasive Cervical Cancer Case

To estimate the number of quality-adjusted life years (QALYs) accumulated in the population, we used the quality adjusted weights from the Cost-Effectiveness Analysis Registry (2011) for the L, C, and CP health states.
3.2.3.2 Data Inferred

After conducting a thorough review of the literature in which we identified the relevant data sources described in the previous section, we found that there is data necessary to populate our model that are not available in the literature. In this section, we describe our methods for using probability computations and peripheral data to estimate the values of data that are not readily available in the literature.

Prevalence of lesions

Data on the true prevalence of lesions are not directly available in the literature. This is because the epidemiology studies only report the lesions that they can detect, and the tests to detect lesions (i.e. Pap, colposcopy) have neither perfect sensitivity nor perfect specificity. As previously described, we interpret the data of the final diagnoses from ALTS (ALTS Group 2003, Kulasingam et al. 2006) as health states. However, all participants in this study had a previous Pap result of ASCUS. Therefore, from ALTS we have data on the prevalence of lesions conditioned on an ASCUS Pap result, which we use to estimate the data on the prevalence of lesions, and on the prevalence of lesions conditioned on oncogenic HPV infection, as follows.

Deriving $P(L_i)$ conditioning on Pap result

Noting that

$$P(L_i|P_j) P(P_j) = P(L_i) P(P_j|L_i)$$

we have

$$P(L_i) = \frac{P(L_i|P_j) P(P_j)}{P(P_j|L_i)} \quad \forall \ i, j$$

From Table 3.4 we have $P(L_i|P_0)$ and $P(P_0)$, therefore we use
From Table 3.4, \( P(P_0|L_i) \) can be found as
\[
P(P_0|L_i) = P(P_{0+}|L_i) - P(P_+|L_i)
\]

We use data from Table 3.5 to estimate \( P(L_i|H_+) \), which we use to assign a lesion state to cases of prevalent infections upon entry to the model, applying the assumption that the HPV DNA test is a perfect indicator of oncogenic HPV infection, so that \( P(I|H_+) = 1 \). We also estimate \( P(L_i|H_-) \), which is necessary to assign a lesion state to women who enter the model free of infection, applying the assumption that \( P(I|H_-) = 1 \). For these calculations we do not consider an age-specific infection probability, instead we use the average infection probability across the population.

Deriving \( P(L_i|H_+) \):
\[
P(L_i|H_+) = \frac{P(H_+|L_i) P(L_i)}{P(H_+)} \tag{2}
\]

Deriving \( P(L_i|H_-) \):
\[
P(L_i) = P(L_i|H_+) P(H_+) + P(L_i|H_-) P(H_-)
\]
\[
P(L_i|H_-) = \frac{P(L_i) - P(L_i|H_+) P(H_+)}{P(H_-)} \tag{3}
\]

The data from the literature that are needed as input for these derivations are presented in Tables 3.4 and 3.5. These involve multiple data sources, and in some cases, conflicting data. Our goal is to select for our calculations values that are as close as possible to being within the reported
confidence intervals, while satisfying the laws of probability as represented in equations (1) – (3). Here we present an optimization model which we developed for this purpose.

The objective function minimizes the square of the deviation of each input parameter from the corresponding mean value reported in the literature, which is represented as \( \hat{p}_i \),

\[
\min \sum_i \left[ P(L_i|P_0) - \hat{p}_{L_i|P_0} \right]^2 + \left[ P(P_0) - \hat{p}_{P_0} \right]^2 + \sum_i \left[ P(P_+|L_i) - \hat{p}_{P_+|L_i} \right]^2 \\
+ \sum_i \left[ P(P_0^+|L_i) - \hat{p}_{P_0^+|L_i} \right]^2 + \sum_i \left[ P(H_+|L_i) - \hat{p}_{H_+|L_i} \right]^2 + \left[ P(H_+) - \hat{p}_{H_+} \right]^2
\]

We calculate \( P(L_i) \) using (1). The prevalence of invasive cervical cancer has been found to be very small: 0.00057 and 0.00074 (ALTS Group 2003, Kulasingam et al. 2002). Therefore we approximate it to zero and consider that the probabilities of being in the health states other than invasive cervical cancer \( (L_0 - L_3) \), should sum to 1:

\[
\sum_i \frac{P(L_i|P_0) P(P_0)}{P(P_0|L_i)} = 1, \quad \text{where } P(P_0|L_i) = P(P_0^+|L_i) - P(P_+|L_i).
\]

Similarly, the probabilities of being in each of the lesion states conditioned on a positive HPV DNA result, \( P(L_i|H_+) \), which are calculated using (2), must sum to 1:

\[
\sum_i \frac{P(H_+|L_i) P(L_i)}{P(H_+)} = 1
\]

The probabilities of being in each of the lesion states conditioned on a negative HPV DNA result \( (P(L_i|H_-), \text{calculated using } (3)) \), must sum to 1:
In addition, we have

\[ 0 \leq P(L_i) \leq 1 \]
\[ 0 \leq P(L_i|H_+) \leq 1 \]
\[ 0 \leq P(L_i|H_-) \leq 1 \]
\[ 0 \leq P(L_i|P_0) \leq 1 \]
\[ 0 \leq P(P_0) \leq 1 \]
\[ 0 \leq P(P_+|L_i) \leq 1 \]
\[ 0 \leq P(P_0+|L_i) \leq 1 \]
\[ 0 \leq P(H_+|L_i) \leq 1 \]
\[ 0 \leq P(H_+) \leq 1 \]

A solution to this optimization model is shown in Table 3.8. The values for each parameter shown in Table 3.8 are as close as possible to those found in the literature (Tables 3.4 and 3.5). In some cases those values are not within the 95% confidence intervals reported in the literature.

| P(L_0|P_0) | P(P_0|L_0) | 0.195 |
| P(L_1|P_0) | P(P_0|L_1) | 0.387 |
| P(L_2|P_0) | P(P_0|L_2) | 0.834 |
| P(L_3|P_0) | P(P_0|L_3) | 0.836 |
| sum       | 1.00      |

| P(P_0) | P(H_-) | 0.215 |
| P(P_+|L_0) | P(H_+|L_0) | 0.180 |
| P(P_+|L_1) | P(H_+|L_1) | 0.504 |
| P(P_+|L_2) | P(H_+|L_2) | 0.953 |
| P(P_+|L_3) | P(H_+|L_3) | 0.944 |

Table 3.8 Input Data Selected with NLP
Using the output in Table 3.8 in combination with (1) – (3), we obtain the results in Table 3.9.

These are estimates of data necessary for our model, which are not directly available in the literature.

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>P(L₀)</td>
<td>0.895</td>
<td>P(L₀</td>
<td>H₁)</td>
</tr>
<tr>
<td>P(L₁)</td>
<td>0.103</td>
<td>P(L₁</td>
<td>H₁)</td>
</tr>
<tr>
<td>P(L₂)</td>
<td>0.0002</td>
<td>P(L₂</td>
<td>H₁)</td>
</tr>
<tr>
<td>P(L₃)</td>
<td>0.002</td>
<td>P(L₃</td>
<td>H₁)</td>
</tr>
</tbody>
</table>

Table 3.9  Data on Probabilities of Lesion Derived Using Probability Computations and Input Data Selected with NLP

Accuracy of Pap results

We have identified the Pap results that occur in our model as: \( P(L_i) \), \( P(P_0 L_i) \), and \( P(P_- L_i) \) \( \forall i \). However, not all the data necessary to represent these are directly available in the literature. Many cost-effectiveness studies (Kim et al. 2002, Sanders et al. 2003, Goldie et al. 2004, Elbasha et al. 2007, Goldhaber-Fiebert et al. 2008) report the values used for the Pap test’s specificity and sensitivity without making further distinctions such as the proportion of lesions that are accurately diagnosed with ASCUS Pap results compared to positive Pap results.

Furthermore, in reporting sensitivity and specificity, information regarding the level of the lesions involved is not provided (Kim et al. 2002, Goldie et al. 2004, Elbasha et al. 2007).

In addition to the prevalence of lesions and the efficacy of the HPV DNA test in detecting them, we estimate the data necessary to represent the following three Pap results: positive, ASCUS, and negative. Note that

\[
P(P_0 L_i) = \frac{P(L_i P_0) P(P_0)}{P(L_i)}
\]

(8)
\[ P(P_+ | L_0) = 1 - [P(P_0 | L_0) + P(P_+ | L_0)] \]  

Combining equations (8) and (9) with the input data selected in Table 3.8 and the data derived in Table 3.9 we obtain

| $P(P_+ | L_0)$ | 0.100 | $P(P_0 | L_0)$ | 0.095 | $P(P_+ | L_0)$ | 0.805 |
|---------------|------|----------------|------|----------------|------|
| $P(P_+ | L_1)$ | 0.250 | $P(P_0 | L_1)$ | 0.137 | $P(P_+ | L_1)$ | 0.613 |
| $P(P_+ | L_2)$ | 0.250 | $P(P_0 | L_2)$ | 0.584 | $P(P_+ | L_2)$ | 0.166 |
| $P(P_+ | L_3)$ | 0.400 | $P(P_0 | L_3)$ | 0.436 | $P(P_+ | L_3)$ | 0.164 |

Table 3.10  Accuracy of Pap Results

3.2.4 Description of the Experiment

The data described in Section 3.2.3 were incorporated in a patient-based simulation model of disease progression and screening constructed using Arena 12 (Rockwell Software). To validate this model we assumed a traditional screening strategy of yearly primary screening and triage with a repeat Pap test, the results of the validation are presented in Section 3.3.1.

To evaluate the performance of the 12 screening strategies on populations of different levels of risk, we simulated hypothetical risk-homogeneous populations with probabilities of infection which varied from 0 to 0.5 in increments of 0.01. In each simulated population of 100,000 women, all women had the same probability of contracting an infection, although the presence/absence of the infection was probabilistically independent from woman to woman. This probability of infection did not change over time (for these hypothetical risk-homogeneous populations we did not use the age-specific probabilities of infection from Table 3.2).
The simulated populations followed an age distribution consistent with that of the US population (Kahn et al. 2007), the age at entry to the model ranged from 14 to 59. The 100,000 women in each population were simulated sequentially, with initial health states, ages, disease progression and test results that evolve independently from woman to woman. Each woman was followed through her screening years until she reached an end state, observing her health state and test results each year. All women that reached age 90 transitioned to the end state DO (death by other cause).

The probability of contracting an oncogenic HPV infection was systematically varied in order to observe differences in the performance of various screening strategies in response to this probability of infection. For each probability considered, 70 different constant risk populations of 100,000 women were studied.

3.3 Results

3.3.1 Validation

In this section we compare the results of our simulation model on the number of incident cervical cancer cases to the reported number of cervical cancer cases in the US population. To make this comparison, we use data from the Surveillance, Epidemiology and End Results (SEER) Program. The incidence rate of cervical cancer for 2004-2008 was 8.1 cases per 100,000 women per year. Evaluating our model using the traditional screening strategy of annual screening and triage with a repeat Pap test, we find the annual incidence rate of cervical cancer to be 8.93 cases per 100,000 women.
We also compare our results to SEER’s distribution of cervical cancer diagnoses by age at diagnosis. In Figure 3.2 we present the cumulative distribution of diagnoses by age as reported by SEER and as observed from our model, and note that although the distributions are not identical, the output from our model is similar to the distribution from SEER. Given the inconsistencies found in the available input data as reviewed in Section 3.2.3 and the fact that our simulation model is only an approximation of reality with simplifying assumptions, it is not surprising that there are some small discrepancies between our model output and SEER data. In addition to this, the higher cervical cancer incidence rate in our model may be due in part to the fact that SEER quantifies the number of cancer cases that are detected and reported, while our output reflects the total number of cancer cases that occur, because the simulation environment allows us to register all health states.

Figure 3.2 Distribution of Diagnoses by Age, Comparison of Our Results to SEER Data
3.3.2 Results of Homogeneous Populations

The performance of screening strategies was assessed based on the total costs incurred and the total QALYs accumulated in a population of 100,000 women. We calculated incremental cost-effectiveness ratios (ICER) in which the additional costs of a strategy are divided by the additional QALYs compared with the next, less costly strategy. Strategies that were both more costly and less effective in terms of QALYs than other strategies were considered dominated. In the US, more costly but more effective strategies costing more than $50,000 per QALY gained are commonly not considered cost-effective (Eichler et al. 2004, Kulasingam and Myers 2003). Others (Goldhaber-Fiebert et al. 2008) also consider strategies that fall within the range of $50,000-$100,000 per QALY to be cost-effective. In our analyses we consider alternatives that cost more than $50,000 per QALY gained to not be cost-effective.

To illustrate, Figure 3.3 shows the performance of each of the screening strategies in terms of Total Costs and QALYs for a homogeneous population with a constant P(I)=0.06. The strategies that are to the right (more costly) and lower (less effective) than other strategies are dominated. Notice that in this case three strategies are not dominated.
In Tables 3.11 and 3.12, we have coded the screening strategies as follows: $P;T$. Where $P$ corresponds to the primary screening strategy and $T$ corresponds to the triage screening strategy. The options for primary screening are $P_n$, denoting a Pap test every $n$ years ($n$: 1, 2, 3); and $PD$, denoting a Pap and an HPV DNA test every 3 years. The options for triage screening are: $P = \text{'repeat Pap in 6 & 12 months'}$; $D = \text{'HPV DNA test'}$; $C = \text{'immediate colposcopy'}$.

In Table 3.11 we present our cost-effectiveness analysis of the three non-dominated screening strategies for the homogeneous population with a constant $P(I)=0.06$ whose results are illustrated in Figure 3.3. The strategies are evaluated in order of increasing costs. The least expensive of these is $P_2;D$. $P_2;P$ is compared with $P_2;D$. With an ICER of 180,058, $P_2;P$ is not considered to be cost-effective in comparison to $P_2;D$. Next we compare $P_1;D$ with $P_2;D$, and $P_1;D$ is found to be cost-effective (since the ICER does not exceed $50,000), resulting in it being the preferred...
strategy among the 12 that we investigated. Applying this process for each value of P(I) that we evaluated, we obtain the results in Table 3.12.

<table>
<thead>
<tr>
<th>Strategy</th>
<th>Total</th>
<th>Incremental</th>
<th>*Comparison</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Costs</td>
<td>QALYs</td>
<td>Costs</td>
</tr>
<tr>
<td>P2;D</td>
<td>485,007,048</td>
<td>7,797,566</td>
<td>-</td>
</tr>
<tr>
<td>P2;P</td>
<td>506,866,138</td>
<td>7,797,687</td>
<td>21,859,090</td>
</tr>
<tr>
<td>P1;D</td>
<td>585,358,890</td>
<td>7,800,444</td>
<td>100,351,842</td>
</tr>
</tbody>
</table>

Table 3.11 Cost-Effectiveness Analysis for Populations with a Constant P(I)=0.06
Preferred strategy is shown in bold, considering a cost-effectiveness threshold of $50,000/QALY. The values presented are the average values across all the replications.

Table 3.12 identifies the preferred strategies for each risk level, with risk levels defined by the constant 6-month incident infection probability (P(I)). That is, for a constant P(I) ≤ 0.03, strategy P3;D (primary screening with Pap test every 3 years, triage screening for ASCUS results via an HPV DNA test); for 0.04 ≤ P(I) ≤ 0.05, strategy P2;D, and for P(I) ≥ 0.06, strategy P1;D. We refer to these ranges as low, medium, and high risk, respectively (Table 3.12 also shows the average percentage of our simulated population that is in each of the risk levels, when simulating a mixed-risk population as in Section 3.4).

<table>
<thead>
<tr>
<th>P(I)</th>
<th>≤ 0.03</th>
<th>0.04 – 0.05</th>
<th>≥ 0.06</th>
</tr>
</thead>
<tbody>
<tr>
<td>Preferred Strategy</td>
<td>P3;D</td>
<td>P2;D</td>
<td>P1;D</td>
</tr>
<tr>
<td>% of Population</td>
<td>25.3%</td>
<td>48.5%</td>
<td>26.3%</td>
</tr>
</tbody>
</table>

Table 3.12 Preferred Screening Strategies for each Risk Level

As expected, the preferred strategy for the lowest risk women is a more conservative screening strategy, and for the highest risk women the preferred strategy consists of a higher frequency of primary screening.
We note that ASCUS triage with the HPV DNA test was found to be the most cost-effective strategy in terms of total costs and QALYs accumulated when combined with primary screening with the Pap test at all frequencies evaluated.

3.4 Extension to Mixed-Risk Population

Given that in the hypothetical risk-homogeneous populations different risk levels were found to be associated with different preferred screening strategies, we next investigated the performance of a risk-differentiated screening strategy applied to a population with varying levels of risk.

We simulated a cohort of 100,000 women at different risks for acquiring an oncogenic HPV infection. Women of different age groups, consistent with the age distribution of the US population, were created in the simulation. An age-dependent probability of acquiring an oncogenic HPV infection, \( P_a(I) \), was assigned to each woman (Table 3.2). In this way, each woman had a risk-specific \( P_a(I) \) based on her age, which is a well known risk factor for HPV infection. Note that this approach is using age as the only risk factor considered in assigning a probability of infection and in implementing the risk-specific strategy. This results in the risk-specific strategy being essentially an age-specific approach. This simulation was performed in the same disease progression and test environment as the simulations of hypothetical populations with a uniform risk level described previously.

We evaluated the 12 screening strategies that are based on current guidelines, and a new, risk-specific screening strategy derived via the results presented in Table 3.12. This risk-specific strategy screened each woman with the strategy deemed most appropriate for her particular risk.
We examined two different approaches to implementing the risk-specific strategy. These differed in the way that each woman’s age-specific probability of infection with oncogenic HPV, $P_a(I)$, was used to classify her as low, medium, or high risk. The first approach was to use each woman’s $P_a(I)$ at the time of screening. The second approach was to use each woman’s infection probability from $i$ years prior to the time of screening ($i : 5, 10, 15, 20, 25$), $P_a(I)$, to determine her risk level, to account for the fact that the lesions and cervical cancer develop over several years. Note that in this way we had six different risk-specific screening approaches.

### 3.4.1 Results of Mixed-Risk Population

The cost-effectiveness analysis is shown in Table 3.13, which evaluates the eight non-dominated screening strategies. This analysis was performed in the same manner as described for Table 3.11. The risk-specific strategy with a 15 year offset (using $P_a(I)$ to determine her risk level) was found to be the most cost-effective. It can be observed that the preferred risk-specific strategy with a 15 year offset is more cost-effective than the less costly and less effective traditional strategies P3;D and P3;P, and that implementing the most costly and most effective traditional strategy P1;D is not cost-effective.

<table>
<thead>
<tr>
<th>Strategy</th>
<th>Total Costs</th>
<th>Total QALYs</th>
<th>Incremental Costs</th>
<th>Incremental QALYs</th>
<th>ICER$^*$</th>
<th>Comparison Strategy</th>
</tr>
</thead>
<tbody>
<tr>
<td>P3;D</td>
<td>339,522,818</td>
<td>7,909,960</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P3;P</td>
<td>355,471,711</td>
<td>7,910,782</td>
<td>15,948,893</td>
<td>821</td>
<td>19,419</td>
<td>P3;D</td>
</tr>
<tr>
<td>Risk-specific, $i = 0$</td>
<td>365,665,387</td>
<td>7,916,520</td>
<td>10,193,676</td>
<td>5,738</td>
<td>1,776</td>
<td>P3;P</td>
</tr>
<tr>
<td>Risk-specific, $i = 5$</td>
<td>374,110,045</td>
<td>7,917,358</td>
<td>8,444,658</td>
<td>838</td>
<td>10,080</td>
<td>Risk-specific, $i = 0$</td>
</tr>
<tr>
<td>Risk-specific, $i = 10$</td>
<td>384,523,876</td>
<td>7,917,683</td>
<td>10,413,831</td>
<td>325</td>
<td>32,033</td>
<td>Risk-specific, $i = 5$</td>
</tr>
<tr>
<td>Risk-specific, $i = 15$</td>
<td>396,209,739</td>
<td>7,917,966</td>
<td>11,685,863</td>
<td>284</td>
<td>41,205</td>
<td>Risk-specific, $i = 10$</td>
</tr>
<tr>
<td>Risk-specific, $i = 20$</td>
<td>404,244,366</td>
<td>7,917,969</td>
<td>8,034,627</td>
<td>3</td>
<td>3,213,851</td>
<td>Risk-specific, $i = 15$</td>
</tr>
<tr>
<td>P1;D</td>
<td>523,363,559</td>
<td>7,919,547</td>
<td>127,153,820</td>
<td>1,581</td>
<td>80,447</td>
<td>Risk-specific, $i = 15$</td>
</tr>
</tbody>
</table>

Table 3.13 Cost-Effectiveness Analysis for a Mixed-Risk Population
Preferred strategy is shown in bold, considering a cost-effectiveness threshold of $50,000/QALY. The values presented are the average values across all the replications.
Table 3.14 shows the way in which the risk-specific screening strategy with a 15 year offset is implemented, using the $P_{a15}(I)$ (Table 3.2) and the results from Table 3.12 to classify women by risk levels and screen accordingly. For women over 85, we found the absence of screening to be most cost-effective.

<table>
<thead>
<tr>
<th>Age</th>
<th>Risk Level 15 yr. earlier</th>
<th>Screening Strategy</th>
</tr>
</thead>
<tbody>
<tr>
<td>≤ 28</td>
<td>Low</td>
<td>P3;D</td>
</tr>
<tr>
<td>29-40</td>
<td>High</td>
<td>P1;D</td>
</tr>
<tr>
<td>41-44</td>
<td>Medium</td>
<td>P2;D</td>
</tr>
<tr>
<td>45-54</td>
<td>High</td>
<td>P1;D</td>
</tr>
<tr>
<td>55-74</td>
<td>Medium</td>
<td>P2;D</td>
</tr>
<tr>
<td>75-84</td>
<td>Low</td>
<td>P3;D</td>
</tr>
<tr>
<td>≥ 85</td>
<td>Low</td>
<td>No screen</td>
</tr>
</tbody>
</table>

Table 3.14 Screening Recommendations, by Age
Using a Risk-Specific Strategy with $i = 15$

3.5 Sensitivity Analysis

In our base case analysis, we assumed that all women have 100% adherence to screening appointments, and that the population is not vaccinated against HPV infections. In our sensitivity analysis we relaxed these assumptions and evaluated the performance of the traditional and risk-specific screening strategies.

3.5.1 Model Including Vaccine

In 2006, the U.S. Food and Drug Administration approved the quadrivalent vaccine Gardasil® for women 9 to 26 years of age to protect against infection with HPV types 6, 11, 16, and 18 (U.S. Department of Health and Human Services 2006), of which only types 16 and 18 are oncogenic;
types 6 and 8 cause genital warts. The costs associated with the vaccine are $120 for each of the three doses, plus administration fees (Elbasha et al. 2007). These vaccination costs are incurred independently of the screening strategy used.

We evaluated the performance of traditional and risk-specific screening strategies in a population of women vaccinated against infection with the oncogenic HPV types 16 and 18 (HPV types 6 and 11 are non-oncogenic). Our objective was to study the effect that a vaccinated population would have on the cost-effectiveness of screening strategies, including risk-specific strategies. Our objective was not to evaluate the cost-effectiveness of having an HPV vaccination program; this has already been done (Kulasingam and Myers 2003, Bergeron et al. 2008, Chesson et al. 2008) and it has been found to be cost-effective.

We modified our mixed risk population model by adjusting the 6-month infection probabilities to reflect the fact that vaccinated women will not contract infections with HPV 16 and 18, which account for 14% of oncogenic HPV infections (Kahn et al. 2007). We also adjusted the lesion progression transition probabilities to reflect the fact that HPV 16 and 18 cause 70% of cervical cancer cases (Munoz et al. 2004, Clifford et al. 2006). To model the lesion progression of a vaccinated population, we would need data on the disease progression rates of all oncogenic strains other than 16 and 18. These data are not readily available (Sanders and Taira 2003, Elbasha et al. 2007, McLay et al. 2010). To represent this, we adjusted the lesion progression transition probabilities from our original model such that the total number of cancer cases is reduced by 70% and the distribution of cancer cases by age continues to be similar to the SEER distribution (McLay et al. 2010). We did this adjustment by using the software system OptQuest in combination with Arena (Rockwell Software) to vary the lesion progression transition probabilities in a simulated population modeled to be 100% vaccinated. The transition
probabilities for lesion progression were lowered, to simulate the effect of the subset of oncogenic HPV types other than 16 and 18, which are less likely to progress to precancerous lesions and cervical cancer (Munoz et al. 2004, Clifford et al. 2006). While lowering the lesion progression transition probabilities, the number of cancer cases was constrained to result in a reduction of approximately 70%, and the percentages of cancer cases for each age group were constrained to remain in the same relative order and close to the base case output, in order to maintain a distribution of cancer cases similar to that of SEER. OptQuest guides a series of simulations, searching for input values using scatter search, tabu search, and neural networks, and evaluating whether the outputs satisfy the objectives and constraints defined (Glover et al. 1999).

We evaluated traditional and risk-specific screening strategies in a population in which all women are vaccinated, and in a population with the current vaccine coverage of approximately 33%, considering women who have received all three doses of the vaccine (CDC 2009). We assumed that vaccinated women have full immunity against the strains covered by the vaccine (Garland et al. 2007) and that immunity does not wane over time. The risk-specific strategies were designed as described in Section 3.3.2: for $P_d(I) \leq 0.03$, strategy P3;D; for $0.04 \leq P_d(I) \leq 0.05$, strategy P2;D; and for $P_d(I) \geq 0.06$, strategy P1;D; using each woman’s $P_d(I)$ from $i$ years prior to the time of screening to determine her risk level.
3.5.2 Model Including Vaccine Results

Table 3.15 shows the cost-effectiveness analysis of the non-dominated screening strategies in a population with 100% vaccination. A risk-specific strategy with a 10 year offset was found to be the most cost-effective.

<table>
<thead>
<tr>
<th>Strategy</th>
<th>Total Costs</th>
<th>Total QALYs</th>
<th>Incremental Costs</th>
<th>Incremental QALYs</th>
<th>ICER*</th>
<th>*Comparison Strategy</th>
</tr>
</thead>
<tbody>
<tr>
<td>P3;D</td>
<td>248,415,926</td>
<td>8,005,261</td>
<td>11,627,727</td>
<td>1,499</td>
<td>44,096</td>
<td>P3;D</td>
</tr>
<tr>
<td>P3;P</td>
<td>260,043,654</td>
<td>8,005,487</td>
<td>14,342,007</td>
<td>4,235</td>
<td>31,519</td>
<td>P3;D</td>
</tr>
<tr>
<td>Risk-specific, i = 0</td>
<td>314,493,429</td>
<td>8,006,759</td>
<td>36,092,007</td>
<td>7,800</td>
<td>47,800</td>
<td>Risk-specific, i = 0</td>
</tr>
<tr>
<td>P2;D</td>
<td>322,137,114</td>
<td>8,007,048</td>
<td>35,013,007</td>
<td>7,000</td>
<td>23,400</td>
<td>P2;D</td>
</tr>
<tr>
<td>Risk-specific, i = 5</td>
<td>333,101,986</td>
<td>8,007,208</td>
<td>37,092,007</td>
<td>8,000</td>
<td>45,000</td>
<td>P2;D</td>
</tr>
<tr>
<td>Risk-specific, i = 10</td>
<td>350,072,890</td>
<td>8,008,079</td>
<td>43,092,007</td>
<td>10,000</td>
<td>43,000</td>
<td>P2;D</td>
</tr>
</tbody>
</table>

Table 3.15 Cost-Effectiveness Analysis for a Population with 100% Vaccination
Preferred strategy is shown in bold, considering a cost-effectiveness threshold of $50,000/QALY. The values presented are the average values across all the replications.

Table 3.16 illustrates how the risk-specific screening strategy with a 10 year offset is implemented.

<table>
<thead>
<tr>
<th>Age</th>
<th>Risk Level 10 yr. earlier</th>
<th>Screening Strategy</th>
</tr>
</thead>
<tbody>
<tr>
<td>≤ 23</td>
<td>Low</td>
<td>P3;D</td>
</tr>
<tr>
<td>24-35</td>
<td>High</td>
<td>P1;D</td>
</tr>
<tr>
<td>36-39</td>
<td>Medium</td>
<td>P2;D</td>
</tr>
<tr>
<td>40-49</td>
<td>High</td>
<td>P1;D</td>
</tr>
<tr>
<td>50-69</td>
<td>Medium</td>
<td>P2;D</td>
</tr>
<tr>
<td>70-84</td>
<td>Low</td>
<td>P1;D</td>
</tr>
<tr>
<td>≥ 85</td>
<td>Low</td>
<td>No screen</td>
</tr>
</tbody>
</table>

Table 3.16 Screening Recommendations for a 100% Vaccinated Population, by Age
Using a Risk-Specific Strategy with i = 10
For a population with the current HPV vaccine coverage of approximately 33% (CDC 2009), we evaluated the following different screening approaches:

A. Screen vaccinated and non-vaccinated subpopulations with the same screening strategy, considering only the traditional strategies.

B. Screen vaccinated and non-vaccinated subpopulations differently, considering only the traditional strategies.

C. Screen the vaccinated and non-vaccinated subpopulations differently, considering risk-specific screening strategies.

Approaches B and C allow flexibility to screen vaccinated and non-vaccinated populations differently; these two populations are at very different risks for cervical cancer.

Table 3.17 shows the cost-effectiveness analysis of the non-dominated screening approaches in a population with 33% vaccination coverage. Approach C was found to be the most cost-effective, which uses the risk-specific strategy that is preferred for each of the subpopulations (vaccinated and non-vaccinated).

<table>
<thead>
<tr>
<th>Approach</th>
<th>Total Costs</th>
<th>QALYs</th>
<th>Incremental Costs</th>
<th>QALYs</th>
<th>ICER*</th>
<th>Comparison</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>354,342,528</td>
<td>7,945,082</td>
<td>26,642,051</td>
<td>2,621</td>
<td>10,163</td>
<td>A</td>
</tr>
<tr>
<td>B</td>
<td>380,984,579</td>
<td>7,947,703</td>
<td>75,974,253</td>
<td>719</td>
<td>105,705</td>
<td>C</td>
</tr>
</tbody>
</table>

Table 3.17 Cost-Effectiveness Analysis for a Population with 33% Vaccination
Preferred strategy is shown in bold, considering a cost-effectiveness threshold of $50,000/QALY. The values presented are the average values across all the replications. The preferred risk-specific approach C consists of screening the vaccinated subpopulation with $i=10$ and the non-vaccinated subpopulation with $i=15$. 
3.5.3 Model Including Adherence

In our base case model we assume that all women attend all screening appointments, which is usually not the case. Paskett et al. (2010) found the percentage of the population that is within risk-appropriate guidelines to be 68%. Additionally, it is well known that women at higher risk for HPV and cervical cancer are always less likely to be adhering to screening guidelines (Marcus et al. 1992, Basen-Engquist et al. 2003, Eggleston et al. 2007, Paskett et al. 2010).

To represent imperfect adherence in our model in a way that reflects these findings, we would need data on the probability distribution of adherence, conditioned on different probabilities of infection. Since these data are not readily available, we made rough estimates of them that are consistent with the literature on adherence to cervical cancer screening, in the following manner.

We classified the mixed population by $P_a(I)$ (age-specific 6-month probability of infection) into low ($\leq 0.03$), medium (0.04-0.05), and high risk (≥0.06). These levels were defined according to the risk levels defined in Table 3.12. We made estimates of probabilities of adherence conditioned on each risk level such that the probability of adherence of the overall population was approximately 70%, and that the probability of adherence was inversely proportional to the probability of infection. We studied three different estimates of this probability distribution, as shown in Table 3.18. We included Estimate 3 to explore the effect of adherence being independent of risk of infection, although we know from the literature that this is not the case in the US.
\begin{table}
\centering
\begin{tabular}{|c|c|c|c|}
\hline
$P_a(I)$ & $\leq 0.03$ & $0.04 - 0.05$ & $\geq 0.06$ \\
\hline
Estimate 1 & 0.90 & 0.75 & 0.50 \\
Estimate 2 & 0.80 & 0.70 & 0.60 \\
Estimate 3 & 0.70 & 0.70 & 0.70 \\
\% of Population & 25.3\% & 48.5\% & 26.3\% \\
\hline
\end{tabular}
\caption{Estimated Probability Distributions of Adherence Conditioned on $P_a(I)$}
\end{table}

In our model these probabilities of adherence represent the probability that each patient will attend screening, at each instance in which she has a screening event programmed (according to the screening strategy and to her previous test results).

For each of our three estimates of adherence probability distributions, we simulate hypothetical homogeneous populations with constant probabilities of infection ranging from 0.01 to 0.50 as described in Section 3.2.4. We evaluate the performance of the 12 traditional screening strategies in each of these homogeneous populations.

\textbf{3.5.4 Model Including Adherence Results}

The results of the homogeneous populations for our three estimates of imperfect adherence are shown in Table 3.19. For each estimate, the most cost-effective strategy varies by probability of infection, as it did in our base case results considering perfect adherence. However, in this case we observe only two risk levels. Estimates 1 and 2 result in the same risk breakpoints and preferred screening strategies, which are the same strategies found to be most cost-effective for the low and high risk levels in our base case.
We next simulate a mixed-risk population with an age-specific $P_a(I)$ that changes over time (as described in Section 3.4), but with the probability distribution of imperfect adherence from Estimate 1. We evaluate the traditional screening strategies and a risk-specific strategy composed of the preferred strategies for Estimate 1 shown in Table 3.19.

Table 3.20 shows the cost-effectiveness analysis of the non-dominated screening strategies in a population with the probability distribution of imperfect adherence of Estimate 1. The risk-specific strategy with a 10 year offset was found to be the most cost-effective.
3.6 Discussion

We evaluated the performance of screening strategies on hypothetical homogeneous populations at different levels of risk for oncogenic HPV infection. We found that the strategies found to be most cost-effective varied for homogeneous populations of different levels of risk, with higher risk populations benefiting from more aggressive strategies. Based on these results, we designed an age-based risk-specific screening strategy composed of three traditional strategies, which when applied to a mixed-risk population was found to be more cost-effective than any of the traditional strategies. In sensitivity analyses, a risk-specific screening strategy was also found to be most cost-effective in a population with imperfect adherence and in a vaccinated population. We consider these findings to be initial evidence of the potential benefits of risk-specific screening policies.

Our findings are largely due to the fact that HPV is a necessary cause for cervical cancer. Therefore, varying the probability of HPV infection results in women being at different risks for cervical cancer. Women that are at higher risks for cervical cancer benefit from more aggressive screening, and the number of cancers that can be prevented with screening in a high-risk population is higher than in low-risk populations. Women that are at lower risks for cervical cancer can be well served with less aggressive screening. If all women are being screened with the same policy, it is likely that high-risk women are being under-screened and low-risk women are being over-screened. Therefore, a risk-specific screening policy can be expected to be more cost-effective than those which make no distinction among women with different risk levels.

We also find that applying a 10 or 15 year offset (depending on whether vaccination and imperfect adherence are being considered) to the age-specific probability of infection used to
classify each woman’s risk level results in the most cost-effective strategies. This is consistent with the fact that precancerous lesions and cervical cancer generally take years to develop (Liaw et al. 1999, Moscicki et al. 2006, Trottier et al. 2009).

Previous studies of the cost-effectiveness of different screening strategies to prevent cervical cancer have not considered risk-specific screening alternatives; they have evaluated different screening strategies when implemented uniformly to a population. Past analyses of triage strategies for ASCUS Pap results have found triage with the HPV DNA test to be the most cost-effective strategy (Kim et al. 2002, Kulasingam et al. 2006), we also found triage with the HPV DNA test to be the preferred strategy at all frequencies of primary screening with the Pap test. McLay et al. (2010) designed dynamic, age-based screening policies, which they found to be more cost-effective than traditional screening strategies. This study and findings are closely related and consistent with ours, with some differences. For their dynamic strategies, they varied how many lifetime screenings to perform and at what ages, but they only considered primary screening with the Pap test and triage screening with direct colposcopy. Their disease model does not include an Infected state, as they consider having an HPV infection to be equivalent to having CIN 1 (McLay et al. 2010).

A strength of our model is that we use data from the ALTS trial to estimate the effectiveness of screening tests, including colposcopy. Previous studies that have evaluated the cost-effectiveness of cervical cancer screening strategies (Sanders and Taira 2003, Goldie et al. 2004, Elbasha et al. 2007, Bergeron et al. 2008) have assumed colposcopy to have perfect or almost perfect sensitivity to detect pre-cancerous lesions, which is not the case. Kulasingam et. al (2006) analyzed data from the ALTS trial and recommend that “the less than perfect sensitivity of colposcopy needs to be accounted for in future clinical guidelines and policy analyses”.

73
Our study’s main limitation is the fact that hypothetical homogeneous populations with a probability of infection which remains constant throughout each woman’s lifetime were used to determine the most cost-effective screening strategies and the risk level breakpoints. These populations are unrealistic; it is well known that the probability of HPV infection changes greatly throughout most women’s lifetime. Our purpose in using them was to explore whether different screening strategies would be most cost-effective for groups of women with different levels of risk, and these homogeneous, constant risk populations were extreme cases that were convenient to analyze. We mitigated the use of these hypothetical populations by then modeling a realistic mixed-risk population (validated using SEER data) with age-specific probabilities of infection. We applied the risk level breakpoints and screening strategies selected from the homogeneous populations to the mixed risk population in the form of risk-specific screening strategies, which were found to be more cost-effective than the current strategies (Table 3.13). This risk-specific strategy is also consistent with what we know about the disease progression being slow, recommending that the $P_a(I)$ of 10 or 15 years prior be used to determine each woman’s risk level for screening purposes. This, and the superior performance of the risk-specific strategies when applied to a realistic population indicates that the results obtained from the hypothetical homogeneous populations are worth using. However, we cannot claim our findings to be optimal, because exploring numerous alternatives for breakpoints and screening strategies directly in mixed-risk populations may result in the design of risk-specific strategies that are superior to the one that we propose.

Another limitation of our study is that in our simulated populations with varying levels of risk, we use a probability of oncogenic HPV infection, $P_a(I)$, that is conditioned only on each woman’s age. Although age is a good indicator of the risk for HPV infection, to fully capture the effect of
risk-specific screening, it would be preferable to use a probability of infection that is conditioned on each woman’s behavioral and demographic characteristics in addition to her age. Additionally, recent guidelines already make some age distinctions (e.g. over and under 30), so our analysis using age-specific $P_a(I)$ is not entirely novel. However, we make more age group distinctions and evaluate the use of age offsets to classify risk levels.

Additionally, some of our simplifying assumptions regarding imperfect adherence are most likely unrealistic. We assume that the probability of a patient deciding to attend screening at each period is independent of whether she attended screening the previous periods, and independent of her past diagnoses. This may be unrealistic because a woman who does not attend regularly may have a different probability of attending an appointment if her last appointment was two years ago rather than five years ago. Also, a woman who has received a serious diagnosis (such as high-grade lesion) may be more likely to attend follow up screening. We make these simplifying assumptions because of a lack of more specific data. Another limitation of our study is that we do not explore populations that have both vaccination and imperfect adherence.

Another limitation is the unavailability and inconsistency of some of the input data, in particular those which refer to the prevalence of lesions and the accuracy of screening tests. These data were estimated using tests that are far from perfect, and even limiting our use to carefully conducted studies, we found discrepancies in the data reported. This resulted in our not being able to constrain some values in our input data to be between the reported upper and lower bounds, while satisfying the laws of probability. Data on the probability distribution of patient adherence conditioned on $P_a(I)$ are also not available, so that the data that we use to model imperfect adherence that is inversely correlated with $P_a(I)$ were very rough estimates based on the findings that support this inverse correlation.
A next step in the study of risk-specific screening is to simulate a population of women whose probability of infection is conditioned on a set of behavioral and demographic characteristics that are known risk factors for HPV (i.e. age, marital status, number of sex partners). To calculate this probability of infection conditioned on patient attributes, we have built a logistic regression model that uses NHANES data (see Chapter 4 of this dissertation). We use the patient’s demographic and sexual behavior characteristics as predictor variables to estimate the probability of a patient being infected with an oncogenic strain of HPV. Future research can also study the design and evaluation of risk-specific screening strategies for other diseases for which this approach is warranted.

Our work is the first to date that explores the possibility of risk-specific screening for cervical cancer prevention. This approach is consistent with the existing literature on known risk factors for cervical cancer and HPV infection, and would be easily implemented. We use age to estimate each patient’s risk; age is already considered in screening decisions. However, current screening guidelines make very few screening distinctions based on age (e.g. over or under 30) and do not explicitly consider time offsets to account for the slow disease progression.

We conclude that the design of risk-specific screening strategies has the potential to be more cost-effective and to result in fewer instances of over-screening and under-screening, which seems to be aligned with the recommendations issued in recent guidelines (ACOG 2009).
3.7 References for Chapter 3


77


Chapter 4: Logistic Regression Analysis to Predict Oncogenic HPV Infection Probability

4.1 Introduction

Infection with oncogenic strains of human papillomavirus (HPV) has been identified as a necessary cause of cervical cancer (Schiffman and Castle 2003, Muñoz 2000). HPV is a sexually transmitted infection (STI), and is considered the most common STI in the United States (Centers for Disease Control and Prevention 2011). The prevalence of HPV infection has been found to be 26.8% among US women of ages 14 to 59 (Dunne et al. 2007), and 15.6% when restricted to known oncogenic strains (Kahn et al. 2007). There are known behavioral and demographic characteristics that influence a woman’s risk of contracting an HPV infection, such as the number of lifetime and recent sex partners, marital status, and a past history of STIs (Kahn et al. 2007, Dunne et al. 2007, Sasagawa et al. 2005, Wright et al. 2004, Sellors et al. 2003, Winer et al. 2003, Moscicki et al. 2001). Questions persist as to the predictive capabilities of models incorporating known risk factors.

Cervical cancer can be prevented very effectively with appropriate screening programs, which include the use of the Papanicolau and of the HPV DNA tests. Improved quantitative assessments of risk-specific screening programs can be facilitated by the construction of predictive models. In the case of cervical cancer, using information about each patient’s risk factors to estimate the probability of having an oncogenic HPV infection can help to model and evaluate risk-specific
screening strategies. In this paper we investigate risk factors associated with oncogenic HPV infection and estimate the probability of a patient being infected, based on her behavioral and demographic characteristics. In particular, we present the results of a logistic regression analysis of data from the 2003-2004 National Health and Nutrition Examination Survey conducted to quantify associations between a patient’s characteristics and her test results for an oncogenic HPV infection, in order to build a predictive multivariate logistic regression model.

4.2 Methods

Our study population is drawn from the National Health and Nutrition Examination Survey (NHANES), which is conducted by the National Center for Health Statistics (NCHS). NHANES collects health-related information by interviewing a representative sample of the US noninstitutionalized civilian population, oversampling people over 60 years old and minorities. NHANES also reports results from physical examinations and laboratory tests, including the HPV DNA test. Details of the survey content and of the data collection and data analysis methodologies are described by the NCHS (2011).

Our study includes 1120 women from the 2003-2004 NHANES cycle, aged 20 to 59. These women answered the demographic, smoking and sexual behavior questionnaires and submitted a self-collected vaginal swab specimen for HPV testing. Details of the specimen collection and HPV genotyping are provided in Dunne et al. (2007) and Kahn et al. (2007). We excluded from the analyses women who submitted an inadequate swab or indicated in the sexual behavior questionnaire that they had never had intercourse. We excluded from the multivariate analysis women who had missing values for any of the variables considered, resulting in a sample size for the multivariate analysis of 1074 women.
We performed a logistic regression analysis to examine the association between patients’ demographic and behavioral characteristics and results of a test for an oncogenic HPV infection. We selected the independent variables included in the model as predictors based on the existing literature about known risk factors for HPV infection (Kahn et al. 2007, Dunne et al. 2007, Sasagawa et al. 2005, Wright et al. 2004, Sellors et al. 2003, Winer et al. 2003, Moscicki et al. 2001). The independent variables consisted of six variables from the Demographics Questionnaire, six variables from the Sexual Behavior Questionnaire, and two variables from the Smoking Questionnaire. The demographic variables were: age, education level, country of birth, race, below poverty index, and marital status, and the variables from the Sexual Behavioral Questionnaire were: age at first intercourse, number of lifetime sex partners, number of recent sex partners (in the past 12 months), history of chlamydia, genital herpes and genital warts. We did not include the variable gonorrhea from the sexual behavior questionnaire in the analyses because there were no observations in our sample of a past history of gonorrhea. We combined two variables from the Smoking Questionnaire into one variable defined as: smoked at least 100 cigarettes in lifetime; if ‘yes’, current smoker; if ‘yes’, former smoker.

The NCHS provides weights to use in statistical analyses of NHANES data to account for the unequal probabilities of selection and to adjust for nonresponse. Because we use data from the MEC Laboratory Component: HPV, we used the medical examination weights provided by the NCHS. Since our multivariate analysis excluded women who had missing responses to any of our variables, we recalculated the medical examination weights to adjust for the women that fell outside the sample of 1074 women that we analyzed. However, we found no significant difference (α=0.05) between the odds ratios (OR) estimates obtained using the adjusted weights.
and those obtained using the original medical examination weights. As a result, we conducted the analyses using the medical examination weights provided by the NCHS.

In the univariate analysis (n=1120), the Wald $\chi^2$ statistic was used to identify variables for inclusion in the multivariate analysis. Odds ratios were used to measure the magnitude of the associations, with each descriptor variable considered to be significantly associated with the dependent variable if the value 1 was not included in the 95% confidence interval of the OR. Variables that were statistically significant at the $p<0.20$ level in the univariate analysis were included in the preliminary multivariate model, even if the confidence interval of their odds ratios included the value 1. For the multivariate analysis (n=1074), the data for 834 women were randomly selected and used to develop the multivariate predictive model, and the remaining 240 were used for validation. There was no significant difference ($\alpha=0.05$) found between the parameter estimates obtained using the sample size of 834 and of those obtained using the full sample (n=1074) to build the multivariate model. This does not appear to depend on the specific sample that was selected, since it was observed in four different random samples of 834 women. All pairwise interactions of variables included in the multivariate model were evaluated.

To validate the predictive model, we used the final multivariate model to estimate the probability of infection for the validation sample (n=240). We entered these estimated probabilities as a continuous variable in a univariate logistic regression model with oncogenic HPV infection as the outcome. Goodness-of-fit for this model and for our final multivariate model was assessed using the Archer and Lemeshow F-adjusted mean residual goodness-of-fit test, which is designed to be used with survey data. All analyses were performed using Intercooled Stata 9.2 (StataCorp, College Station, Texas).
4.3 Results

Univariate Analysis

Table 4.1 shows the variables that were found to be significantly associated with oncogenic HPV infection at the p<0.20 level in the univariate analysis, and their unadjusted odds ratios. These variables are: age, marital status, country of birth, below poverty index, smoking status, age at first intercourse, number of lifetime sex partners, number of recent sex partners, and a history of genital warts. For example, women who have never been married and are not currently living with a partner are more likely to have an oncogenic HPV infection compared to married women (OR: 2.43; 95% CI: 1.34-4.04); and women who have had five or more partners in their lifetime are more likely to have an oncogenic HPV infection compared to women who have had less than 5 lifetime partners (OR: 1.76; 95% CI: 1.01-3.06).
<table>
<thead>
<tr>
<th>Characteristics</th>
<th>No.</th>
<th>OR (95% CI)</th>
<th>p Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20-24 (reference)</td>
<td>153</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>25-29</td>
<td>154</td>
<td>0.54 (0.29-1.01)</td>
<td>0.053</td>
</tr>
<tr>
<td>30-59</td>
<td>813</td>
<td>0.42 (0.20-0.85)</td>
<td>0.020</td>
</tr>
<tr>
<td>Marital status</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Married (reference)</td>
<td>600</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>Widowed, divorced, separated</td>
<td>204</td>
<td>1.61 (0.85-3.03)</td>
<td>0.131</td>
</tr>
<tr>
<td>Never married</td>
<td>223</td>
<td>2.43 (1.34-4.04)</td>
<td>0.006</td>
</tr>
<tr>
<td>Living with partner</td>
<td>93</td>
<td>1.18 (0.51-2.71)</td>
<td>0.675</td>
</tr>
<tr>
<td>Country of birth</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>US (reference)</td>
<td>909</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>Mexico</td>
<td>128</td>
<td>0.94 (0.48-1.83)</td>
<td>0.834</td>
</tr>
<tr>
<td>Other</td>
<td>83</td>
<td>0.31 (0.09-1.02)</td>
<td>0.053</td>
</tr>
<tr>
<td>Below poverty index*</td>
<td>198</td>
<td>1.37 (0.92-2.05)</td>
<td>0.113</td>
</tr>
<tr>
<td>Smoked at least 100 cigarettes</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No (reference)</td>
<td>650</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>Yes and former smoker</td>
<td>200</td>
<td>1.18 (0.59-2.37)</td>
<td>0.612</td>
</tr>
<tr>
<td>Yes and current smoker</td>
<td>268</td>
<td>1.32 (0.89-1.96)</td>
<td>0.160</td>
</tr>
<tr>
<td>Age at first intercourse</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt; 16 (reference)</td>
<td>305</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>≥ 16</td>
<td>810</td>
<td>1.44 (0.95-2.18)</td>
<td>0.078</td>
</tr>
<tr>
<td>Lifetime partners</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 (reference)</td>
<td>217</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>2-4</td>
<td>325</td>
<td>0.98 (0.58-1.65)</td>
<td>0.925</td>
</tr>
<tr>
<td>≥ 5</td>
<td>578</td>
<td>1.76 (1.01-3.06)</td>
<td>0.045</td>
</tr>
<tr>
<td>Recent partners</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt; 2 (reference)</td>
<td>953</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>≥ 2</td>
<td>167</td>
<td>1.48 (0.86-2.56)</td>
<td>0.146</td>
</tr>
<tr>
<td>History of genital warts</td>
<td>74</td>
<td>1.65 (0.76-3.59)</td>
<td>0.193</td>
</tr>
</tbody>
</table>

Table 4.1 Univariate Association Between HPV and Selected Demographic and Behavioral Variables

*Poverty index threshold based on family income and adjusted for family size.

**Multivariate Analysis**

The final multivariate model includes the variables that were significant in the univariate analysis, excluding the ones that became non-significant in the multivariate analysis. Each variable that was not significant in the univariate analysis was added to the multivariate model to see if the
parameter estimates or the goodness-of-fit of the model changed due to confounding. The variable below poverty index was included in the final multivariate model because it changed parameter estimates by more than 20% and resulted in an improved model fit. Interaction terms were not found to be significant at the 0.05 level. The final multivariate model and the adjusted odds ratios are shown in Table 4.2.

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>OR (95% CI)</th>
<th>Coefficient</th>
<th>p Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20-24 (reference)</td>
<td>1.00 (1.0-1.0)</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>25-29</td>
<td>0.33 (0.14-0.82)</td>
<td>-1.094</td>
<td>0.020</td>
</tr>
<tr>
<td>29-59</td>
<td>0.36 (0.18-0.72)</td>
<td>-1.029</td>
<td>0.007</td>
</tr>
<tr>
<td>Marital status</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Married or living with partner</td>
<td>1.00 (1.0-1.0)</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Not married or living with partner</td>
<td>1.72 (1.06-2.79)</td>
<td>0.541</td>
<td>0.031</td>
</tr>
<tr>
<td>Below poverty index *</td>
<td>1.86 (0.94-3.67)</td>
<td>0.621</td>
<td>0.070</td>
</tr>
<tr>
<td>Lifetime partners &lt; 5 (reference)</td>
<td>1.00 (1.0-1.0)</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Lifetime partners ≥ 5</td>
<td>1.78 (1.04-3.04)</td>
<td>0.577</td>
<td>0.037</td>
</tr>
<tr>
<td>Intercept (constant)</td>
<td>NA</td>
<td>-1.880</td>
<td></td>
</tr>
</tbody>
</table>

Table 4.2 Predictor Variables Included in the Final Multivariate Model
* Poverty index threshold based on family income and adjusted for family size.

The predictive equation was calculated with the following parameters:

\[
P(I) = \frac{1}{1 + e^{-z}}, \text{ where } z = -1.880 + [ -1.094 \text{ (if age 25-29)} - 1.029 \text{ (if age 29-59)} + 0.541 \text{ (if not married and not living with partner)} + 0.621 \text{ (if below poverty index)} + 0.577 \text{ (if lifetime partners ≥ 5)} ].
\]

Goodness of fit for the final multivariate model was assessed using the Archer and Lemeshow F-adjusted mean residual goodness-of-fit test, which did not indicate a lack of fit between the data and the model (p=0.852). The univariate model used for validation in a separate data set consisted
of the estimated probabilities of infection as a continuous independent variable and the oncogenic HPV infection as the outcome. The fit of this validation model was also assessed and the Archer and Lemeshow F-adjusted mean residual goodness-of-fit test did not indicate a lack of fit between the data and the model (p=0.762). A description of the Archer and Lemeshow F-adjusted mean residual goodness-of-fit test and an illustration of its results using our validation dataset are provided in the Supplement for Chapter 4.

4.4 Discussion

We found that a predictive relationship exists between demographic and behavioral characteristics and a woman’s likelihood of having an oncogenic HPV infection. Of particular importance are the risk factors younger age (under 25), marital status of single and not living with partner, and a higher number of lifetime partners (five or more).

Our logistic regression model quantifies the relationship between demographic and behavioral characteristics and the likelihood of having an oncogenic HPV infection, and provides a multivariate prediction model that can be used to estimate the probability of a patient having an oncogenic HPV infection. Our final multivariate model was used to estimate the probability of infection in a separate validation sample, and the Archer and Lemeshow F-adjusted mean residual goodness-of-fit test did not indicate a lack of fit between the validation sample data and the predictive model.

It is important to note that our focus is on estimating probabilities of oncogenic HPV infection, which can be used as a key input parameter to model a population for the prospective analysis of risk-specific cervical cancer screening strategies. Our goal is not to classify whether a woman is
or is not infected, but to obtain a probabilistic representation of the population that can be used for modeling for quantitative analysis.

In performing our analyses, several modeling choices were required. In the multivariate analysis, we found that defining the variable *Marital Status* as binary (married or living with partner, not married or living with partner), resulted in an improved level of significance and model fit. The variable *Lifetime Partners* was not found to be significant as a continuous variable, but it was significant when defined as a binary variable to distinguish between having had more or less than five partners. Similarly, the variable *Recent Partners* was only found to be significant in the univariate analysis as a binary variable to distinguish between having had 1, or 2 or more partners in the past year. The variables *Age* and *Age at First Intercourse* were also found to be significant only when defined as categorical variables.

There are several limitations to our study. Some limitations that refer to the collection of the NHANES data that we use, such as the differences between nonresponders and responders, and the fact that the vaginal swab was self-collected have been described previously (e.g. Dunne et al. 2007). In addition to this, the data from the sexual behavior questionnaire for respondents under 20 years old were not available to the general public, so our analysis is limited to women aged 20 to 59. We only analyzed the variables most commonly associated with HPV infection. There are other factors which can affect the immune system (e.g., nutrition and stress) and the exposure to the HPV virus (e.g., time in a new relationship before intercourse, HPV status of the partner) which are not included in our analysis.

There are differences and similarities in our results and those obtained in previous studies. Dunne et al. (2007) analyzed data from women that had been tested for HPV infections and had
answered the demographic and sexual behavior questionnaires in the 2003-2004 NHANES (the same data set that we studied, but including data for women ages 14 through 19, to which they had access). They report the prevalence of any HPV infection (i.e., oncogenic and non-oncogenic types) by demographic and behavioral characteristics. Dunne et al. (2007) found younger age, unmarried marital status, and increasing number of lifetime and recent partners to be significant in their multivariate model (n=1288). We analyzed almost the same data set but did not find the number of recent partners to be significant in our multivariate model. This difference may be due in part to the fact that the dependent variable for Dunne et al. (2007) was “any HPV infection”, whereas we studied only oncogenic HPV infections. Kahn et al. (2007) also analyzed data from the 2003-2004 NHANES and report the demographic factors associated with oncogenic HPV infection, but did not analyze sexual behavior variables. Kahn et al. (2007) found age, race and marital status to be significant predictors of oncogenic HPV infection when adjusting for all other covariates. Both of these previous studies were focused on identifying risk factors for HPV infection, whereas our focus was on developing a predictive model to estimate the probability of oncogenic HPV infection. A key difference between our work and these two previous studies is that their goal was not predictive or classification, but to estimate the prevalence and identify the factors associated with HPV infection.

Our results suggest that our predictive model can be used to estimate a patient’s probability of oncogenic HPV infection as a function of demographic and behavioral characteristics. This is relevant for estimating her risk of cervical cancer, and thus might prove useful in devising, analyzing, and implementing risk-specific screening strategies. Quantitative studies that are used for assessing cervical cancer screening strategies and policies (Goldie et al. 1999, Myers et al. 2000, Kim et al. 2002, Sanders & Taira 2003, Canfell et al. 2004, Kohli et al. 2006, Bergeron et al. 2008) include probabilistic representations of the population. Having a probabilistic predictive
model like this can support this type of analysis in investigating demographic and behavioral risk-specific screening strategies.
4.5 Supplement for Chapter 4

We use the Archer & Lemeshow F-adjusted mean residual goodness-of-fit test to evaluate the fit of our models. Because this test is fairly recent, we provide a description of it.

We begin with a description the traditional Hosmer-Lemeshow goodness of fit test. This test is designed to be used with data obtained from random sampling, and it is not intended to be applied to data from a survey sample (Archer & Lemeshow 2006). This test partitions the population into groups, usually 10, by the ordered probabilities of outcome occurrence (in our case, of oncogenic HPV infection). The first group contains the \( n_1 = n/10 \) subjects having the smallest estimated probabilities, and the last group contains the \( n_{10} = n/10 \) subjects having the largest estimated probabilities.

As an illustration, Table 4.3 depicts Stata output from a Hosmer-Lemeshow goodness of fit test applied to our validation sample. For each group, Obs 1 is the number of people in the group that were found to have an infection, and Exp 1 is the expected number of infections based on the calculated probabilities for each person in the group. In this case the test grouped the 1st and 2nd deciles, the 4th and 5th deciles, and the 7th and 8th deciles because of ties in the greatest probability value corresponding to each of these pairs of deciles.
In survey data, the sampling weight represents the number of people that each sampled observation represents in the total population. Multiplying each observation of a survey dataset by its sampling weight leads to a representation of the total population. Therefore, the sum of the number of people represented in the groups of the goodness-of-fit test should equal the total population size. (Archer & Lemeshow 2006).

Goodness of fit for our final multivariate model and for the validation sample were assessed using the Archer and Lemeshow F-adjusted mean residual goodness-of-fit test, which was developed to be used with data obtained by a survey sample design (Archer & Lemeshow 2006, Archer et al. 2007). In this test the groups are not formed based only on the ordered estimated probabilities of outcome occurrence. The sampling weights are also considered, so that each group represents one tenth of the people in the total population. The first group contains the smallest estimated probability values such that the sum of the weights of the observations correspond to one tenth of the population, the second group contains the next smallest estimated probability values such that the sum the sum of the weights of the observations correspond to one tenth of the population, and the 10th group contains the largest estimated probability values such that the sum of the weights of the observations correspond to one tenth of the population. An output in table form of the groups formed by the Archer & Lemeshow F-adjusted mean residual goodness-of-fit test is not

<table>
<thead>
<tr>
<th>Group</th>
<th>Probability</th>
<th>Obs 1</th>
<th>Exp 1</th>
<th>No. of people in group</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>0.1003</td>
<td>5</td>
<td>5.5</td>
<td>55</td>
</tr>
<tr>
<td>3</td>
<td>0.1177</td>
<td>4</td>
<td>3.3</td>
<td>28</td>
</tr>
<tr>
<td>5</td>
<td>0.1203</td>
<td>3</td>
<td>4.4</td>
<td>37</td>
</tr>
<tr>
<td>6</td>
<td>0.1475</td>
<td>5</td>
<td>4.6</td>
<td>34</td>
</tr>
<tr>
<td>8</td>
<td>0.1528</td>
<td>6</td>
<td>6.1</td>
<td>40</td>
</tr>
<tr>
<td>9</td>
<td>0.2023</td>
<td>6</td>
<td>4.7</td>
<td>27</td>
</tr>
<tr>
<td>10</td>
<td>0.3326</td>
<td>3</td>
<td>5.2</td>
<td>19</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Total sample: 240</td>
</tr>
</tbody>
</table>

Table 4.3 Stata Output from a Hosmer-Lemeshow Goodness-of-fit Test
provided by Stata. Table 4.4 shows a table that we constructed using our validation sample and
the description of the test.

<table>
<thead>
<tr>
<th>Group</th>
<th>Probability</th>
<th>Obs 1</th>
<th>Exp 1</th>
<th>Sum of group sampling weights</th>
<th>Proportion of the population</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.0517</td>
<td>4</td>
<td>1.8</td>
<td>1,237,937</td>
<td>0.097</td>
</tr>
<tr>
<td>2</td>
<td>0.0517</td>
<td>0</td>
<td>0.7</td>
<td>1,234,242</td>
<td>0.097</td>
</tr>
<tr>
<td>3</td>
<td>0.0834</td>
<td>3</td>
<td>1.2</td>
<td>1,268,338</td>
<td>0.100</td>
</tr>
<tr>
<td>4</td>
<td>0.0885</td>
<td>2</td>
<td>2.3</td>
<td>1,277,506</td>
<td>0.100</td>
</tr>
<tr>
<td>5</td>
<td>0.0885</td>
<td>3</td>
<td>2.0</td>
<td>1,300,908</td>
<td>0.102</td>
</tr>
<tr>
<td>6</td>
<td>0.0922</td>
<td>2</td>
<td>1.5</td>
<td>1,309,899</td>
<td>0.103</td>
</tr>
<tr>
<td>7</td>
<td>0.1429</td>
<td>4</td>
<td>4.6</td>
<td>1,282,764</td>
<td>0.101</td>
</tr>
<tr>
<td>8</td>
<td>0.1429</td>
<td>4</td>
<td>3.7</td>
<td>1,280,262</td>
<td>0.101</td>
</tr>
<tr>
<td>9</td>
<td>0.2076</td>
<td>5</td>
<td>3.1</td>
<td>1,313,138</td>
<td>0.103</td>
</tr>
<tr>
<td>10</td>
<td>0.3359</td>
<td>5</td>
<td>7.2</td>
<td>1,220,082</td>
<td>0.096</td>
</tr>
</tbody>
</table>

Total Population: 12,725,076

Table 4.4 Groups Constructed According to the Description of the Archer & Lemeshow F-adjusted Mean Residual Goodness-of-fit Test
4.6 References for Chapter 4


Chapter 5: Evaluating Risk-Specific Cervical Cancer Screening Strategies

5.1 Introduction

Cervical cancer is the second leading cause of female cancer mortality worldwide with 250,000 deaths each year (World Health Organization 2011). The National Cancer Institute (2011) estimates that 12,710 cases of invasive cervical cancer were diagnosed in the US in 2011, and approximately 4,290 women will die from cervical cancer.

A necessary cause of cervical cancer is infection with an oncogenic strain of human papillomavirus (HPV) (Schiffman and Castle 2003, Muñoz 2000); HPV is the most common STI in the United States (Centers for Disease Control and Prevention 2011). Several studies have identified relationships between specific demographic and behavioral characteristics (e.g. marital status, number of lifetime and recent partners) and the probability of contracting HPV infections (Kahn et al. 2007, Dunne et al. 2007, Sasagawa et al. 2005, Wright et al. 2004, Sellors et al. 2003, Winer et al. 2003, Moscicki et al. 2001). However, current screening guidelines do not incorporate the patient’s risky behaviors in the screening decisions (Smith et al. 2010, American College of Obstetricians and Gynecologists 2009, Wright et al. 2007).

Chapter 3 of this dissertation reports the results of an initial exploration of the use of risk-specific screening strategies. These were based on a patient-based discrete event simulation model of
disease progression (from oncogenic HPV infection towards cervical cancer) and screening tests used to prevent invasive cervical cancer.

We evaluated screening strategies consistent with the current ASCCP consensus guidelines (Wright et al. 2007). Regular preventive screening practices for cervical cancer are called “primary screening”. When the primary screening test results are inconclusive, follow up tests called “triage screening” are performed. For primary screening, we evaluated the use of the Papanicolau (Pap) test every 1, 2, and 3 years, and a combination of the Pap test and the HPV DNA test every 3 years. For triage screening, we evaluated the use of a repeat Pap test, the HPV DNA test, and colposcopy. We also designed and evaluated risk-specific screening strategies which screen each woman with the strategy deemed most appropriate for her particular risk level. The risk-specific strategies were developed by simulating risk-homogeneous populations with various probabilities of infection with oncogenic HPV and determining the most cost-effective of the traditional screening strategies for each infection probability. We considered risk-specific strategies that screen women as a function of their current risk level for oncogenic HPV infection, and strategies that screen women as a function of their risk level in prior years to account for the slow progression of the disease. To assess the cost-effectiveness of different screening strategies, we compared the total costs and the total quality-adjusted life years (QALYs) accumulated in the population with each strategy.

The data that populated the model were obtained from various sources. The disease progression and regression rates had been inferred from cohort studies and used in previous cost-effectiveness analyses; health state transitions occurred at 6 month intervals in the model. Data on the accuracy of screening tests conditioned on the true disease state were taken from clinical trials. Some data such as the true prevalence of precancerous lesions in the population are not directly available in
the literature and we inferred values while aiming to minimize the deviation from the existing data. The cost of an invasive cervical cancer case was estimated based on the cost of treatment and the economic burden per years of life lost for each person. The number of quality-adjusted life years (QALYs) accumulated in the population was estimated using the quality adjusted weights from the Cost-Effectiveness Analysis Registry. Details of the model structure and data used can be found in Chapter 3 of this dissertation.

We found the risk-specific screening strategies to be more cost-effective than the traditional, invariant strategies. However, the risk-specific strategies were derived from unrealistic homogeneous populations, and evaluated in populations in which each woman’s risk was a function only of her age. In this chapter, we assess how the risk-specific strategies might be expected to perform in a more realistic population. In particular, we consider a population in which each woman’s risk for oncogenic HPV infection varies over the course of her life as a function of a set of risk factors.

5.2 Methods

In order to model the manner in which the risk for oncogenic HPV infection varies over time, we investigate the quantitative relationship between known risk factors for HPV infection and a woman’s probability of being infected with an oncogenic HPV strain (we use the term “infection” to refer to infection with an oncogenic HPV strain). We model how the risk factors change within the population over time, and how changes in the risk factors affect the probability of HPV infection throughout the lives of each woman in our simulated population.
We used data from the National Health and Nutrition Examination Survey (NHANES) (National Center for Health Statistics (NCHS) 2011), which includes patient demographic and behavioral characteristics, as well as the presence or absence of oncogenic HPV infections. In Section 5.2.1 we describe a logistic regression model derived from the NHANES data, which we use to quantify how a patient’s risk profile impacts her probability of being infected with an oncogenic HPV strain. In Section 5.2.2 we describe our model of a population of women with varying risk factors for oncogenic HPV infection, and the manner in which the characteristics that affect their risk shift over time. These shifts are modeled using probability distributions obtained primarily from US Census data.

**5.2.1 Logistic Regression Model**

In our study, we define a woman’s “risk” as the probability that she has an oncogenic HPV infection. To study risk-specific screening strategies, we begin with a logistic regression model to estimate the probability of oncogenic HPV infection, based on a woman’s demographic and behavioral characteristics (the development of this logistic regression model is described in further detail in Chapter 4 of this dissertation).

We use NHANES data to construct our logistic regression model. NHANES collects health-related information by interviewing a representative sample of the US noninstitutionalized civilian population. NHANES also reports results from physical examinations and laboratory tests, including the HPV DNA test. Details of the survey content and of the data collection and data analysis methodologies are described by the NCHS (2011). Our study included data from 1120 women from the 2003-2004 NHANES cycle, aged 20 to 59. These women answered the demographic, smoking and sexual behavior questionnaires and submitted a self-collected vaginal
swab specimen for HPV testing. Details of the specimen collection and HPV genotyping processes are provided in Dunne et al. (2007) and Kahn et al. (2007).

We selected the independent variables included in the model as predictors based on the existing literature about known risk factors for HPV infection (Moscicki et al. 2001, Sellors et al. 2003, Winer et al. 2003, Wright et al. 2004, Sasagawa et al. 2005, Dunne et al. 2007, Kahn et al. 2007). The independent variables considered for use in the model were drawn from the Demographics, Sexual Behavior, and Smoking Questionnaires. The demographic variables were: age, education level, country of birth, race, below poverty index, and marital status; the variables from the Sexual Behavioral Questionnaire were: age at first intercourse, number of lifetime sex partners, number of recent sex partners (in the past 12 months), history of Chlamydia, history of genital herpes, history of genital warts; the variables from the Smoking Questionnaire were: smoked at least 100 cigarettes in lifetime, is a current smoker.

Our final predictive model uses the following variables to estimate a patient’s probability of having an oncogenic HPV infection: age, marital status, below poverty index, lifetime partners. The predictive model was validated on a separate data sample and the Archer and Lemeshow F-adjusted mean residual goodness-of-fit test did not indicate a lack of fit between the data and the model. The details of this logistic regression analyses and of the final predictive model can be found in Chapter 4 of this dissertation.

5.2.2 Modeling Varying Risk in a Simulated Population

We simulate a population of 100,000 women whose demographic profile and life changes are consistent with that of the US population. We use US Census data to establish a woman’s, age,
marital status, poverty status, and number of lifetime partners at entry to the simulation, as well as the manner in which these characteristics change over time.

Within our simulation, each woman is represented by characteristics that affect her probability of infection: age, marital status, poverty status, number of partners. The set of characteristics for a given woman at a given time is denoted by \( w \). We use \( P_I(w) \) to denote the 6-month probability of infection of a woman with characteristics represented by \( w \). Thus, \( P_I(w_t) \) for \( t \geq 0 \) will represent a woman’s changing infection probability over time \( t \), which changes as a result of changes in her personal characteristics, \( w \).

We assume that the probability of infection for those who are under the age of 13 is zero. The data that we have available from NHANES are for women aged 20 to 59, therefore our logistic regression model only applies to women within this age group. For women 14 to 20, we use the age-specific probability of infection from Kahn et al. (2007), and for women 60 and over, we use the probability of infection for women over 60 from Myers et al. (2000).

Using our logistic regression equation we calculate the probability that a woman has an oncogenic HPV infection (infection prevalence). However, to model the health state transitions in the simulation we need the 6-month probability of acquiring a new infection (infection incidence), which we have defined as \( P_I(w_t) \). We estimate this from the infection prevalence using the epidemiologic relationship: \( \text{Prevalence} = \text{Incidence} \times \text{Duration} \) (Aschengrau and Seage 2003) and using an infection duration of 15 months (Richardson et al. 2003, Munoz et al. 2004).
5.2.2.1 Changes in Risk Factors

Recall that $w_t$ represents a woman’s age, marital status, number of lifetime partners, and poverty status at a given time. These characteristics change throughout her life. We model these life changes, based on the data that we describe in this section.

Marital Status

We model the following changes in marital status: first marriage, end of first marriage (through divorce or death of husband), second marriage, end of second marriage. We did not find sufficiently detailed data to model changes in cohabitation.

We randomly assign an age of first marriage to women entering the simulation, based on the probability distribution shown in Table 5.1 (U.S. Census Bureau 2001). Women who were not married by age 50 are assumed to never marry.

<table>
<thead>
<tr>
<th>Age</th>
<th>Probability of 1st Marriage at Age</th>
</tr>
</thead>
<tbody>
<tr>
<td>≤ 20</td>
<td>0.233</td>
</tr>
<tr>
<td>21-25</td>
<td>0.335</td>
</tr>
<tr>
<td>26-30</td>
<td>0.207</td>
</tr>
<tr>
<td>31-35</td>
<td>0.091</td>
</tr>
<tr>
<td>36-40</td>
<td>0.034</td>
</tr>
<tr>
<td>41-45</td>
<td>0.016</td>
</tr>
<tr>
<td>46-50</td>
<td>0.010</td>
</tr>
<tr>
<td>never married</td>
<td>0.074</td>
</tr>
</tbody>
</table>

Table 5.1 Probability Distribution of Age of First Marriage

In addition, we randomly assign a duration of first marriage to each woman. We use Centers for Disease Control and Prevention data (CDC/NCHS 2002) for probabilities of marriage durations of up to 10 years that are specific to the woman’s age at marriage, and U.S. Census Bureau
(2001) data for probabilities of marriage durations greater than 10 years. These two data sources show consistency in the instances in which they can be compared. The data used to randomly assign a duration of first marriage to women in the simulation are shown in Table 5.2.

<table>
<thead>
<tr>
<th>Age at 1st marriage</th>
<th>First Marriage Duration (years)</th>
<th>&lt; 1</th>
<th>1-3</th>
<th>3-5</th>
<th>5-10</th>
<th>10-15</th>
<th>15-20</th>
<th>20-25</th>
<th>25-30</th>
<th>30-35</th>
<th>35-40</th>
<th>&gt; 40</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;20</td>
<td></td>
<td>0.09</td>
<td>0.13</td>
<td>0.10</td>
<td>0.14</td>
<td>0.07</td>
<td>0.06</td>
<td>0.05</td>
<td>0.04</td>
<td>0.04</td>
<td>0.24</td>
<td></td>
</tr>
<tr>
<td>20-25</td>
<td></td>
<td>0.05</td>
<td>0.09</td>
<td>0.07</td>
<td>0.15</td>
<td>0.07</td>
<td>0.06</td>
<td>0.05</td>
<td>0.04</td>
<td>0.04</td>
<td>0.34</td>
<td></td>
</tr>
<tr>
<td>≥26</td>
<td></td>
<td>0.04</td>
<td>0.04</td>
<td>0.06</td>
<td>0.10</td>
<td>0.07</td>
<td>0.06</td>
<td>0.05</td>
<td>0.04</td>
<td>0.04</td>
<td>0.46</td>
<td></td>
</tr>
</tbody>
</table>

Table 5.2 Probability Distribution of First Marriage Duration

The U.S. Census Bureau (2001) first marriage duration data account for marriages that end through the husband’s death as well as through divorce, for marriage durations of up to 40 years. For first marriages that last longer than 40 years (and are considered in our model to not end in divorce), we used the National Center for Health Statistics age-specific all-cause mortality rates for men (from years 2005-2007), to model the possibility of women with marriages longer than 40 years becoming widows.

We next assign a duration of divorce, that is, the number of years between the end of the first marriage and the start of the second marriage, if there is one. The probability distribution of divorce duration is shown in Table 5.3 and depends on whether the woman was younger or older than 25 at the time of her divorce (Bramlett and Mosher 2002). The data used to model the time between first and second marriages are applied in our model to all women whose first marriage ended, even though a minority of them are widows instead of divorced.
Table 5.3  Probability Distribution of Duration of Divorce

For women who remarry, we randomly assign a duration of second marriage, using the probability distribution shown in Table 5.4 (U.S. Census Bureau 2001). The U.S. Census Bureau (2001) second marriage duration data account for marriages that end through the husband’s death as well as through divorce, for marriage durations of up to 20 years. For second marriages that last longer than 20 years (and are considered in our model to not end in divorce), we used the National Center for Health Statistics age-specific all-cause mortality rates for men, to model the possibility of these women becoming widows.

Table 5.4  Probability Distribution of Second Marriage Duration

We do not model third and higher-order marriages because they are rare; only 2.3% of women marry three or more times (CDC/NCHS 2002).

Changes in marital state occur in the following manner in our simulation model. When women that have never been married reach their age of first marriage, their marital status changes to “married”. When married women reach their duration of first marriage, their marital status
changes to “divorced or widowed”. When they reach the duration of divorce assigned, their marital status changes to “married” again (representing their second marriage). When they reach the duration of second marriage assigned, their marital status changes back to “divorced or widowed”. Women with a first marriage duration greater than 40 years or with a second marriage duration greater than 20 years (which are assumed to not end in divorce) may enter the “divorced or widowed” state if their husband dies. The “divorced or widowed” state is used to keep track of changes in marital status. For the purposes of calculating each woman’s infection probability using the logistic regression equation, the only relevant marital states are “married” and “not married”, which includes the “divorced or widowed” state.

**Number of Lifetime Partners**

We model changes in the number of lifetime partners using the probability distributions of women acquiring new partners each year. The probability of unmarried women having 0 to 3 or more new partners each year, by age, is shown in Table 5.5; the probability of married women having 1 new partner each year is 0.06 (O’Dowd 2003). We randomly assign unmarried women to one of the four categories shown in Table 5.5 based on the probability distribution associated with their age group; women who are assigned to the category “3 or more” are assigned 3 new partners in the simulation model. These data are only available for women aged 18 to 64; in our model women over 65 are assumed to not have new partners.
Table 5.5 Yearly Probability Distribution of New Partners in Unmarried Women

<table>
<thead>
<tr>
<th>New Partners</th>
<th>18-24</th>
<th>25-34</th>
<th>35-44</th>
<th>45-54</th>
<th>55-64</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.690</td>
<td>0.690</td>
<td>0.780</td>
<td>0.890</td>
<td>0.885</td>
</tr>
<tr>
<td>1</td>
<td>0.186</td>
<td>0.186</td>
<td>0.132</td>
<td>0.066</td>
<td>0.069</td>
</tr>
<tr>
<td>2</td>
<td>0.050</td>
<td>0.050</td>
<td>0.035</td>
<td>0.018</td>
<td>0.018</td>
</tr>
<tr>
<td>3 or more</td>
<td>0.074</td>
<td>0.074</td>
<td>0.053</td>
<td>0.026</td>
<td>0.028</td>
</tr>
</tbody>
</table>

Poverty Status

We model transitions in and out of poverty using data from the U.S. Census Bureau (2001-2003).

In Table 5.6 we present the yearly probabilities of entering and exiting poverty, by marital status.

Table 5.6 Yearly Probability of Entering and Exiting Poverty

| Yearly probability of entering poverty, if out of poverty the previous year |
|-----------------------------|-----------------|-----------------|-----------------|-----------------|-----------------|
| Married                     | 0.008           |                 |                 |                 |                 |
| Never Married               | 0.021           |                 |                 |                 |                 |
| Divorced or Widowed         | 0.025           |                 |                 |                 |                 |

| Yearly probability of exiting poverty, if in poverty the previous year |
|-----------------------------|-----------------|-----------------|-----------------|-----------------|-----------------|
| Married                     | 0.137           |                 |                 |                 |                 |
| Never Married               | 0.088           |                 |                 |                 |                 |
| Divorced or Widowed         | 0.049           |                 |                 |                 |                 |

5.3 Results

5.3.1 Validation

We model the progression of the disease from the point of infection with an oncogenic HPV strain through the various stages of precancerous lesions that lead to invasive cervical cancer; we also model the preventive screening events. We simulate 100,000 women through their screening
years. Over the course of those years, their demographic characteristics change, affecting their risk of acquiring an HPV infection, which ultimately affects their probability of developing cervical cancer. In this section we compare the results of our simulation model on the number of incident cervical cancer cases to the known number of cervical cancer cases in the US population. To make this comparison, we use data from the Surveillance, Epidemiology and End Results (SEER) Program.

The incidence rate of cervical cancer reported by SEER for 2004-2008 was 8.1 cases per 100,000 women per year. Evaluating our model using the traditional screening strategy of annual screening and triage with the Pap test, we find the annual incidence rate of cervical cancer to be 8.65 cases per 100,000 women. We also compare our results to SEER’s distribution of cervical cancer diagnoses by age at diagnosis. In Figure 5.1 we present the cumulative distribution of diagnoses by age as reported by SEER and as observed from our model, and note that although the distributions are not identical, the distribution from our model resembles the shape of the distribution from SEER.
The discrepancies between our results and the SEER data can be due to a number of factors. It is impossible to know the true value of some of the input data, because even the best diagnostic tests available are far from perfect. As a result of this, there are inconsistencies in the input data, these are described at length in Chapter 3 of this dissertation. We make some simplifying assumptions in our model, so that it is only an approximation of reality. In addition to this, the higher cervical cancer incidence rate in our model may be due in part to the fact that SEER quantifies the number of cancer cases that are detected and reported, while our output reflects the total number of cancer cases that occur, because the simulation environment allows us to register all health states.

5.3.2 Cost-Effectiveness Analysis

We evaluated 12 screening strategies based on current guidelines, and risk-specific screening strategies. The performance of screening strategies was assessed based on the total costs incurred and the total QALYs accumulated in a simulated population of 100,000 women. We calculated

Figure 5.1 Distribution of Diagnoses by Age, Comparison of our Results to SEER Data
incremental cost-effectiveness ratios (ICER) in which the additional costs of a strategy are divided by the additional QALYs compared with the next, less costly strategy. Strategies that were both more costly and less effective in terms of QALYs than other strategies were considered dominated. In the US, more costly but more effective strategies costing more than $50,000 per QALY gained are commonly not considered cost-effective (Eichler et al. 2004, Kulasingam and Myers 2003). Others (Goldhaber-Fiebert et al. 2008) also consider strategies that fall within the range of $50,000-$100,000 per QALY to be cost-effective.

We have coded the screening strategies as follows: $P; T$, where $P$ corresponds to the primary screening strategy and $T$ corresponds to the triage screening strategy. The options for primary screening are $P_n$, denoting a Pap test every $n$ years ($n$: 1, 2, 3); and $PD$, denoting a Pap and an HPV DNA test every 3 years. The options for triage screening are: $P = 'repeat Pap in 6 & 12 months'; D = 'HPV DNA test'; C = 'immediate colposcopy'. The risk-specific strategy, which was derived in Chapter 3 of this dissertation, screens each woman with the strategy deemed most appropriate for her particular risk level, as shown in Table 5.7.

<table>
<thead>
<tr>
<th>Infection Probability</th>
<th>≤ 0.03</th>
<th>0.04 – 0.05</th>
<th>≥ 0.06</th>
</tr>
</thead>
<tbody>
<tr>
<td>Risk Level</td>
<td>Low</td>
<td>Medium</td>
<td>High</td>
</tr>
<tr>
<td>Preferred Strategy</td>
<td>$P_3;D$</td>
<td>$P_2;D$</td>
<td>$P_1;D$</td>
</tr>
</tbody>
</table>

Table 5.7 Screening Strategies Used for Each Risk Level in Risk-Specific Strategies

Recall that $P_i(w_i)$ represents a woman’s changing infection probability over time as her personal characteristics change. The risk-specific strategy is implemented as a function of $P_i(w_i)$ at a given point in time. Additionally, it may be offset by an amount of time $i$ years in the past to account for the slow progression of the disease. In this case the infection probability used to classify the woman’s level of risk is $P_i(w_{i−i})$. 111
We also assess the benefits of implementing these risk-specific strategies (flexible strategies as a function of a set of risk factors), as compared with the benefits of implementing flexible strategies as a function of age only (as in Chapter 3 of this dissertation). We evaluate the performance of the age-specific strategy that was found to be most cost-effective in Chapter 3 (age-specific strategy with a 15 year offset, Table 3.14), in this more realistic population in which risk varies as a function of a set of risk factors.

In Table 5.8 we present the cost-effectiveness analysis of the six non-dominated screening strategies. The strategies are evaluated in order of increasing costs. The risk-specific strategy without a time offset is compared with P3;D and is found to be cost-effective (since the ICER does not exceed $50,000). Next we compare the risk-specific strategy with a five year offset, with the non-offset strategy and find the five year offset strategy to be cost-effective. Next the risk-specific strategy with a ten year offset is found to be cost-effective in comparison to the five year offset risk-specific strategy. The age-specific strategy with a fifteen year offset is not found to be more cost-effective than the risk-specific strategy with a 10 year offset. Finally, P1;D is compared with the most cost-effective strategy up to that point (the risk-specific with a ten year offset) and is not found to be cost-effective.

<table>
<thead>
<tr>
<th>Strategy</th>
<th>Total Costs</th>
<th>Total QALYs</th>
<th>Incremental Costs</th>
<th>Incremental QALYs</th>
<th>Incremental ICER</th>
<th>Comparison Strategy</th>
</tr>
</thead>
<tbody>
<tr>
<td>P3;D</td>
<td>325,301,969</td>
<td>7,917,211</td>
<td>25,730,744</td>
<td>6,026</td>
<td>4,270</td>
<td>P3;D</td>
</tr>
<tr>
<td>Risk-specific, i = 0</td>
<td>351,032,712</td>
<td>7,923,237</td>
<td>9,742,775</td>
<td>708</td>
<td>13,761</td>
<td>Risk-specific, i = 0</td>
</tr>
<tr>
<td>Risk-specific, i = 5</td>
<td>360,775,487</td>
<td>7,923,945</td>
<td>13,581,933</td>
<td>469</td>
<td>28,959</td>
<td>Risk-specific, i = 5</td>
</tr>
<tr>
<td><strong>Risk-specific, i = 10</strong></td>
<td><strong>374,357,420</strong></td>
<td><strong>7,924,414</strong></td>
<td><strong>13,581,933</strong></td>
<td><strong>469</strong></td>
<td><strong>28,959</strong></td>
<td><strong>Risk-specific, i = 10</strong></td>
</tr>
<tr>
<td>Age-specific, i = 15</td>
<td>380,750,526</td>
<td>7,924,423</td>
<td>6,393,106</td>
<td>9</td>
<td>710,345</td>
<td>Risk-specific, i = 10</td>
</tr>
<tr>
<td>P1;D</td>
<td>501,691,553</td>
<td>7,926,014</td>
<td>127,334,133</td>
<td>1,600</td>
<td>79,584</td>
<td>Risk-specific, i = 10</td>
</tr>
</tbody>
</table>

Table 5.8 Cost-Effectiveness Analysis of Screening Strategies
Preferred strategy is shown in bold, considering a cost-effectiveness threshold of $50,000/QALY
The risk-specific strategy with a 10 year offset (i.e., using $P_I(w_{t-10})$, the probability of infection of 10 years before the time of screening to determine a woman’s risk level) was found to be the most cost-effective strategy. To implement this screening strategy, the logistic regression equation would be used to estimate the $P_I(w_{t-10})$ of each woman using her risk factors from 10 years earlier.

5.4 Sensitivity Analysis

In our base case analysis, we assumed that all women have 100% adherence to screening appointments, and that the population is not vaccinated against HPV infections. In our sensitivity analysis we relaxed these assumptions and evaluated the performance of the traditional and risk-specific screening strategies.

5.4.1 Model Including Vaccine

In 2006, the U.S. Food and Drug Administration approved the quadrivalent vaccine Gardasil® for women 9 to 26 years of age to protect against infection with HPV types 6, 11, 16, and 18 (U.S. Department of Health and Human Services 2006). The costs associated with the vaccine are $120 for each of the three doses, plus administration fees (Elbasha et al. 2007). These vaccination costs are incurred independently of the screening strategy used. In order to study the effect that a vaccinated population would have on the cost-effectiveness of risk-specific strategies, we evaluated the performance of traditional and risk-specific screening strategies in a population of women vaccinated against infection with the oncogenic HPV types 16 and 18 (types 6 and 11 are non-oncogenic). The details of our model of a vaccinated population have been described in
Chapter 3 of this dissertation. We consider a population in which all women are vaccinated, and a population with the current vaccine coverage of approximately 33%, considering women who have received all three doses of the vaccine (CDC 2009).

Table 5.9 shows the cost-effectiveness analysis of the non-dominated screening strategies in a population with 100% vaccination. The risk-specific strategy with a 10 year offset was found to be the most cost-effective.

<table>
<thead>
<tr>
<th>Strategy</th>
<th>Total Costs</th>
<th>Total QALYs</th>
<th>Incremental Costs</th>
<th>Incremental QALYs</th>
<th>ICER*</th>
<th>Comparison Strategy</th>
</tr>
</thead>
<tbody>
<tr>
<td>P3;D</td>
<td>241,487,393</td>
<td>8,009,747</td>
<td>60,190,293</td>
<td>1,348</td>
<td>44,652</td>
<td>P3;D</td>
</tr>
<tr>
<td>Risk-specific, ( i = 0 )</td>
<td>301,677,686</td>
<td>8,011,095</td>
<td>16,266,296</td>
<td>106</td>
<td>153,456</td>
<td>Risk-specific, ( i = 0 )</td>
</tr>
<tr>
<td>Risk-specific, ( i = 5 )</td>
<td>317,943,982</td>
<td>8,011,201</td>
<td>32,458,916</td>
<td>819</td>
<td>39,632</td>
<td>Risk-specific, ( i = 0 )</td>
</tr>
<tr>
<td>Risk-specific, ( i = 10 )</td>
<td>334,136,602</td>
<td>8,011,914</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 5.9  Cost-Effectiveness Analysis for a Population with 100% Vaccination
Preferred strategy is shown in bold, considering a cost-effectiveness threshold of $50,000/QALY.

For a population with the current HPV vaccine coverage of 33% (CDC 2009), we evaluated the following different screening approaches:

A. Screen vaccinated and non-vaccinated subpopulations with the same screening strategy, considering only the traditional strategies.

B. Screen vaccinated and non-vaccinated subpopulations differently, considering only the traditional strategies.

C. Screen the vaccinated and non-vaccinated subpopulations considering risk-specific screening strategies.

Table 5.10 shows the cost-effectiveness analysis of the non-dominated screening approaches in a population with 33% vaccination coverage. Approach C was found to be the most cost-effective,
which uses the risk-specific strategy that is preferred for each of the subpopulations (vaccinated and non-vaccinated), which in this case was the strategy with a 10 year offset for both subpopulations.

<table>
<thead>
<tr>
<th>Approach</th>
<th>Total Costs</th>
<th>Total QALYs</th>
<th>Incremental Costs</th>
<th>Incremental QALYs</th>
<th>ICER*</th>
<th>*Comparison Approach</th>
</tr>
</thead>
<tbody>
<tr>
<td>B</td>
<td>316,797,629</td>
<td>7,950,197</td>
<td>22,762,443</td>
<td>2</td>
<td>11,496,184</td>
<td>B</td>
</tr>
<tr>
<td>A</td>
<td>339,560,072</td>
<td>7,950,199</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>361,084,550</td>
<td>7,953,289</td>
<td>44,286,921</td>
<td>3,092</td>
<td>14,324</td>
<td>B</td>
</tr>
</tbody>
</table>

Table 5.10 Cost-Effectiveness Analysis for a Population with 33% Vaccination
Preferred strategy is shown in bold, considering a cost-effectiveness threshold of $50,000/QALY.

5.4.2 Model Including Adherence

In our base case model we assume that all women attend all screening appointments, which is usually not the case. Paskett et al. (2010) found the percentage of the population that is within risk-appropriate guidelines to be 68%. Additionally, women at higher risk for HPV and cervical cancer are less likely to adhere to screening guidelines (Marcus et al. 1992, Basen-Engquist et al. 2003, Eggleston et al. 2007, Paskett et al. 2010). To study the effect of imperfect adherence in the cost-effectiveness of risk-specific screening strategies, we estimated probabilities of adherence conditioned on each risk level such that the probability of adherence of the overall population was approximately 70%, and the probability of adherence was inversely proportional to the probability of infection. Our model of imperfect adherence is described in detail in Chapter 3 of this dissertation.

Table 5.11 shows the cost-effectiveness analysis of the non-dominated screening strategies in a population with imperfect adherence. The risk-specific strategy with a 10 year offset was found to be the most cost-effective.
<table>
<thead>
<tr>
<th>Strategy</th>
<th>Total Costs</th>
<th>Total QALYs</th>
<th>Incremental Costs</th>
<th>Incremental QALYs</th>
<th>ICER*</th>
<th>*Comparison</th>
</tr>
</thead>
<tbody>
<tr>
<td>P3;D</td>
<td>325,840,530</td>
<td>7,915,327</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Risk-specific, $i=0$</td>
<td>326,643,188</td>
<td>7,920,407</td>
<td>802,658</td>
<td>5,080</td>
<td>158</td>
<td>P3;D</td>
</tr>
<tr>
<td>Risk-specific, $i=5$</td>
<td>336,545,432</td>
<td>7,921,162</td>
<td>9,902,244</td>
<td>755</td>
<td>13,116</td>
<td>Risk-specific, $i=0$</td>
</tr>
<tr>
<td>Risk-specific, $i=10$</td>
<td>349,481,581</td>
<td>7,922,057</td>
<td>12,936,149</td>
<td>895</td>
<td>14,454</td>
<td>Risk-specific, $i=5$</td>
</tr>
<tr>
<td>P1;D</td>
<td>453,379,896</td>
<td>7,923,688</td>
<td>103,898,316</td>
<td>1,631</td>
<td>63,702</td>
<td>Risk-specific, $i=10$</td>
</tr>
</tbody>
</table>

Table 5.11 Cost-Effectiveness Analysis for a Population with Imperfect Adherence
Preferred strategy is shown in bold, considering a cost-effectiveness threshold of $50,000/QALY.

A summary of our results is provided in Table 5.12.

<table>
<thead>
<tr>
<th>Strategy</th>
<th>Total Costs</th>
<th>Total QALYs</th>
<th>Incremental Costs</th>
<th>Incremental QALYs</th>
<th>ICER*</th>
<th>*Comparison</th>
</tr>
</thead>
<tbody>
<tr>
<td>P3;D</td>
<td>325,301,969</td>
<td>7,917,211</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Risk-specific, $i=0$</td>
<td>351,032,712</td>
<td>7,923,237</td>
<td>25,730,744</td>
<td>6,026</td>
<td>4,270</td>
<td>P3;D</td>
</tr>
<tr>
<td>Risk-specific, $i=5$</td>
<td>360,775,487</td>
<td>7,924,141</td>
<td>9,742,775</td>
<td>708</td>
<td>13,761</td>
<td>Risk-specific, $i=0$</td>
</tr>
<tr>
<td>Risk-specific, $i=10$</td>
<td>374,357,420</td>
<td>7,924,414</td>
<td>13,581,933</td>
<td>469</td>
<td>28,959</td>
<td>Risk-specific, $i=5$</td>
</tr>
</tbody>
</table>

Table 5.12 Summary of Results
Approaches for a population with 33% vaccination: A: Screen vaccinated and non-vaccinated subpopulations with the same strategy, considering only the traditional strategies. B: Screen vaccinated and non-vaccinated subpopulations differently, considering only the traditional strategies. C: Screen the vaccinated and non-vaccinated subpopulations with risk-specific screening strategies ($i=10$).
5.5 Discussion

We evaluated the cost-effectiveness of traditional and risk-specific screening strategies to prevent cervical cancer in a simulated population. Women in the simulated population have a varying risk for oncogenic HPV infection throughout their lives, which varies in accordance to a set of risk factors. We found risk-specific strategies to be most cost-effective, in particular one that screened women according to their risk for HPV infection 10 years prior. The same risk-specific screening strategy was found to be most cost-effective in a population with imperfect adherence and in a vaccinated population. We consider these findings to be initial evidence in favor of adopting risk-specific screening policies, and that these would likely result in more cost-efficient and medically effective overall screening than the current uniform policies.

Our findings are consistent by design with the fact that having an oncogenic HPV infection is a necessary cause for cervical cancer and that the risk for acquiring an infection varies over the course of women’s lives. Consequently, women are at different risks for developing cervical cancer throughout their lives, and adjusting the frequency of screening to accommodate this varying risk can be expected to be cost-effective. The fact that in most cases it takes years for precancerous lesions to appear and develop into cervical cancer is likely the reason why we find risk-specific screening considering each woman’s risk for HPV infection from 10 years earlier to be most cost-effective.

A number of previous studies have compared the different traditional (invariant) screening strategies to prevent cervical cancer. McLay et al. (2010) have been the first to design dynamic, age-based screening policies, which they found to be more cost-effective than traditional screening strategies. However, they only consider triage screening with direct colposcopy and
their disease model does not include an Infected state, but considers the presence of an HPV infection as equivalent to having CIN 1. Our work is the first to date that evaluates the performance of risk-specific strategies, using a model of how risk factors affect the probability of HPV infection, and a simulated population whose risk factors change over time in a manner consistent with the US population. When we evaluate the performance of the most cost-effective age-specific strategy from Chapter 3 (with a 15 year offset), in the population simulated in this chapter with varying risk factors, we find that the risk-specific strategy with a 10 year offset is more cost-effective. The fact that the risk-specific strategy outperformed the age-specific strategy is significant because the age-specific strategy is easier to implement, and would otherwise be an attractive alternative.

Our study’s main limitation is the fact that the risk-specific strategies evaluated were designed using unrealistic homogeneous populations with a probability of infection which remains constant throughout each woman’s lifetime (see Chapter 3 of this dissertation). We evaluated the performance of the risk-specific strategies in a realistic population with varying levels of risk, and the fact that they were found to be the most cost-effective suggests that they are valuable alternatives to the current screening policies. However, we cannot consider them to be optimal solutions, as exploring different combinations of strategies and different risk level breakpoints in a realistic population may result in finding superior risk-specific strategies.

Another limitation is the unavailability and inconsistencies in some of the input data that populate our model, in particular those which refer to the prevalence of lesions and the accuracy of screening tests; this is described at length in Chapter 3. Additionally, as with any model, we make simplifying assumptions, so that it is only an approximation of reality.
A strength of our study is that we use a disease progression and screening model that did not make some common and inaccurate assumptions, such as those referring to the accuracy of screening tests (see Chapter 3). We also use data representative of the US population from reliable sources (NHANES, US Census) to develop a model of risk for oncogenic HPV infection that allows us to explore dynamic screening strategies that extend beyond a woman’s age.

This work is an initial exploration of risk-specific strategies for cervical cancer prevention. More analyses are needed, most importantly the development of risk-specific strategies within the context of a realistic population, which in this work we use only to evaluate the risk-specific strategies developed in homogeneous populations (Chapter 3). Different combinations of traditional strategies used to form part of risk-specific strategies and different risk level breakpoints need to be explored in a model of a population with varying levels of risk.

We found risk-specific screening strategies that consider a woman’s past risk of HPV infection to be most cost-effective. However, to implement these strategies each patient would need to recall factors that define her infection risk ten years in the past. Recalling age and marital status does not appear to be difficult, but it may not be easy to recall poverty status or the number of partners at a past point in time. It is possible that a more thorough electronic health record system that follows the patient’s risk factors could provide this kind of historical data in the future. When collecting risk information from patients, having them fill out a form as opposed to interviewing them would be recommended; the NHANES Sexual Behavior Questionnaire is filled out in private.

Our work is the first to date that explores the possibility of risk-specific screening for cervical cancer prevention. This approach is consistent with the existing literature on known risk factors.
for cervical cancer and HPV infection. We conclude that risk-specific screening strategies have the potential to be more cost-effective than the traditional invariant strategies. The design of risk-specific strategies for cervical cancer prevention should be explored further.
5.6 References for Chapter 5


Chapter 6: Conclusion

Cervical cancer is the second leading cause of female cancer mortality worldwide, but most deaths by cervical cancer can be prevented with appropriate screening programs. In the United States cervical cancer was once the leading cause of cancer death for women; however in the past forty years, the number of cervical cancer cases and the number of deaths from cervical cancer have decreased significantly due to regular screening.

More recently developed screening tests and HPV vaccines, continuing disparities in cervical cancer incidence, and a better understanding of the risk factors and disease progression of cervical cancer make screening policies an evolving issue. Recent guidelines call for screening to begin later in life and suggest that greater intervals between screenings can be appropriate for some age groups. These are reasonable moves to reduce over-screening, but risk-specific screening strategies have yet to be considered.

We have performed an initial exploration of the potential of risk-specific screening to prevent cervical cancer cost-effectively. Our findings suggest that screening women according to their level of risk may be more cost-effective than the traditional screening strategies. The implementation of risk-specific strategies has the potential to reduce over-screening without neglecting women when they are at higher risk for developing cervical cancer. Risk-specific screening can provide the possibility of a more efficient allocation of public health resources; this can offer societal benefits beyond those directly associated with the reduction in cervical cancer.
Our findings indicate that risk-specific screening for cervical cancer prevention can be beneficial, and invite further study of this possibility. The risk-specific strategies were designed using unrealistic homogeneous populations with a probability of infection which remains constant throughout each woman’s lifetime. They were found to be the most cost-effective when evaluated in a realistic population with varying levels of risk, however exploring different combinations of strategies and risk level breakpoints in a realistic population may result in finding superior risk-specific strategies. Therefore, it will be important to explore risk-specific strategies that are developed, not only evaluated, within the context of realistic simulated populations. Furthermore, future research can explore the use of a different measure of risk, such as the probability of cervical cancer as opposed to the probability of HPV infection, to determine risk levels and screening approaches. Risk-specific strategies for cervical cancer can also be explored in lower resource and higher risk countries. Preventive screening programs for other diseases for which risk factors have been identified and can be measured also have the potential to benefit from a risk-specific approach. Public health policies that consider individuals’ needs and adjust to them when it is feasible and cost-effective to do so, can provide a better service to society.


