OCCURRENCE, DETERMINANTS AND DYNAMICS OF HPV COINFECTIONS IN A COHORT OF MONTREAL UNIVERSITY STUDENTS

By

Michaela Anne Smith

A thesis submitted to the Department of Community Health and Epidemiology
in conformity with the requirements for
the degree of Master of Science

Queen’s University
Kingston, Ontario, Canada

(April, 2011)

Copyright © Michaela Anne Smith, 2011
Abstract

Background: Coinfections with multiple types of human papillomavirus (HPV) are a common occurrence among HPV-infected individuals, but the clinical significance and etiology of these infections remain unclear. Though current evidence suggests that women with coinfections have increased HPV exposure (i.e. more sexual partners), it is also hypothesized that these women may represent a subgroup with increased HPV susceptibility, though this has been rarely studied to date.

Purpose: The purpose of this project was to examine the occurrence, determinants and dynamics of HPV coinfections in a cohort of university students in order to explore the relationship(s) between coinfections, lifestyle factors and immunological susceptibility.

Methods: This project is based on a secondary analysis of data from the McGill-Concordia Cohort, a longitudinal study of the natural history of HPV infection in 621 female university students in Montreal, Quebec. Participants were followed for 2 years at 6-month intervals. At each visit, cervical specimens were collected for cytology and HPV testing, and women completed a questionnaire about lifestyle and risk behaviours. Two definitions of coinfections were used: cumulative coinfection over follow-up and concurrent coinfection at each visit. Kaplan-Meier techniques were used to estimate incidence and duration of coinfections and multiple logistic regression was used to identify determinants of coinfections and associations between coinfections and squamous intraepithelial lesions (SIL).

Results: More than half of the cohort became infected with HPV and of those, over 60% acquired multiple HPV types over follow-up. Incidence of coinfections was significantly increased among
HPV-infected women at enrollment. The most important determinant of coinfection occurrence was number of sexual partners (both lifetime and new), though some genes of the immune response (HLA-DQB1*06:02, HLA-G*01:01:03 and HLA-G*01:01:05) were also significant predictors. Women with coinfections, particularly those with 4+ HPV types, also had longer infection durations and greatly increased odds of SIL.

**Conclusions:** Women with coinfections acquire new HPV types at an increased rate and have greater HPV persistence and occurrence of SIL, which may indicate immunological susceptibility. HPV coinfections mainly occur due to increased sexual activity but a decreased immune response to the virus may also be involved in a subset of women.
Statement of Contributions and Support

This thesis represents the work of Michaela Smith in collaboration with her supervisors, Dr. Harriet Richardson and Dr. Eduardo Franco. The original epidemiologic study on which this thesis is based was designed by Dr. Franco and coordinated by Dr. Richardson at McGill University. As such, assembly of the McGill-Concordia Cohort and all data collection (including the laboratory analyses) were completed prior to the beginning of this thesis project. For this study focusing on HPV coinfections, Michaela Smith conceived the research objectives for the project, conducted all statistical analyses and wrote the chapters of this thesis. Drs. Richardson and Franco provided input and advice at all stages of the project as well as editorial feedback on the final written document.

Financial support for the McGill-Concordia Cohort study was provided by the Canadian Institutes of Health Research (CIHR) (Grants MT-13649 and MOP-53111 to Dr. Franco as PI). Michaela Smith also received additional financial support from the Queen's University Research Initiation Grant (Dr. Richardson) and the Cancer Research Society Division of Epidemiology, Department of Oncology, McGill University (Dr. Franco).
Acknowledgements

Several individuals and organizations deserve recognition for their contributions to this endeavor.

First, I would like to acknowledge my supervisors Dr. Harriet Richardson and Dr. Eduardo Franco. To Harriet: thank you for your guidance and advice over the course of this project and for the continued interest and perspective you brought to this topic despite the new direction your current research has taken. Your willingness to discuss and exchange ideas on almost any subject has greatly contributed to my learning and development throughout this process. To Eduardo: thank you for providing your expertise to this project and for the time and ideas you have contributed to its development—they have been very valuable to me.

Thank you as well to the faculty, staff, and students of the Department of Community Health and Epidemiology for academic, administrative and moral support. I would particularly like to acknowledge Dr. Paul Peng for his statistical guidance, Anne Grundy for her kindness and advice, Mark McPherson for introducing me to the “retain” statement, and Erin Lillie for sharing my devotion to coffee, yoga and Buffy. These were all invaluable contributions to this process. I would also like to acknowledge the financial support of the Ontario Graduate Scholarship and the Queen’s Graduate Award.

Finally, a special thanks to my parents and to Eric: thank you for your patience, encouragement and advice throughout this process and for your belief in me when it was needed. This would not have been possible without your support.
Table of Contents

Abstract .............................................................................................................................................. ii
Statement of Contributions and Support ......................................................................................... iv
Acknowledgements ........................................................................................................................... v
Table of Contents ............................................................................................................................. vi
List of Figures ........................................................................................................................................ ix
List of Tables ......................................................................................................................................... x
List of Acronyms and Abbreviations ............................................................................................... xi

Chapter 1 Introduction ...................................................................................................................... 1
  1.1 Overview and rationale .............................................................................................................. 1
  1.2 Thesis objectives ....................................................................................................................... 2
  1.3 Overview of study design ......................................................................................................... 3
  1.4 Thesis organization .................................................................................................................... 3

Chapter 2 Literature Review ............................................................................................................. 4
  2.1 Cervical cancer: Overview of epidemiology and etiology ....................................................... 4
    2.1.1 Distribution and global burden ............................................................................................ 4
    2.1.2 Human papillomavirus: a necessary cause ....................................................................... 5
  2.2 Human papillomavirus: Overview and role in cervical carcinogenesis .............................. 6
    2.2.1 Taxonomy and classification ............................................................................................... 6
    2.2.2 HPV and cervical carcinogenesis ....................................................................................... 8
      2.2.2.1 Natural history ............................................................................................................... 8
      2.2.2.2 Cofactors ....................................................................................................................... 9
      2.2.2.3 Screening and new prevention approaches .................................................................. 12
    2.2.3 Diagnosis of HPV infection ............................................................................................. 14
  2.3 Human papillomavirus: Epidemiology .................................................................................... 16
    2.3.1 Prevalence and incidence .................................................................................................. 16
    2.3.2 Risk factors for HPV infection ......................................................................................... 17
    2.3.3 HPV persistence ............................................................................................................... 21
  2.4 HPV coinfections ..................................................................................................................... 22
    2.4.1 Definitions ......................................................................................................................... 23
    2.4.2 Prevalence and incidence of coinfections ....................................................................... 24
    2.4.3 Coinfection dynamics ....................................................................................................... 25
      2.4.3.1 Clustering ..................................................................................................................... 25
Chapter 3 Methods ........................................................................................................... 33
  3.1 Purpose and objectives ............................................................................................... 33
  3.2 Study design ............................................................................................................... 34
    3.2.1 Overview ............................................................................................................. 34
    3.2.2 Study population and accrual of subjects ........................................................... 34
  3.3 Data collection ......................................................................................................... 35
    3.3.1 Laboratory procedures ....................................................................................... 35
    3.3.2 Questionnaires ................................................................................................ 38
  3.4 Defining coinfections .............................................................................................. 40
  3.5 Statistical analyses ................................................................................................. 41
    3.5.1 Objective 1: Occurrence of coinfections ............................................................ 41
    3.5.2 Objective 2: Determinants of coinfections ....................................................... 43
    3.5.3 Objective 3: Clinical implications of coinfections ............................................. 48
    3.5.4 Missing data .................................................................................................... 50
  3.6 Ethical considerations ............................................................................................ 51
Chapter 4 Results .............................................................................................................. 52
  4.1 Objective 1: Occurrence of coinfections ................................................................. 52
    4.1.1 Prevalence ...................................................................................................... 52
    4.1.2 Incidence ....................................................................................................... 56
  4.2 Objective 2: Determinants of coinfections ............................................................... 61
    4.2.1 Cumulative coinfection status: bivariate results .............................................. 62
    4.2.2 Cumulative coinfection status: multivariate results ....................................... 66
    4.2.3 Concurrent coinfection status: multivariate results ....................................... 68
  4.3 Objective 3: Clinical implications of coinfections ................................................... 70
    4.3.1 Infection duration ......................................................................................... 70
    4.3.2 Cervical lesions ............................................................................................. 74
  4.4 Brief summary ....................................................................................................... 76
Chapter 5 Discussion ........................................................................................................... 77
  5.1 Summary of results and interpretations of findings ................................................ 77
5.1.1 Prevalence and incidence of coinfections ..............................................77
5.1.2 Risk factors for coinfection occurrence..................................................79
5.1.3 Coinfections, infection duration and lesions ..........................................85
5.2 Methodological issues and limitations ......................................................89
  5.2.1 Loss to follow up......................................................................................89
  5.2.2 Measurement error and misclassification...............................................90
    5.2.2.1 HPV status.......................................................................................90
    5.2.2.2 Other variables ...............................................................................91
  5.2.3 Generalizability .....................................................................................91
5.3 Strengths and significance ..........................................................................93
5.4 Conclusions and future directions .............................................................94
References........................................................................................................96
Appendix A Assessment of Differential Loss to Follow up...............................105
Appendix B Assessment of Generalizability .....................................................106
Appendix C Ethics Approval ............................................................................107
Appendix D Consent Form ..............................................................................108
Appendix E Enrollment Questionnaire ............................................................110
Appendix F Follow-up Questionnaire ..............................................................125
List of Figures

Chapter 2 Literature Review
Figure 2.1 – Overview of cervical carcinogenesis.................................................................9
Figure 2.2 – Visual representation of different definitions of HPV coinfections..................24

Chapter 3 Methods
Figure 3.1 – Conceptual diagram of thesis objectives..........................................................33
Figure 3.2 – Overview of data collection procedures..............................................................39
Figure 3.3 – The challenge of defining coinfections in practice ...........................................40
Figure 3.4 – Conceptual model of risk factors for HPV coinfections in the cohort.................46

Chapter 4 Results
Figure 4.1 – Cumulative probabilities of (A) concurrent coinfection and (B) infection with new HPV types based on HPV status at baseline .................................................................59
Figure 4.2 – Flow charts for Objective 2................................................................................61
Figure 4.3 – Time to clearance of longest enduring type-specific infection, stratified by coinfection status for (A) incident infections and (B) all infections.................................................72
Figure 4.4 – Time to clearance of longest enduring type-specific infection, stratified by cumulative number of types for (A) incident infections and (B) all infections...........................................73
List of Tables

Chapter 2 Literature Review
Table 2.1 – Classification of cytological precursors .......................................................13

Chapter 3 Methods
Table 3.1 – Overview and conceptualization of exposure variables ..................................45

Chapter 4 Results
Table 4.1 – Prevalence of concurrent and cumulative coinfections in entire cohort and in women who were ever HPV-positive ..........................................................52
Table 4.2 – Maximum number of HPV types detected per woman at different points over follow-up (n=621) ........................................................................................................53
Table 4.3 – Type-specific HPV occurrence in single- and multiple-type infections ..........54
Table 4.4 – Incidence rates of concurrent coinfection (n=550) ............................................57
Table 4.5 – Incidence rates of new HPV infection (n=621) ................................................57
Table 4.6 – Incidence rates of coinfections stratified by visits completed ......................60
Table 4.7 – Incidence rates of new HPV infections stratified by visits completed ............60
Table 4.8 – Frequency distributions and bivariate ORs between baseline variables and cumulative coinfection status .........................................................................................63
Table 4.9 – Multivariate ORs for associations between baseline exposures and cumulative coinfection status ......................................................................................................67
Table 4.10 – Multivariate ORs for associations between baseline and time-dependent exposures and concurrent coinfection status at individual visits ........................................69
Table 4.11 – Mean duration of type-specific HPV infections based on cumulative HPV infection status and stratified by cumulative number of types ..............................................................71
Table 4.12 – Prevalence of cytological outcomes in the cohort ........................................74
Table 4.13 – Associations between HPV coinfection status and SIL cumulatively over follow-up and concurrently at the same visit ................................................................75
Table 4.14 – Associations between number of HPV types and SIL at different time points ....76
# List of Acronyms and Abbreviations

<table>
<thead>
<tr>
<th>Acronym</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AIDS</td>
<td>Acquired immunodeficiency syndrome</td>
</tr>
<tr>
<td>ASCUS</td>
<td>Atypical squamous cells of undetermined significance</td>
</tr>
<tr>
<td>CI</td>
<td>Confidence interval</td>
</tr>
<tr>
<td>CIN</td>
<td>Cervical intraepithelial neoplasia</td>
</tr>
<tr>
<td>DNA</td>
<td>Deoxyribonucleic acid</td>
</tr>
<tr>
<td>FTP</td>
<td>Full-term pregnancy</td>
</tr>
<tr>
<td>GEE</td>
<td>Generalized estimating equations</td>
</tr>
<tr>
<td>HCII</td>
<td>Hybrid Capture II</td>
</tr>
<tr>
<td>HIV</td>
<td>Human immunodeficiency virus</td>
</tr>
<tr>
<td>HLA</td>
<td>Human leukocyte antigen</td>
</tr>
<tr>
<td>HPV</td>
<td>Human papillomavirus</td>
</tr>
<tr>
<td>HR-HPV</td>
<td>High-risk human papillomavirus</td>
</tr>
<tr>
<td>HSIL</td>
<td>High-grade squamous intraepithelial lesion</td>
</tr>
<tr>
<td>IARC</td>
<td>International Agency for Research on Cancer</td>
</tr>
<tr>
<td>IR</td>
<td>Incidence rate</td>
</tr>
<tr>
<td>LR-HPV</td>
<td>Low-risk human papillomavirus</td>
</tr>
<tr>
<td>LSIL</td>
<td>Low-grade squamous intraepithelial lesion</td>
</tr>
<tr>
<td>MHC</td>
<td>Major histocompatibility complex</td>
</tr>
<tr>
<td>NHANES</td>
<td>National Health and Nutrition Examination Survey</td>
</tr>
<tr>
<td>OC</td>
<td>Oral contraceptive</td>
</tr>
<tr>
<td>OR</td>
<td>Odds ratio</td>
</tr>
<tr>
<td>ORF</td>
<td>Open reading frame</td>
</tr>
<tr>
<td>PCR</td>
<td>Polymerase chain reaction</td>
</tr>
<tr>
<td>Acronym</td>
<td>Description</td>
</tr>
<tr>
<td>---------</td>
<td>------------------------------------</td>
</tr>
<tr>
<td>PYLL</td>
<td>Potential years of life lost</td>
</tr>
<tr>
<td>RR</td>
<td>Relative risk / Rate ratio</td>
</tr>
<tr>
<td>SAS</td>
<td>Statistical Analysis Software</td>
</tr>
<tr>
<td>SIL</td>
<td>Squamous intraepithelial lesion(s)</td>
</tr>
<tr>
<td>SPSS</td>
<td>Statistical Package for the Social Sciences</td>
</tr>
<tr>
<td>STI</td>
<td>Sexually transmitted infection</td>
</tr>
<tr>
<td>TOL</td>
<td>Tolerance</td>
</tr>
<tr>
<td>VIF</td>
<td>Variance inflation factor</td>
</tr>
</tbody>
</table>
Chapter 1

Introduction

1.1 Overview and rationale

Human papillomavirus (HPV) is one of the most common sexually transmitted infections worldwide (1, 2). Over 150 different types of HPV have been identified thus far, and at least 40 of these types have been found to infect the anogenital tract (3). These types are typically classified as either oncogenic or non-oncogenic depending on their association with cervical carcinoma(s) (4). In fact, in the last decade it has been determined that oncogenic HPV infection is a necessary, but not sufficient, cause of cervical cancer—the first such relationship to be discovered in cancer epidemiology (5). As such, much of the HPV research in recent years has been toward the goal of identifying other “cofactors” involved in the etiology of cervical cancer, that is, factors which increase the risk of cervical cancer among women infected with HPV (1, 6). Such cofactors are broadly considered to be characteristics of the host (e.g. genetic susceptibility), virus (e.g. viral load, genetic variants) and/or environment (e.g. smoking, oral contraceptive use) (6).

Due to the high prevalence of HPV and the high multiplicity of HPV types in existence, infections with multiple HPV types commonly occur. Such multiple-type infections are largely known as HPV coinfections, and occur most frequently among young, sexually active individuals (1, 7). In practice, coinfections may be defined in several different ways, but are generally considered to be infections in which ≥2 or more types of HPV are detected at the same point in time or at different points in time (8, 9). Following from this definition, an HPV infection that only includes one type of HPV has been called an HPV monoinfection (10) (and will be referred to as such throughout this thesis).
Although research into coinfections has been growing in recent years, many questions still remain regarding the occurrence, etiology, and clinical significance of these multiple-type infections (1, 7). For example, there are relatively few longitudinal estimates of how many women develop coinfections over time, as well as how many different HPV types women with coinfections typically acquire. Similarly, it remains unclear why certain women acquire coinfections and other women do not, as well as whether the risk factors for HPV coinfections are the same as for HPV monoinfections. Finally, and perhaps most importantly, there is still some debate over the clinical implications of coinfections, in terms of whether coinfections convey any additional risk of cervical carcinogenesis, the ultimate concern for all HPV-infected women. As such, whether the presence of multiple types of HPV should be considered a viral cofactor of cervical cancer risk remains undetermined at this time.

These lingering issues form the rationale for this project, which aims to examine the epidemiology of HPV coinfections among young, sexually active women focusing on the occurrence, determinants and dynamics of coinfections in a cohort of female university students. Of particular interest is whether coinfections may be a marker of increased immunological susceptibility to HPV, as has been recently proposed (11), or whether coinfections are simply the result of increased sexual activity.

### 1.2 Thesis objectives

To this end, this thesis project has three main research objectives:

I. To describe the occurrence of HPV coinfections over time, using several measures of prevalence and incidence;

II. To examine both behavioural and immunological determinants of HPV coinfections; and

III. To explore the clinical significance (if any) of HPV coinfections among young women.
1.3  Overview of study design

Subjects for this research come from the McGill-Concordia Cohort, a prospective cohort study of the natural history of HPV infection and cervical neoplasia among 621 female university students in Montreal, Quebec (12). Recruitment took place between November 1996 and January 1999 among female students attending either the McGill or Concordia University health clinics. Participants were eligible for the study if they planned to remain in Montreal for at least two years and had not been treated for cervical abnormalities in the past 12 months. All eligible women were asked to return to the clinic every 6 months over a 2-year period, for a total of 5 visits each. At each visit, a cervical specimen was collected for HPV DNA testing and Pap cytology, and participants completed a questionnaire about lifestyle and sexual behaviour.

1.4  Thesis organization

This thesis is written in the traditional style and contains five chapters. Following this initial introductory chapter, the second chapter provides a broad overview of the epidemiology of HPV and highlights the literature on HPV coinfections. The third chapter summarizes the methods used in this project, including the study design, data collection process and statistical analysis techniques. The fourth chapter presents the results of this thesis, organized by research objective. Finally, the fifth chapter contains a discussion of our findings and evaluates the strengths, limitations and significance of this project. Additionally, there are six appendices, containing two supplementary tables (Appendix A and B), the ethics approval for this thesis (Appendix C), the consent form filled out by participants (Appendix D), and the two study questionnaires (Appendix E and F).
Chapter 2

Literature Review

This chapter is intended to be a broad introduction to the literature necessary to understand the epidemiology of HPV coinfections. The chapter begins with a brief overview of the epidemiology and etiology of cervical cancer, with a focus on the role played by HPV. Next, an introduction to HPV is provided, focusing on its classification, role in carcinogenesis, and methods of detection. The last two sections highlight the epidemiology of HPV infections in general, and HPV coinfections in particular. Where possible, emphasis is placed on studies conducted among younger women, as these are of greatest relevance to this thesis.

2.1 Cervical cancer: Overview of epidemiology and etiology

2.1.1 Distribution and global burden

Worldwide, cervical cancer is the second most common cancer among women, with over half a million incident cases and over 300 000 attributable deaths estimated in 2007 (13, 14). Occurrence of the disease varies widely by geographic region, however, with the burden disproportionately concentrated in less developed areas, where over 80% of the world’s cases occur (14). In regions with the highest incidence, namely Latin America, sub-Saharan Africa, Southern Asia and the Caribbean, the age-standardized incidence rates of cervical cancer are at least triple the rates in Western Europe and North America (>25 per 100 000 v. <10 per 100 000, respectively) (13).

Though mortality rates are substantially lower than incidence rates, cervical cancer tends to affect women at a younger age than many other cancers, and as such, can be a major cause of potential years of life lost (PYLL). In fact, in the developing world, where it accounts for 15% of
all female cancers, cervical cancer is the single largest cause of PYLL to cancer (15). In some regions, particularly Latin America, Eastern Europe and the Caribbean, there are more PYLL to cervical cancer than to tuberculosis or AIDS (15). Moreover, the societal importance of the disease is accentuated by the fact that it often affects multiparous women who are still raising their families, the death of which can have major social consequences for a community (16-18).

As in most industrialized countries, rates of cervical cancer are quite low in Canada with an age-standardized incidence rate of 7 per 100,000 in 2010 (19). Overall, cervical cancer was the 13th most common cancer among Canadian women with 1300 incident cases and 370 deaths attributable to the disease in 2010 (19). These relatively low rates are largely a testament to the success of cervical cancer screening programs introduced in the last 50 years, as before this, rates of cervical cancer in most developed countries were comparable to the rates seen in developing nations today (13, 15). Worth noting, however, is that not all Canadian women have experienced equal declines in rates of cervical cancer. Like many diseases, cervical cancer takes a particularly heavy toll on Aboriginal populations, comprising up to 35% of cancers among Inuit women and 29% of cancers among Aboriginal women in Saskatchewan (20, 21). Similarly, the mortality rate from cervical cancer is estimated to be 4 to 6 times higher among Aboriginal women than among white women in Canada (22, 23). Other subgroups that are at an elevated risk of cervical cancer include immigrant women, elderly women, rural women and women of low socio-economic status (23-25). Despite the fact that these vulnerable populations represent a relatively small proportion of Canadian women, they represent the majority of women at risk for developing cervical cancer (23).

2.1.2 Human papillomavirus: a necessary cause

Though the correlation between cervical cancer and sexual activity has been observed since the 1800s, the identification of the sexually transmitted etiologic agent underlying the association has been a relatively recent achievement (26). Indeed, the central causal role played
by infection with human papillomavirus (HPV) was not unequivocally recognized until the late 1990s/early 2000s, following important advances in molecular biology that permitted the detection of HPV DNA in cervical samples with improved sensitivity and specificity (1). In particular, a landmark study which used a highly sensitive method of HPV detection to analyze an international sample of over 1000 tumour specimens revealed that HPV DNA was, in fact, present in 99.7% of cervical cancer cases worldwide (27). At the same time, several observational epidemiologic studies began to report relative risks (RR) in the triple digits (>150) for the association between HPV and cervical cancer—risk estimates that are among the largest ever found in cancer epidemiology (28). As such, due to the international consistency in the linkage and the sheer magnitude of the association, it has been determined that oncogenic HPV infection is a necessary cause of cervical cancer—the first such relationship to be demonstrated in cancer research (5, 27, 29, 30). In practical terms, this means that cervical cancer does not and will not occur without the presence of HPV in the cervix (5). As no other human cancer has been found to have such a clear necessary cause, this finding has had important implications for both primary and secondary prevention of the disease (1).

2.2 Human papillomavirus: Overview and role in cervical carcinogenesis

2.2.1 Taxonomy and classification

Papillomaviruses (PVs) are a highly diverse group of viruses that infect several different species of mammals and birds, though they have been most intensely studied in humans (4). Like all PVs, HPVs are epitheliotropic viruses that infect either the skin or mucosal lining of the anogenital and oral tracts (4, 18). The HPV virion has a double-stranded circular DNA genome of approximately 8000 base pairs with eight overlapping open reading frames (ORFs) divided into early (E) and late (L) genes as well as an untranslated long control region (3, 31). Despite its
small size, its molecular biology is quite complex: three oncogenes (E5, E6, and E7) disrupt the cell cycle and modulate the transformation process, two regulatory genes (E1 and E2) control transcription and DNA replication, and two structural genes (L1 and L2) encode the major and minor capsid proteins of the virus (4, 31). Infection occurs when HPV comes into contact with the basal cells of the cervical epithelium, most likely through micro-abrasions or microscopic tears in the epithelial tissue (7). Following infection, the virus begins to replicate in the infected cells and advances into the upper layers of the epithelium as the infected cells mature (7).

More than 150 different genotypes (types for short) of HPV have been discovered so far, of which more than 40 infect the mucosa of the genital tract (3). As the L1 ORF is the most conserved gene in the HPV genome, new types of HPV are identified on the basis of L1 ORF DNA sequence homology with currently existing types (4). A new type is classified when DNA sequence homology differs by more than 10% from previously identified types (4). Within a type, subtypes and genetic variants can also be distinguished, differing in sequence homology by 2-10% and less than 2%, respectively (4).

Types of HPV are classified according to their biological niche (mucosal or cutaneous), oncogenic potential and phylogenetic grouping (4, 31). In terms of oncogenic potential, types are classified as either oncogenic/high-risk (HR) or non-oncogenic/low-risk (LR) based on their association with cervical carcinomas and associated precursor lesions (28). Over the past 15 years, up to 18 different types have been considered oncogenic (HPV16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, 59, 68, 73 and 82) or probably oncogenic (HPV26, 53, and 66) (28, 32). Most recently, however, combined results from three large international samples have indicated that 8 types in particular (HPV16, 18, 45, 31, 33, 35, 52, and 58) are consistently the most frequently detected in cervical tumours worldwide (2). Of these, HPV16 and 18 alone contribute to approximately 70% of cervical cancer cases globally (2, 32). HPV types are also classified based
on their phylogeny, as it has been shown that the oncogenic potential of the virus is directly correlated with phylogenetic groupings of viral genotypes (33). In fact, all HPV types that are known to be oncogenic belong to a single genus (Alpha group 2) all descending from a common oncogenic ancestor (34). HPV classification is a continually evolving area of research, however, and there is ongoing debate regarding the status of so-called “weakly carcinogenic” types (e.g. HPV53 and 66) (34, 35).

2.2.2 HPV and cervical carcinogenesis

2.2.2.1 Natural history

Crucial to cervical carcinogenesis is an area of the cervix known as the transformation zone, so named because it is where the squamous epithelial cells of the ectocervix are replaced by the columnar epithelial cells of the endocervix (18, 36). For reasons that are not entirely understood, HPV seems to target the transformation zones between different types of epithelium as the virus has been found to cause several human cancers at these sites: cervical, anal and oropharyngeal, for example (18). Further illustrating the importance of the transformation zone is the fact that HPV infections are equally common in both cervical and vaginal specimens, but vaginal cancers are extremely rare (18).

 Though HPV infection is necessary for cervical cancer, it is not a sufficient cause of the disease, indicating that the etiology of cervical carcinogenesis is more complex than HPV alone (6). In the simplest view, there are three important steps in the natural history of cervical carcinogenesis: (1) infection with HPV, (2) progression of infection to precancerous lesions, and (3) invasion of the cervical basement membrane (Figure 2.1) (37). Importantly, though these steps generally follow each other sequentially, each step does not necessarily lead to the one succeeding it, as HPV infections often clear spontaneously and cervical lesions may regress (37). In fact, it is estimated that overall, <1% of HR-HPV infections will progress to invasive cervical cancer (2,
HPV persistence, a critical intermediate step between infection and progression, will be discussed in detail in a subsequent section.

**Figure 2.1 - Overview of cervical carcinogenesis. Adapted from (6) and (37).**

2.2.2.2 Cofactors

On account of the central role played by HPV in cervical carcinogenesis, many of the factors classically associated with cervical cancer (e.g. sexual activity) have now been re-categorized as risk factors for HPV infection (1, 26). Nevertheless, the fact that not all infections with HPV develop into cervical cancer strongly suggests that there are other factors also involved in progression of the disease (6). Because these factors cannot independently result in cervical cancer (and are thus mainly important among HPV-infected women), they are often referred to as cervical cancer cofactors. Such cofactors may act at various stages of cervical carcinogenesis (Figure 2.1) through mechanisms that increase HPV susceptibility, influence infection persistence and/or promote oncogenic progression (1). All cofactors can be grouped into characteristics either of the host or of the virus. As was mentioned earlier, HPV type is a particularly important prognostic factor determining cervical cancer risk, however, other viral factors such as HPV DNA load and coinfection with multiple types (the major focus of this thesis) may also influence cancer risk (2).
In terms of host cofactors, there is consistent evidence to show that tobacco smoking, multiparity and long-term use of oral contraceptives (OCs) can substantially increase the risk of cervical cancer among women with HPV (6, 18, 39). Three recent collaborative re-analyses of data from over 16,000 women with cervical cancer (and over 30,000 women without) have attempted to quantify the increased risks, using all available studies and accounting for HPV infection status where possible. According to the results of this collaboration, current smokers (v. nonsmokers), women with >6 full-term pregnancies (FTP) (v. 1-2 FTP), women who were <17 at first FTP (v. ≥25), and women who use OCs for >5 years (v. never) all have nearly double the risk of developing cervical cancer (RRs ranging from 1.76-1.95), assuming they become infected with HPV (40-42).

Biological mechanisms for these increased risks are plausible but remain undetermined. Smoking, for example, may have a direct carcinogenic effect on the cervix, as nicotine metabolites have been identified in the cervical mucous of smokers (21). Tobacco smoking could thus potentially exert its influence by contributing to additional DNA damage in HPV-infected cervical cells or by reducing immune responses in the cervix, thereby increasing susceptibility to HPV infection (1, 41). In addition, elevated concentrations of hormones (as with OC use and during pregnancy) as well as delivery-related cervical trauma each result in greater exposure of the cervical transformation zone, which may make the cervix more vulnerable to HPV (18, 39, 40, 42). Similarly, immunosuppression caused by pregnancy could also favour acquisition of HPV, or promote its oncogenic potential after an infection is established (6, 42). As well, dietary factors (e.g. consumption of fruits and vegetables) and coinfection with other STIs (particularly Chlamydia trachomatis) are also suspected to be cofactors, however, more conclusive evidence is needed in order to confirm their role (6, 43, 44).

In addition to these largely modifiable risk factors, there is also evidence to suggest there is a genetic component to cervical cancer risk. For example, in a study using data from the
Swedish cancer registry (45), it was found that both biological mothers and sisters of women with cervical cancer were at significantly increased risk of also developing the disease (RR=1.83; 95% CI: 1.77-1.88 and RR=1.93; 95% CI: 1.85-2.01, respectively). By comparison, no increased risk of cervical cancer was observed among adoptive mothers and sisters of cervical cancer cases. Of the genetic factors that may potentially be involved in cervical cancer risk, some of the strongest evidence has been found for genes in the major histocompatibility complex (MHC), as these genes play a crucial role in the immune response and in determining an individual’s resistance to infections (3). Known as the human leukocyte antigen (HLA) system in humans, the MHC comprises a cluster of highly polymorphic genes that encode for various cell-surface HLA molecules, which are divided into several different classes, Class Ia (A, B, C), Class Ib (E, F, G) and Class II (DR, DQ, DP) (46-48). Class Ia molecules are typically found on all nucleated cells in the body and are responsible for presenting endogenous antigens to cytotoxic T lymphocytes (CD8+ T cells) (46, 47). In contrast, Class II molecules are generally only found on professional antigen-presenting cells (e.g. macrophages, dendritic cells) and mainly function to present exogenous antigens to helper T cells (CD4+ T cells) (46, 47). In addition to T cells and antigen-presenting cells, HLA Class Ib molecules also interact with surface receptors on natural killer cells, and thus may also play an important role in regulation of the innate immune response (48).

Through their role in antigen presentation, HLA molecules play a critical role in cell-mediated immune recognition and viral clearance, and as such, there is strong biological plausibility for the HLA system to play a role in HPV-related disease (46). Theoretically, HLA molecules that “present” HPV antigen with high affinity will be more protective, whereas HLA molecules that do not recognize and bind HPV antigens as efficiently will be associated with increased cervical disease (46). Thus far, the most convincing evidence for HLA alleles has come from studies on cervical neoplasia and cancer (3, 46, 48). In these studies, several Class II alleles have been consistently linked to risk of cervical disease. In particular, the HLA-DRB1*13 alleles
have shown a consistent protective effect (ORs significantly <1.0) for cervical lesions and carcinomas, whereas the DQB1*03, DRB1*15:01 and DQB1*06:02 alleles have shown only a somewhat consistent effect of increased risk of cervical disease (ORs >1.0, but not always significant) (3, 46). In general, there are less consistent results for Class I genes, but these have been studied less frequently due to an early research bias in the availability of robust genotyping assays for these alleles (46). Thus far, results from several studies (48, 49) have found both positive and negative associations between cervical neoplasia and a variety of Class I genes (particularly HLA-B and HLA-G), however, additional studies are needed for more conclusive results (46).

2.2.2.3 Screening and new prevention approaches

Because it takes many years for HPV infections to progress to invasive carcinomas, cervical cancer is an ideal disease for screening, and prevention programs have been very successful thus far. Classical prevention of cervical cancer has involved Pap test screening, a form of secondary prevention; however, the establishment of the HPV-cervical cancer link has recently led to the development of two new approaches in cervical cancer prevention: screening for HPV DNA, and immunization against HPV infection (50).

Pap testing, also known as cytology screening, involves taking a small sample of cells from the cervical transformation zone, smearing the sample onto a glass slide and sending it to a cytology laboratory where it is examined to determine if the cells are normal (51). The current terminology used to report and classify cytology diagnoses is the Bethesda system, introduced in 1988 and updated in 2001, though a prior classification system based on cervical histology is also still used to classify cytological precursors (52-54). Differences in the various classification terminologies can be seen in Table 2.1. The most important changes initiated by the Bethesda system were the combination of cervical intraepithelial neoplasia (CIN) 2 and 3 categories into one category called “high-grade squamous intraepithelial lesion” (HSIL) as well as the
an introduction of “atypical squamous cells of undetermined significance” (ASCUS) as a category for cervical smears with equivocal squamous epithelium changes (54).

### Table 2.1 – Classification of cytological precursors. Adapted from (53), p. 251 and (9), p.4.

<table>
<thead>
<tr>
<th>Dysplasia terminology</th>
<th>Histology: CIN(^a) terminology</th>
<th>Cytology: Bethesda terminology</th>
</tr>
</thead>
<tbody>
<tr>
<td>normal</td>
<td>normal</td>
<td>within normal limits</td>
</tr>
<tr>
<td>atypia/inflammation</td>
<td>CIN 1</td>
<td>ASCUS(^b)</td>
</tr>
<tr>
<td>mild dysplasia</td>
<td>CIN 2</td>
<td>LSIL(^c)</td>
</tr>
<tr>
<td>moderate dysplasia</td>
<td>CIN 3</td>
<td>HSIL(^d)</td>
</tr>
<tr>
<td>severe dysplasia</td>
<td>invasive cervical carcinoma in situ</td>
<td>invasive cervical carcinoma in situ</td>
</tr>
</tbody>
</table>

\(a\): Cervical intraepithelial neoplasia;  
\(b\): Atypical squamous cells of undetermined significance;  
\(c\): Low-grade squamous intraepithelial lesion;  
\(d\): High-grade squamous intraepithelial lesion

However, despite its effectiveness at reducing rates of cervical cancer, there are many limitations associated with Pap cytology, namely, that its accuracy depends on optimally collected cervical samples and is based on subjective interpretations of cellular abnormalities (55). In fact, though the specificity of the Pap test is estimated to be 98%, the average sensitivity of a single test is estimated at just over 50%, meaning that false-negative results occur in nearly half of specimens (21). As such, the success of Pap screening is largely a result of repeated testing at regular intervals, so any false-negatives can be caught on subsequent tests (50, 56). Unfortunately, this need for frequent repeated screening cycles makes Pap testing relatively expensive and burdensome on the health care system, as well as impractical in resource-poor settings (50).

Fortunately, the causative link between HPV and cervical cancer has paved the way for new HPV-targeted approaches to prevention. One such approach is screening for HPV DNA, a technique that involves testing cervical cells for the presence of oncogenic HPV types (31). One commercially available assay for HPV screening is known as Hybrid Capture II (HCII), a test which is reported to be much more sensitive (96.1%) although slightly less specific (90.7%) for the detection of high-grade cervical lesions (CIN2+), compared to Pap cytology (57). Moreover, a
cluster randomized control trial in a rural Indian setting (58) has recently shown that a single round of HPV testing using HCII is associated with a significant reduction in both the incidence of and deaths from advanced cervical cancers, suggesting that screening for HPV DNA has the potential to save many lives in the developing world where the burden of cervical cancer is highest (59).

A second new approach has been the development of two prophylactic vaccines against certain HPV types, both representing major breakthroughs in the primary prevention of cervical cancer. One vaccine is quadrivalent (Gardasil, Merck and Co, Inc) and offers protection against the HPV types that cause 70% of cervical cancers (HPV16 and 18) as well as 90% of genital warts (HPV6 and 11) (60). The other vaccine (Cervarix, GlaxoSmithKline) is bivalent and protects against HPV16 and 18 only (60). To date, both vaccines have proved highly efficacious at preventing cervical and other anogenital lesions caused by HPV16 and 18 (61). As well, significantly decreased rates of genital warts are already being observed among young women (and young men, due to herd immunity) in Australia (62). Although it will be many years before any impact on rates of cervical cancer are seen and questions remain regarding the duration of vaccine-induced HPV immunity, it is thought that these vaccines, in combination with HPV testing, represent great advances in our ability to control and prevent cervical cancer (55).

2.2.3 Diagnosis of HPV infections

HPV cannot be grown in conventional cell cultures, and serological assays have only limited accuracy, as not all women develop antibodies to the virus after natural infection (31). As a result, accurate diagnosis of HPV infection relies on the detection of viral nucleic acid in exfoliated epithelial cells. Importantly, HPV DNA testing only reflects current HPV status at the time of the test, which means it is not necessarily a measure of cumulative lifetime exposure to the virus (31, 63). HPV detection can be performed using the same cervical specimen collected for cytological examination, which is an important logistical aspect of its use both clinically and
in epidemiological studies (31). Though HPV DNA can be detected using various techniques, nucleic acid amplification methods are most often used, as they are able to detect HPV with the greatest sensitivity and specificity (1, 31).

Mentioned already, the most basic method of HPV DNA detection is a test known as Hybrid Capture II (HCII). Specifically, this approach tests for the presence of the 13 most common HR-HPV types in a sample of cervical cells, and can thereby distinguish whether a HR-HPV infection is present, but not which HPV type(s) in particular (31). As such, approaches that can detect the presence of specific HPV types, namely methods using polymerase chain reaction (PCR), have typically been used in epidemiologic natural history studies. PCR-based methods are highly sensitive and can detect <10 copies of HPV genome in a sample of cervical cells (12, 31). By means of a thermocycling process, PCR involves repeatedly amplifying certain sections of DNA using oligonucleotide primers in the presence of a thermostable DNA polymerase (31). The most widely used PCR techniques are those which employ consensus (general) primers for the PCR reaction, enabling a broad-spectrum of HPV types to be amplified in the same reaction. Three different designs of general PCR primers can be used for broad-spectrum detection of HPV DNA: GP5+/6+, MY09/11 and most recently, PGMY09/11 and SPF10 (31). When comparing the results of different studies it is important to consider which PCR primers have been used, because this can affect which HPV types (and the multiplicity of types) that are subsequently detected (64). For example, certain primers are known to preferentially amplify some HPV types over others, which can have an effect on the HPV type distributions that are reported in different studies (2, 65). Moreover, and perhaps most importantly in the context of this thesis, certain primers (e.g. PGMY09/11) may be more sensitive for the detection of multiple HPV types in the same specimen (66, 67).
2.3 Human papillomavirus: Epidemiology

2.3.1 Prevalence and incidence

HPV is the most commonly detected STI in the world and most sexually active individuals will be exposed to the virus at some point in their lifetime (1, 2). Previous studies have estimated the prevalence of HPV among asymptomatic women in the general population to be between 2 to 44% (1, 68), and recent data from a large, global meta-analysis of nearly 160 000 women found that 10.4% (95% CI: 10.2-10.7) of women worldwide have HPV DNA present in their cervix at a given point in time (69). HPV16 is the most prevalent type in all world regions with a point prevalence of 2.6% (95% CI: 2.5-2.8) (2). HPV18 is the second most common type globally, though it exhibits some variation between continents. In North America, for example, HPV18 is fourth most common after HPV16, HPV53 and HPV52 (2).

Similar to rates of cervical cancer, HPV prevalence is significantly higher in less developed regions (13.4%; 95% CI: 13.1-13.7) than in more developed regions (8.4%; 95% CI: 8.3-8.6) (2), suggesting that the high cervical cancer burden in the developing world is not entirely due to lack of screening. Rates of HPV infection also vary greatly according to age, however, and typically peak among young adults after the initiation of sexual activity. In all regions of the world except Asia, the prevalence of HPV is as high as 30% among women under age 25, then steadily decreases among older age groups (2). Interestingly, a second peak in HPV prevalence among middle-aged women has also been observed in some populations (most strikingly in Latin America and Europe) to which the cause is unclear (2, 70-72). Reasons for the second peak have been debated, but recent research on global sexual patterns (73) and HPV reinfection among older women (74) has revealed that the most likely cause is new sexual partners among women aged 40 and over, possibly related to increasing rates of divorce and remarriage occurring among this age group in many countries.
Prospective cohort studies collecting multiple measurements of HPV status over time have been used to assess the incidence of cervical HPV infections in several different populations. In general, incidence has either been expressed as the cumulative probability of HPV infection over time or as an incidence rate using person-time as a denominator. Results from these studies have typically revealed that rates of new HPV infections are quite high among women who are initially HPV-negative (1). As with prevalence, HPV incidence is also highest among younger women soon after the onset of sexual activity. For example, in several cohort studies of young women conducted in England (75), the US (76, 77), and Canada (78), the cumulative incidence of HPV infection approached or exceeded 40% after 3 years among individuals who were HPV-negative at enrollment. HPV incidence also usually declines with increasing age, exhibiting a sharp decrease after age 30 (2). However, despite the fact that prospective cohorts are usually an optimal approach for assessing incidence, an important caveat of these studies is that when an HPV type appears for the first time in a woman, we cannot be sure that the type is actually “new” and not a reactivation of a latent infection that occurred prior to enrollment (1). Naturally, the best way of overcoming this limitation is to recruit young women before they are sexually active or shortly thereafter, as this provides greater assurance that an incident HPV infection is truly new, however this is not always feasible.

2.3.2 Risk factors for HPV infection

Several key risk factors for HPV infection have been established from epidemiological studies and can be broadly grouped into factors affecting either exposure or susceptibility to the virus. As virtually all genital HPV infections are sexually transmitted (79), the main route of exposure to HPV is through sexual activity. As such, the most important determinants of HPV infection are consistently found to be sexual behaviours, such as early age at first intercourse, having many sexual partners or having partners with numerous sexual partners (1, 2, 79, 80). Number of sex partners is, in effect, a proxy measure of HPV exposure (81) as having more
partners increases the chances of encountering an HPV-infected partner and provides more exposure opportunities for transmission. Typically, number of recent partners is found to be stronger risk factor than number of lifetime partners, as infections from partners in the more distant past tend to clear and become undetectable over time (82). Importantly, however, a woman’s risk of HPV exposure is determined not only by her own sexual behaviour, but also the sexual behaviour of her male partner(s), reflecting the sexual networks that underlie STI transmission (2, 79, 83). Among women with only one lifetime sex partner for example, the risk of exposure to HPV (and also of developing cervical cancer) can be linked directly to the number of lifetime sex partners of their partner/husband (2, 83).

Condom use, another factor that may play a role in exposure to HPV, does not seem to have the strong protective effect it does for other STIs (1, 84). Moreover, while many studies have found equivocal results for HPV and condom use, some studies have paradoxically found condom use to be associated with increased risk of HPV infection. Explanations for this finding have largely been attributed to a form of indication bias, whereby condoms tend to be used with partners that are considered ‘higher risk’ (e.g. sex workers), although issues of measurement and misclassification can also not be ruled out (1, 85). Interestingly, particularly in light of the findings for condom use, several studies have found male circumcision to be an effective way of reducing HPV infection among women (37). In a recent randomized trial, for example, female partners of circumcised men had significantly lower prevalence (prevalence rate ratio=0.72; 95% CI: 0.60-0.85) and incidence (incidence rate ratio=0.77; 95% CI: 0.63-0.93) of HPV than female partners of non-circumcised men (86).

Besides markers of sexual activity, the most consistent risk factor for HPV infection is young age. Indeed, the peak HPV occurrence observed among young adults is thought to be due to their unique position of being both maximally exposed to the virus (through more sexual partners) as well as maximally susceptible to it (2). This susceptibility is thought to come from
two sources: increased cervical ectopy (87) and lack of HPV acquired immunity (88). Greater cervical ectopy refers to a topography of the cervix such that there is increased exposure of the cervical transformation zone, a situation which is thought to increase vulnerability to HPV infection, but which tends to decrease as a woman ages (11, 36, 87). In addition, younger women also typically have minimal HPV acquired immunity compared to older women, as adaptive immune responses to HPV develop over time in response to natural infection (1). Highlighting the important role played by adaptive immunity, a study conducted among Danish sex workers (88) revealed a consistent decrease in HPV prevalence corresponding with increasing age, despite continuously high exposure to new sex partners in this population over time. Consistent with this study, global HPV seroprevalence data indicate that HPV16/18 seroprevalence increases from young to middle age, contrary to the patterns observed for HPV DNA prevalence (63). Worth noting however, is that natural immunity likely does not last indefinitely: in a recent study of immunity among older women, previous HPV exposure did not appear to protect against reinfection with the same HPV type or with different types (74).

In addition to the trends regarding natural adaptive immunity, evidence for immunological status as a risk factor for HPV infection also comes from studies of HIV-infected individuals. In particular, HIV-infected women have been found to have extremely high prevalence of HPV (98% in one study) (89) as well as greater HPV persistence (90, 91) and increased frequency of lesion development (92, 93). As it is thought that HIV-related immune suppression both decreases the ability of the immune system to clear the virus and allows it to replicate to greater levels (89), it is strongly suspected that immunological differences between women in the general population may also modulate individual susceptibility to HPV.

As mentioned earlier, one source of such immunological variation may be polymorphisms in the HLA genes; however, there is relatively little data on the role of HLA alleles in HPV acquisition and persistence. Nevertheless, there are a few longitudinal studies that have found
significant associations for some HLA alleles and HPV dynamics. One study conducted in a Brazilian population (94) found that several HLA Class II haplotypes were significantly associated with HPV persistence. Another study from a Canadian population (47) found that the DRB1*13 allele group was significantly associated with cumulative HPV positivity (OR=2.6, 95% CI: 1.3-5.2), HR-HPV positivity (OR=2.8; 95% CI: 1.3-5.9) and HPV16 positivity (OR=3.3; 95% CI: 1.4-8.2) among women at high risk for HPV exposure. A more recent study conducted among the same study population also found that women with the HLA-G*01:01:02 allele had significantly increased risk of both HPV16 infection (OR=2.10, 95% CI: 1.11-3.96) and HPV16 persistence (OR=2.07, 95% CI: 1.16-3.68) (95). Despite the fact that the most consistent associations to date have been found for HLA alleles and risk of cervical neoplasia/cancer, these results suggest that it is possible for HLA molecules to be involved in immunological susceptibility to HPV infection more “upstream” of cervical disease. Additional studies in diverse populations are needed, however, before conclusions can be drawn.

As with HLA alleles, factors such as smoking, OC use and multiparity have also been studied as possible determinants that may mediate susceptibility to HPV infection, due to the consistently increased risks of cervical cancer found for these factors. To date, results from individual studies that have looked at these factors have been largely inconsistent, possibly due to differences in study populations, statistical methods and/or HPV detection techniques (1, 79, 80). Further complicating the issue, all three factors are strongly associated with sexual activity and it can be very difficult to adequately control for the confounding effect of sexual behaviour without large sample sizes (1, 96, 97).

However, recent findings based on a global sample of over 10,000 women from the IARC prevalence studies may provide some of the most comprehensive data available on these associations to date. According to these studies, only smoking was associated with increased risk of HPV-positivity—not high parity, early age at first FTP or long-term OC use (96, 97). Based on
these data, it is plausible that tobacco smoking may influence the natural history of HPV at several stages, whereas hormonal factors may predominantly influence the progression and malignant transformation of HPV (96). As smoking has been reported to impair immune responses in the cervix (98-100), it is hypothesized that this increased risk may operate through some form of immune modulation (97).

2.3.3 HPV persistence

Evidence from numerous natural history studies has revealed that the vast majority (90-95%) of HPV infections are transient and are cleared by the immune system within a short period of time (1, 2). As such, HPV persistence should be viewed as a relatively rare but critical step in cervical carcinogenesis. In fact, persistent infection with a HR-HPV type is now widely believed to be a necessary early precursor event in the carcinogenic progression to invasive cervical cancer (1). Despite its virtually unequivocal importance, however, there has been surprisingly little consensus as to what constitutes a persistent HPV infection, which has complicated the process of comparing findings and drawing conclusions between studies (1, 7). The most common definition used in the literature has typically been ≥2 HPV-positive visits, however, ≥3 HPV-positive visits, proportion of HPV-positive visits, and HPV-positivity throughout follow-up have also been used (101). Unfortunately, when defined by number of visits, persistence of infection(s) will vary markedly depending on the time interval between visits (7). For example, the shorter the interval, the more likely an infection will be deemed persistent (7).

Further complicating the issue of defining persistence are differences between studies in categorization of HPV infection, such as whether the analysis is restricted to type-specific infections, as well as oncogenic types, individual types or HPV positivity in general (101). Assessment of type-specific persistence is likely of particular importance, as it is persistence of the same HR-HPV type over time that is thought to be the main driver of cervical carcinogenesis. Furthermore, including subjects whose original HPV infection was replaced by a new HPV
infection with a different type (i.e. not type-specific) will cause persistence estimates to be exaggerated.

In an effort for a more clear and consistent way of assessing persistence, some investigators have opted to study the duration of infection (i.e. time to clearance) instead, an approach which requires the utilization of time-dependent survival analysis methods (1). These duration studies have indicated that nearly all HPV infections clear within 1-2 years of initial acquisition, and that HR-HPV infections tend to have longer durations then LR-HPV, confirming the established relationship between persistence and oncogenic risk (1, 10, 78). Infections that persist for longer than 20-24 months have a greatly increased risk of progression as it has been observed that the longer an infection persists, the longer it will continue to persist (102).

As well, two recent studies have revealed results that confirm the importance of persistence despite ambiguities in its definition. The first, a systematic review and meta-analysis of 41 epidemiological studies involving 22 500 women (101), found that HPV persistence was strongly and consistently associated with CIN2+/HSIL+ regardless of differences in definitions and study methods. In this analysis, RRs for HPV persistence and cervical neoplasia ranged from 1.3 (95% CI: 1.1-1.5) to 813.0 (95% CI: 168.2-3229.2) with 92% of the RRs over 3.0. Similarly, the second study, which analyzed participants enrolled in the placebo arm of an HPV vaccine trial (103), found that 65% of young women with persistent HPV16/18 infection >6 months developed SIL/CIN, and that persistence increased the risk regardless of the definition used or whether cytology or histology was used as the outcome.

2.4 HPV coinfections

The term ‘coinfection’ refers to infection with two or more (≥2) types of HPV, either at the same point in time or at different points in time. In the literature, HPV coinfections go by
many names (e.g. multiple HPV infections, multiple-type HPV infections, repeated HPV infections, subsequent HPV infections) and are defined in several different ways, which can make comparing the results of different studies challenging. Early interest in coinfections arose from the experimental research into the development of the polyvalent HPV vaccines, as there was a need to better understand the dynamics of infection between individual HPV types (e.g. co-occurrence, acquisition, persistence) as well as whether interactions exist between different oncogenic types (9). This section will describe what is known and what remains unknown about HPV coinfections, focusing on their distribution, dynamics and risk factors.

2.4.1 Definitions

At present, three different definitions of HPV coinfections exist in the literature (Figure 2.2) (8, 104, 105). The first, concurrent coinfection, is the simplest but also the most restrictive definition, constituting the detection of ≥2 types of HPV at the same point in time, often in the same cervical specimen. This definition is the most commonly reported in the literature because it does not require longitudinal data. A second definition refers to sequential coinfection, which is defined as the detection of ≥2 types of HPV at different points in time (i.e. HPV16 detected at a given study visit and HPV18 detected at a subsequent visit). This is the definition which best reflects coinfection dynamics, in terms of whether one HPV type influences the acquisition of subsequent HPV types. The last definition, cumulative coinfection, encompasses both of the previous two definitions, and thus refers to the detection of ≥2 types of HPV at any point during a follow-up period. This definition is the most liberal and the least prone to misclassification of the coinfection outcome as it does not limit detection to any particular study visit (104). For example, a coinfection missed at one visit because only one HPV type was detected may still be captured at another visit under a cumulative definition.
2.4.2 Prevalence and incidence of coinfections

Coinfection with multiple types is a relatively common finding among individuals infected with HPV. Pooled data from the IARC prevalence surveys indicate that among HPV-positive women, the global prevalence of concurrent coinfection is 32.2% overall, ranging from 18.5% in Korea to 46.0% in Argentina (65). Like HPV infection in general, the highest rates of HPV coinfection are found among younger women. In two studies conducted among female adolescents in the United States, the prevalence of HPV coinfections ranged from 50 to over 80% among HPV-positive women (106, 107). In two studies involving women of all ages from diverse regions (Scotland and Mozambique), the prevalence of HPV coinfection was found to be lower, though still quite high, with over 40% of HPV-positive women having multiple types detected (108, 109).

In North America, an HPV prevalence study conducted among women aged 14-59 in the NHANES 2003-2004 population found that among HPV-positive women the prevalences of coinfection with 2 or with ≥3 HPV types were 24% and 16%, respectively (110). A study from
Montreal by Richardson et al (78) reported that 38% of HPV-positive women had a concurrent coinfection with both HR-HPV and LR-HPV types; however, no Canadian studies have explicitly examined the prevalence of HPV coinfections using all possible combinations of HPV types. Though there is relatively little data on HPV coinfection among men, there is some evidence to suggest that the prevalence is similar to that found among women. For example, one study conducted among men aged 18-40 from Arizona found that 27.0% of all men and 41.9% of HPV-positive men tested positive for multiple types (111).

Relatively few studies have reported on the incidence of HPV coinfections, as most studies have been cross-sectional investigations without longitudinal data. In a large cohort of Brazilian women, Rousseau et al (105) found that the incidence of concurrent coinfections was 1.73 per 1000 woman-months (95% CI: 1.30-2.16) among all women and 4.18 per 1000 woman-months (95% CI: 2.6-5.8) among women aged 18-24. Two longitudinal studies conducted among younger women in England and New Jersey reported cumulative incidences of sequential coinfection with a new HPV type of 26% and 70%, respectively, among women HPV-positive (for a different HPV type) initially (75, 112). One recent study that looked at the incidence of repeated HPV infections after an initial infection found that 69% of women had a second infection, and of those with a second infection, 63% had a third infection within 3 years (11). Currently, there are no such data on the incidence of HPV coinfections in Canada.

2.4.3 Coinfection Dynamics

2.4.3.1 Clustering

Many of the early longitudinal studies involving coinfections were designed to investigate coinfection dynamics, as there have been concerns about the possibility of HPV type replacement following vaccination against certain HPV types (8, 105, 113, 114). Type replacement, a microbial population dynamics phenomenon, refers to the elimination of some types of HPV
(such as those targeted by current vaccines), followed by an increase in the incidence (and potentially, the carcinogenicity) of previously less common HPV types (2, 115). For type replacement to occur however, two conditions would have to apply: first, there must be partial competition between individual HPV types during natural infection, and second, HPV vaccines must not provide any cross-protection against types naturally competed against by types included in the vaccines (e.g. HPV16 and 18) (115).

The best way to examine natural competition between types is to determine whether certain HPV types tend to be found together (clustering) or not found together in coinfections, as the latter scenario would indicate some degree of competition. As such, several studies have examined whether clustering exists in natural infection, with a particular focus on the types in current HPV vaccines. For example, one longitudinal study that focused on type-specific clustering reported excess risks for several two-type combinations involving vaccine-covered types (116). In particular, it was found that infection with HPV6 or 11 at baseline yielded a 14-fold increase (OR=14.1; 95% CI: 2.1-95.4) in subsequent HPV18 infection. Similarly, infection with HPV16 or 18 at baseline was associated with a nearly 6-fold increase (OR=5.7; 95% CI: 2.2-15.1) in subsequent HPV58 infection. Though these associations are strongly indicative of clustering, it should be noted that they are not suggestive of type replacement, as this would be indicated by protective associations against subsequent infections. In general, however, the literature on clustering remains inconclusive: while several studies have found evidence of clustering between certain HPV types and phylogenetic groupings (113, 116-118), there has been little type-specific consistency in the findings between studies, and others have found no significant associations at all (8, 119). Recently, a pooled analysis of data from nearly 14 000 women has led to the suggestion that type-specific clustering may actually be an artifact of the method used to detect HPV DNA, as the clustering results in that study differed depending on the genotyping method that was used (65).
Though type-specific clustering remains somewhat unclear, clustering of HPV infections within women in general has been consistently observed both cross-sectionally and longitudinally and when using different methods to detect HPV DNA. In other words, rather than offering protection against subsequent HPV infections (as would be expected due to adaptive immunity) several studies have shown that having one (or more) HPV infection(s) increases a woman’s risk of acquiring additional infections with new HPV types (8, 11, 112-114). In a cohort of Hawaiian women, for example, Goodman et al (120) found that the risk of several incident HR-HPV types was significantly increased among women with ≥2 coexisting HPV infections. In particular, the RR of incident HPV53 was 10.0 (95% CI: 2.3-42.8) and the RR of incident HPV16 was 7.0 (95% CI: 2.1-23.5) among women with ≥2 HPV types compared to women with monoinfections.

Taken together, the current evidence on coinfection dynamics does not suggest that type replacement will occur, although long-term data on HPV type distributions in vaccinated women would provide the best assurance of this fact (115). The evidence does suggest, however, that the occurrence of multiple HPV types is concentrated within certain women, and that these women are at increased risk of further HPV acquisition. This latter fact, which has been generally attributed to the common transmission routes (e.g. sexual activity) shared by all HPV types, has raised additional questions as to why certain women appear to acquire more HPV infections than others.

2.4.3.2 Persistence

Similar to the findings regarding clustering, there is a continued debate about the relationship between HPV persistence and detection of coinfections. First, however, it is important to note that these two issues are fundamentally intertwined, as a woman who is unable to clear her HPV infection (and thus has persistence) is more likely to have multiple types of HPV detected at any given point in time, should she continue to acquire new HPV infections (11). Despite this link, many of the earlier studies on this relationship found conflicting results, with as
many studies finding positive associations between persistence and coinfections (75, 77, 121) as those finding no associations (102, 113, 114). This discrepancy has since been attributed to variations in definitions and statistical methodologies for assessing persistence (e.g. logistic regression v. survival techniques) as well as important differences between study populations (e.g. cytologically normal v. exclusively abnormal women) (10).

Encouragingly, however, two recent studies with at least 5 years of follow-up may have begun to elucidate the matter. The first study, which looked at durations of type-specific infections among Brazilian women, found that coinfection with multiple types was associated with significantly increased duration of infection (10). The second study, which looked at the incidence of repeated infections among women in California, found that persistence of an initial HPV type to the point of detection of a second HPV type was the strongest risk factor for acquiring a second HPV infection (HR: 4.51; 95% CI: 3.78-5.37) and, in particular, a much stronger risk factor than number of new sex partners (11). Perhaps most importantly, however, both sets of investigators contend that rather than coinfections leading to persistence (or vice versa), it is more likely that both outcomes may actually be an indication of immune dysfunction in a woman, such that she is both unable to clear her infections and more susceptible to additional HPV infections upon exposure to the virus. Notably, this hypothesis is consistent with the evidence regarding coinfection acquisition, whereby multiple infections appear to cluster within women, and where women with prior HPV infections tend to acquire additional infections over time. Thus, rather than coinfections being entirely due to higher-risk sexual behaviours, as has been suggested (112, 120), it is possible that the development of coinfections may also reflect some degree of increased biological susceptibility to the virus.

2.4.4 Risk factors for coinfections

As with determinants of HPV infections in general, the most important risk factors for HPV coinfections also seem to be younger age and markers of sexual activity. Given the finding
that coinfections seem to concentrate within certain women, however, an important question has been whether these women are simply exposed to HPV at greater frequencies (such as through additional sexual partners), or whether these women are actually inherently more susceptible to HPV upon “normal” exposure to the virus. In simple terms, several studies have tried to answer these questions by examining whether there are specific risk factors for coinfections in addition to the risk factors for HPV infections more generally.

To date, only six studies have explicitly examined risk factors for HPV coinfections among women, and virtually all of them have found positive associations for number of sexual partners and younger age (11, 70, 71, 104, 112, 122). The one study that did not find an association with age was conducted among a population of university students, and as a result, likely had an age range that was too narrow to detect important age differences (112). Other factors that appeared in some studies but not in others include: OC use (112), condom use (70), previous STI infections (11, 122), and being Hispanic, Black, or mixed-race (11, 112). These discrepancies are most likely attributable to the vast differences between study populations in terms of age, socio-demographic factors and location, as well as in study design. The most important design limitation is probably that many of these studies have been cross-sectional, a design which underestimates the cumulative diversity of HPV types that may be acquired over time, and which may thus lead to misclassification of coinfection status (104, 123). Other important limitations of these studies include: small numbers of events (71), limited coinfection definitions (11, 71, 112), and older study populations with a lower risk of acquiring HPV coinfections (70, 71, 104).

As well, given the hypothesis that women who acquire coinfections may have increased susceptibility to HPV, surprisingly little attention has been given to studying factors associated with the immune response. In addition to the high prevalence of coinfections observed among young women with relatively little exposure to HPV, evidence for a relationship between
coinfections and immunity comes from the fact that HPV coinfections are extremely common among individuals with HIV, with nearly 80% prevalence and an average of >3 HPV types per woman observed in one study (89). Despite this theoretical link, only one study published to date has examined aspects of the immune response and risk of HPV coinfection (112). That study, conducted among college-age women in the US, found that women who had persistent antibodies (detected at 2 consecutive visits, 12 months apart) to HPV16 had a reduced risk of subsequent HPV infections (sequential coinfections). Interestingly, the association was strongest for types genetically related to HPV16 and weakest for unrelated types. Though the finding with related types was attributed to cross-protection, the investigators hypothesized that the association with unrelated types may have been due to persistent antibodies being a marker of the host’s ability to mount a strong, effective immune response against any HPV type. This explanation is in accordance with the idea that women who experience coinfections may have some degree of immune dysfunction, and warrants further investigation in additional studies.

2.4.5 Coinfections and cervical carcinogenesis

Similar to the findings for persistence, there has been some debate about whether HPV coinfections increase the risk of developing cervical lesions and cancer. Though the majority of studies have found positive associations between coinfections and cervical disease (105, 109, 119, 123-127), some studies have found no greater risk among those with coinfections compared to those with monoinfections (28, 108, 128, 129), particularly for monoinfections involving HPV16 (130). Due to the nature of these outcomes, however, many of these studies have not been longitudinal, as prospective studies are limited by the fact that cervical disease cannot be ethically allowed to progress once precancerous lesions are identified. As such, even when the two outcomes are detected together, non-longitudinal designs make it difficult to ascertain whether the coinfection(s) actually preceded lesion development. Furthermore, as mentioned earlier, cross-
sectional designs also may be prone to underestimating the presence of coinfections since exposure assessment only reflects one point in time.

Worth noting, however, two of the most recent studies—both sufficiently large and methodologically robust—have found convincing evidence that there is a relationship between coinfections and cervical lesions. One study, a longitudinal analysis of over 2100 women (123), found extremely strong associations (OR>250) for coinfections with 4-6 HPV types and HSIL, even after excluding HPV16. The second study, conducted among more than 1300 women undergoing colposcopy (125), found that having ≥3 types of HPV or multiple HR-HPV were each strongly associated with CIN2+/carcinoma: OR=817 (95% CI: 251-2655) and OR=270.3 (95% CI: 101.2-722), respectively. Despite these convincing results, it remains unclear whether multiple types play a causal, possibly synergistic role in the progression of cervical disease, or simply happen to be more commonly present in women with cervical lesions/carcinomas due to increased HPV susceptibility in these women. In either case, there seems to be growing evidence that HPV coinfections may be important markers of disease, either directly or indirectly.

2.5 Summary and Rationale

In summary, though HPV is a necessary cause of cervical cancer, it is not a sufficient cause, and only a small proportion of HPV-infected women will go on to develop the disease. As a result, much of the HPV research in recent years has attempted to uncover other cofactors that may identify women at greater risk of progression after infection, and HPV coinfections may be one such factor. Though it is known that coinfections are common, particularly among young women, many of the issues surrounding HPV coinfections remain poorly understood, and it is possible that these women—or at least a proportion of them—may represent a subset of women with increased susceptibility to HPV.
Although there is fairly consistent evidence to suggest that HPV-infected women often continue to acquire additional HPV types over time, there is relatively little data describing how many HPV infections women typically acquire, especially longitudinally as it is thought that the cumulative diversity of types to which women are exposed may play a role in risk of further disease (123). In addition, there are still many gaps in our understanding of the etiology of HPV coinfections, particularly why certain women experience them and others do not. Though increased sexual activity is clearly involved, the role of the immune response is a relatively unexamined risk factor for which there is strong biological plausibility. Lastly, the clinical implications of HPV coinfections also remain unclear, as the relationship(s) between coinfections, HPV persistence and cervical neoplasia are far from resolved. Through an investigation of the occurrence, determinants and dynamics of HPV coinfections, this thesis will attempt to address many of these lingering questions, providing a comprehensive examination of the epidemiology of multiple HPV infections among young women.
Chapter 3

Methods

3.1 Purpose and objectives

The purpose of this thesis was to examine the occurrence, determinants and dynamics of HPV coinfections in a cohort of female university students. Specifically, there were three main objectives that, while separate from each other, come together to create an overall picture of the epidemiology of coinfections among young women. These objectives are as follows:

I. To describe the prevalence and incidence of HPV coinfections, at different time points and using multiple definitions;

II. To examine baseline and time-dependent determinants of HPV coinfections, including HLA alleles as potential immune factors; and

III. To explore the clinical implications (e.g. persistence and cervical neoplasia) associated with HPV coinfections in young women.

A conceptual diagram of how these objectives fit into the overall model of cervical carcinogenesis can be seen in Figure 3.1.

Figure 3.1– Conceptual diagram of thesis objectives
3.2 Study design

3.2.1 Overview

This thesis project is based on a secondary analysis of data collected in the McGill-Concordia cohort study, a prospective study of the natural history of HPV infection and cervical neoplasia among female university students in Montreal, Quebec (12). Recruitment of the cohort took place between November 1996 and January 1999 and follow-up was completed in November 2001. Women were asked to participate in the study over a 2-year period with pre-scheduled return visits every 6 months, for a total of 5 study visits. At each visit, participants were asked to complete a self-administered questionnaire and cervical specimens were collected for Pap cytology and HPV testing. HLA genotyping was performed on DNA extracted from cervical specimens collected at enrollment.

3.2.2 Study population and accrual of subjects

Women were recruited into the study while attending either the McGill University or Concordia University Health Clinics. A nurse practitioner was present at each site and was responsible for introducing the study to all females waiting to see a nurse or physician in each clinic’s reception area. The study was also advertised in university newspapers and radio stations and in person (by H. Richardson) to first-year classes at McGill. Women were eligible for the study if they planned to remain in Montreal for at least two years and had not been treated for cervical disease or had an abnormal Pap smear in the year prior to enrollment. Eligible participants were asked to sign a consent form, fill out a personal data sheet and complete a baseline questionnaire. To improve subject compliance at the pre-scheduled return visits, women were informed that they would receive $20 for each return visit that they completed (12).

In total, 635 women originally consented to be in the study, however, 14 women were subsequently withdrawn because they did not return for the second visit and either did not fill out
the baseline questionnaire (13 women) or had a cervical sample that could not be analyzed (1 woman) at the enrollment visit (78). This left 621 women with complete data for analysis. On average, women returned every 7 months for their follow-up visits with approximately 10% of the cohort lost to follow-up at each visit (78). Of the original 621 women, 424 (68%) completed all 5 pre-scheduled visits. For the purposes of this thesis, these 621 women contributed a total of 2675 visits and 13767 months of follow-up.

3.3 Data collection

3.3.1 Laboratory procedures

As described previously by Richardson et al (12, 78), samples of endocervical and ectocervical cells were simultaneously collected using two Accelon cervical biosamplers (Medscand Inc, Hollywood FL) from each woman at each visit. A Pap smear was prepared with the first sampler, and the cervical cells collected with the second sampler were used for HPV DNA testing. After collection, the plastic sampler tips from the second sampler were transferred to a tube of saline-buffer that was subsequently agitated to release the exfoliated cervical cells. After this, the sampler tips were discarded and the tubes were frozen for transport to the laboratory of Dr. François Coutlée at the Université de Montréal, who was responsible for overseeing HPV DNA testing (12).

HPV DNA detection

Details of HPV DNA detection have been described previously by Coutlée et al (131). In brief, when the samples arrived at the laboratory, they were first purified using QIAamp columns (Qiagen), followed by PCR amplification of β-globin DNA (with PC04 and GH20 primers) to verify the absence of inhibitors and the integrity of processed DNA (78). If β-globin DNA could not be detected in a sample, the DNA sample was considered inadequate for further HPV testing.
ß-globin positive specimens were further tested for the presence of HPV DNA using a PCR protocol with consensus primers MY09/11 and HMB01. These primers target the highly conserved L1 region of the HPV genome and amplify a 450 base pair fragment (31). This detection system is a highly sensitive DNA amplification method that can detect as few as 1-10 molecules of HPV in a genital sample, while amplifying a broad spectrum of HPV types in the same reaction (12). The amplification process occurred in a TC 9600 thermal cycler as follows: amplification of *AmpliTaq Gold* DNA polymerase at 95°C for 9 minutes; 95°C denaturation for 1 minute, annealing at 55°C for 1 minute, extension at 72°C for 1 minute for 40 cycles; and 5 minute terminal extension at 72°C (131). Genotyping of the HPV PCR products was done using the reverse line blot method (Roche Molecular Systems), whereby the HPV amplicons were hybridized to an array of immobilized oligonucleotide probes on a single “strip” (131). This technique enabled the identification of 27 different HPV genotypes (HPV6, 11, 16, 18, 26, 31, 33, 35, 39, 40, 42, 45, 51, 52, 53, 54, 55, 56, 57, 58, 59, 66, 68, 73, 82, 83, 84) fixed as distinct lines on each strip. For additional details regarding specific laboratory procedures, please refer elsewhere (12, 131).

Halfway through the duration of the study the MY09/11 primers were redesigned to create a new set of primers known as the PGMY09/11 system (12). This new system was reported to be an improvement over previous primer sets, as it appeared to be slightly more sensitive for HPV detection overall as well as for the detection of multiple HPV types, though this latter difference was not statistically significant (66, 67). Despite the proposed advantages of this new system, it was not feasible to change the PCR protocol in the middle of follow-up since the earlier test results would not have been comparable to the test results from the latter half of the study (12). Moreover, as the MY09/11 primer set has been used extensively in many epidemiological studies published to date, the HPV results from the McGill-Concordia Cohort are directly comparable to much of the current HPV literature.
Pap cytology

Using the cells from the first Accelon biosampler, Pap smears were prepared by fixing the cells onto glass slides with 95% ethanol. The cytology slides were processed and read by Ms. Juliette Robitaille in the laboratory of Dr. Alex Ferenczy at the Jewish General Hospital in Montreal (12). Classification of abnormalities was based on the Bethesda system as described in Table 2.1. The cytology reports were photocopied and the original was kept in the student’s medical file. Women with a diagnosis of HSIL (high-grade lesion) were immediately referred for colposcopy and did not complete any additional study visits. Women with a diagnosis of ASCUS or LSIL continued to be followed every 6 months for monitoring, in accordance with standard clinical practice guidelines at the time of the study. If, however, a women was LSIL+ at two consecutive visits she was also censored and referred for colposcopy (personal communication, Dr. Richardson).

HLA genotyping

Due to the high cost of genotyping at the time of the original study, not all 621 women in the cohort could be genotyped for all HLA alleles. As a result, the decision was made to restrict genotyping to women who met one of two conditions: (a) women who participated in the study for at least 3 visits (~18 months) or (b) women who had a persistent HPV infection with the same HPV type for two consecutive visits (47, 132). Women who had β-globin negative DNA were also excluded (47). At the time of this analysis, complete data was available for 548 women in the cohort for all HLA alleles under investigation.

As has been described previously (47, 95), HLA alleles were typed from purified DNA obtained from the cervical specimens collected at enrollment, using PCR amplification of sequence-specific primers. Several different HLA alleles/allele groups were typed: HLA-B*07, DQB1*03, DQB1*06:02, DRB1*13, DRB1*15:01, E*01:01, E*01:03, G*01:01:01, G*01:01:02,
G*01:01:03, G*01:01:05, G*01:01:07, G*01:01:08, G*01:03, G*01:04:01, G*01:04:03, G*01:05N. Slightly different laboratory procedures were used in the genotyping of each allele, the details of which can be found elsewhere (47, 95, 132). Briefly, oligonucleotide primers designed on the basis of HLA polymorphic sequences were used to amplify specific Class Ia, Class Ib and Class II alleles using PCR analysis (47, 95, 132). Negative and positive controls were included in all amplification runs and all PCR products were sequenced in both directions (47). Visualization of the PCR products was done through electrophoresis performed on 1% agarose gel stained with ethidium bromide and visualized under ultraviolet illumination (47).

3.3.2 Questionnaires

Content and procedures

To investigate the role of risk behaviours associated with acquisition and dynamics of HPV infection(s), women were asked to complete a self-administered questionnaire at each study visit. The content of the questionnaire was based on questions developed and validated by the National Cancer Institute and the IARC for the measurement of different dimensions of sexual behaviour and hygiene as well as the use of contraceptives, tobacco and alcohol (12). As well, before use in this study, the questionnaire was previously tested in a cross-sectional pilot study conducted earlier in Montreal in the same target population as this cohort (12, 85). In addition to the risk factors mentioned above, the questionnaire also inquired about socioeconomic information, race, diet, reproductive history and medical history. Questionnaires were available in both English and French and participants were given the opportunity to complete their answers in a private office at the health clinic, with the study nurse available at all times to answer any questions about the questionnaire or the study (12).
Creation of variables

Two versions of the questionnaire were administered: an in-depth baseline version (Appendix E) and an abridged follow-up version (Appendix F) that was used to measure recent changes in sexual behaviour or lifestyle practices. Information from the questionnaires was used to create both time-fixed and time-dependent exposure variables for the analysis. Time-fixed variables reflected data that was only collected at the baseline visit (e.g. diet) or characteristics that did not change over time (e.g. HLA alleles). In contrast, time-dependent variables referred to behaviours or activities that could vary over time at each follow-up visit (e.g. number of sexual partners; smoking status). All time-dependent covariates captured exposure information on behaviours that had occurred since the previous study visit. At baseline, however, the corresponding question may have referred to behaviour(s) throughout the subject’s adult lifetime or in specific time periods such as the year prior to enrollment. A visual representation of the data collection process can be seen in Figure 3.2.

Figure 3.2 – Overview of data collection procedures

<table>
<thead>
<tr>
<th>Procedures &amp; instruments</th>
<th>Entry</th>
<th>6</th>
<th>12</th>
<th>18</th>
<th>24</th>
</tr>
</thead>
<tbody>
<tr>
<td>Laboratory:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HPV DNA detection &amp; typing</td>
<td>✓  ✓  ✓  ✓  ✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pap testing</td>
<td>✓  ✓  ✓  ✓  ✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HLA genotyping</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Questionnaire:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Diet</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sociodemographics</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tobacco and alcohol consumption</td>
<td>✓  ✓  ✓  ✓  ✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sexual history</td>
<td>✓  ✓  ✓  ✓  ✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contraceptive history</td>
<td>✓  ✓  ✓  ✓  ✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Medical history</td>
<td>✓  ✓  ✓  ✓  ✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
3.4 Defining coinfections

One of the challenges of this project was deciding what would constitute a coinfection and which definition(s) of coinfections would be included. Though the definitions presented earlier in Figure 2.3 are fairly straightforward, in practice, the dynamics of HPV acquisition, clearance and re-acquisition are rarely so clear-cut. As can be seen in Figure 3.3, for example, the same coinfection can be deemed sequential (HPV31 followed by HPV16) or concurrent (HPV31 and HPV16 at the same visit) depending on how one chooses to view the situation. To get around this ambiguity, some authors that have studied sequential coinfections have chosen to ignore visits with multiple overlapping types (e.g. visit 3, Figure 3.3) and only measure coinfections that occur in separate visits (8, 112). However, due to previous work in this cohort (78), it was known a priori that there were many overlapping infections in this study population, so we did not feel that this would be an appropriate approach to take with this data. As a result, it was ultimately decided to focus on two definitions: concurrent coinfection, to take advantage of the high number of overlapping infections in the dataset; and cumulative coinfection, to make use of the longitudinal nature of the study and to capture the accumulation of multiple HPV types over time. As such, because sequential coinfections were not a focus of our analysis, the situation depicted in visit 3 of Figure 3.3 was counted as a concurrent coinfection in this thesis.

Figure 3.3 – The challenge of defining coinfections in practice
Both coinfection definitions were expressed dichotomously in the analyses: cumulative coinfection over the study period (yes/no) and concurrent coinfection at each study visit (yes/no) based on the occurrence of the 27 different HPV types that were genotyped in the cohort. It should be noted however, that these coinfection outcomes were not mutually exclusive, as a woman with a concurrent coinfection at one visit was also counted as having a cumulative coinfection overall, and thus contributed to the analyses for each outcome. Additionally, women with “purely” sequential (i.e. non-overlapping) coinfections were also captured in the cumulative outcome. Also of note, though the term “outcome” has been used here for simplicity, both coinfection variables also acted as exposures for the analyses in objective 3.

3.5 Statistical Analyses

Two statistical packages were used for the analyses in this thesis: SAS (Version 9.2, SAS Institute, Cary, NC) and SPSS (Version 19, IBM Corporation, Somers, NY).

3.5.1 Objective 1: Occurrence of coinfections

Prevalence of coinfections

For this objective, the occurrence of HPV coinfections was described in several ways. First, in terms of prevalence, the period prevalence of both cumulative and concurrent coinfections over follow-up was described. The prevalence of cumulative coinfections was calculated as the number of women in which ≥2 types of HPV were detected at any point during follow-up, divided by the total number of women in the cohort (n=621). Similarly, the prevalence of concurrent coinfections was calculated as the number of women in which ≥2 types of HPV were detected at the same point in time, divided by the total number of women in the cohort. In the same method as above, the prevalence of cumulative and concurrent coinfections among HPV-positive women (i.e. women in which HPV was detected at any point during the study) was
also calculated. Additionally, the total number of HPV types accumulated over follow up and at individual study visits for each woman was reported. In an effort to assess whether some types of HPV were found more commonly as coinfections, the cumulative type-specific occurrence of each type overall and stratified into monoinfections (1 type only) or coinfections (≥2 types) was also calculated.

*Incidence of coinfections*

In terms of incidence, the incidence rate of concurrent coinfections was calculated as the number of women in which ≥2 types of HPV were *newly* detected at the same point in time over follow-up (i.e. not at enrollment), divided by follow-up time in woman-months. The rates were then stratified by HPV infection status at enrollment to see if women who were already infected with one type of HPV acquired a higher rate of coinfections over time. In an effort to extend this analysis, and include all women in the cohort, the incidence of any new HPV infection was also calculated stratifying on enrollment status. For this analysis, women were grouped into HPV-negative, HPV monoinfection or HPV coinfection strata at enrollment and followed until they acquired a new type of HPV that was not present initially (or their first type, for the HPV-negative group). For the coinfection group, this would correspond to at least their third type of HPV, although it may have been greater depending on how many types were detected at the baseline visit. Using these same stratifications, the Kaplan-Meier technique was used to estimate the cumulative probabilities of acquiring either a concurrent coinfection or an infection with a new HPV type as a function of follow-up time. Log rank tests were performed to assess whether the cumulative probabilities of acquisition differed based on enrollment infection status. Kaplan-Meier actuarial analysis was also used to calculate the mean time to acquisition of each outcome for each stratified group. Finally, to assess the possibility of bias related to detection opportunity, a set of additional analyses was performed whereby the incidence rates for each outcome were grouped and stratified by the number of visits each woman completed.
3.5.2 Objective 2: Determinants of coinfections

For objective 2, potential determinants of HPV coinfections were investigated using two different approaches. First, determinants of cumulative coinfection were investigated using exposure variables assessed at baseline as potential predictors of developing a coinfection over follow-up. The goal of this analysis was to identify more long-term, non-modifiable risk factors, such as those that do not vary over time. Second, potential determinants of concurrent coinfections at individual study visits were also examined using a repeated measures analysis with both baseline and time-dependent exposure variables as candidate predictors. The goal of the repeated measures model was to attempt to explain the presence of a coinfection at a given visit using updated data from the time interval prior to that visit. In particular, it was hoped that by capturing the effects of recent time-dependent exposures, we might identify modifiable risk factors that could be adopted (or avoided) to decrease the risk of developing coinfections over time.

Unconditional multivariate logistic regression was used to estimate associations between exposure variables and each dichotomous coinfection outcome, for both sets of analyses. For each outcome, women with coinfections were contrasted with HPV-negative women and women with HPV monoinfections in separate models, based on cumulative infection status (cumulative outcome) and infection status at each visit (concurrent outcome). This approach was taken in order to gain the most information possible about the etiology of HPV coinfections as we hypothesized that each comparison might yield slightly different results. In particular, we hypothesized that the coinfection v. HPV-negative comparison would yield the most “extreme” results as these groups of women are likely to be the most different from each other. Here, for example, is where we thought sexual activity might play the biggest role, since it represents the major route of exposure to HPV, and the most important determinant of infection(s). By contrast, we hypothesized that sexual activity would be less important in the coinfection v. monoinfection
(all HPV-positive) comparison and that this might be where more subtle determinants of risk (e.g. immunological factors) might emerge.

For the model involving repeated measures and time-dependent variables, the generalized estimating equations (GEE) extension of logistic regression was used to adjust the ORs and standard errors for the within-subject correlation between multiple visits by the same individual (133). As it was assumed that observations from visits closer together in time would be more correlated than visits farther apart, a first-order auto-regressive working correlation matrix was used in the SAS PROC GENMOD procedure. Although we originally intended to include all concurrent coinfections in this analysis, we later decided that prevalent coinfections should be excluded as these outcomes would reflect the effect(s) of baseline exposures that were already captured in the cumulative outcome model. As such, the repeated measures analysis only examined determinants of incident concurrent coinfections.

Potential determinants

The choice of which exposures to investigate as potential determinants was made on the basis of the published literature as well as on what was available in our dataset. With the exception of the HLA alleles, all information on exposures was obtained from the baseline and follow-up questionnaires described previously. A list of all exposure variables that were initially under consideration, and their conceptualizations, can be seen in Table 3.1. For the HLA alleles, a ‘dominant’ allele approach was used, such that women who were homozygous or heterozygous for a given allele were compared to women in whom the allele was absent.
### Table 3.1 Overview and conceptualization of exposure variables

<table>
<thead>
<tr>
<th>Variable</th>
<th>Conceptualizations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age at enrollment</td>
<td>continuous / grouped as 17-20, 21-23, 24-26, ≥270</td>
</tr>
<tr>
<td>Race</td>
<td>White, Black, Hispanic, Asian</td>
</tr>
<tr>
<td>Age at first intercourse</td>
<td>continuous / grouped as &lt;16, 16-18, ≥19</td>
</tr>
<tr>
<td>Number of lifetime sex partners§</td>
<td>continuous / grouped as 0-1, 2-4, 5-9, ≥10</td>
</tr>
<tr>
<td>Number of recent sex partners§</td>
<td>continuous / grouped as 0, 1, ≥2</td>
</tr>
<tr>
<td>Number of new sex partners§</td>
<td>continuous / grouped as 0, 1, ≥2</td>
</tr>
<tr>
<td>Anal sex§</td>
<td>never, ever</td>
</tr>
<tr>
<td>Oral contraceptive use§</td>
<td>regularly, sometimes, never</td>
</tr>
<tr>
<td>Duration of OC use§</td>
<td>never, &lt;1 year, 1-5 years, &gt;5 years</td>
</tr>
<tr>
<td>Condom use§</td>
<td>regularly, sometimes, never</td>
</tr>
<tr>
<td>Frequency of sex per week§</td>
<td>&lt;1 time, 1-3 times, 4-6 times, ≥7 times</td>
</tr>
<tr>
<td>Recent vaginal irritation§</td>
<td>no, yes</td>
</tr>
<tr>
<td>History of yeast infection§</td>
<td>no, yes</td>
</tr>
<tr>
<td>History of STIs§</td>
<td>(Includes trichomoniasis, chlamydia, herpes, syphilis, gonorrhea, genital sores)</td>
</tr>
<tr>
<td>Vegetable consumption§</td>
<td>≥1 day, &gt;1 week, 1 per week, rarely</td>
</tr>
<tr>
<td>Alcohol consumption*§</td>
<td>continuous / grouped as 0, 1-3, ≥4</td>
</tr>
<tr>
<td>Cigarettes smoked per day§</td>
<td>nonsmoker, &lt;1, 1-5, &gt;5</td>
</tr>
<tr>
<td>Pack years smoked§</td>
<td>continuous / grouped as nonsmoker, &lt;1, 1-2, &gt;2</td>
</tr>
<tr>
<td>HLA alleles</td>
<td>absent, present</td>
</tr>
</tbody>
</table>

**Class Ia**

- HLA-B*07

**Class Ib**

- HLA-E*01:01, E*01:03; HLA-G*01:01:01, G*01:01:02,
  - G*01:01:03, G*01:01:05, G*01:01:07, G*01:01:08,
  - G*01:03, G*01:04:01, G*01:04:03, G*01:05N

**Class II**

- HLA-DQB1*03, DQB1*06:02; HLA-DRB1*13, DRB1*15:01

§In lifetime

◊ In lifetime prior to enrollment

* In the last year prior to enrollment

† In the past 5 years

‡ Exposures for which time-dependent data was available

** Time-dependent exposures to which there was a cumulative component (i.e. values at each subsequent visit were added to values from previous visits)
A conceptual model for the proposed role of these potential risk factors in the development of HPV coinfections can be seen in Figure 3.4. In the model, the determinants under consideration are grouped on the basis of whether they were thought to modulate exposure or susceptibility to HPV, and subsequently by the putative mechanisms through which they may operate. The relative size of the open-ended arrows reflects our hypotheses (as stated above) regarding where we predict each set of risk factors to play a more dominant role.

**Figure 3.4 – Conceptual model of risk factors for HPV coinfections in the cohort**

![Conceptual Model](image)

**Model building**

In the cumulative model, model building began by constructing bivariate logistic regression models for baseline exposures and cumulative HPV outcomes using SAS PROC LOGISTIC. Variables with a significant Wald test at p<0.25 were then considered as candidates for inclusion in the multivariate models. For continuous variables, linearity in the logit was tested using the Box Tidwell transformation (134). Using age as an example, this test involved creating
a new variable of the form age*\log(\text{age}) and entering it into the model along with the original age variable. Significance of the term at p<0.05 indicated non-linearity, in which case continuous variables were modeled categorically. Using this approach, pack years smoked was the only candidate variable modeled continuously (though it was not retained in any of the final models). Selection of independent variables into the final multivariate models was based on backward elimination using a significance level of 0.1. This cut-off point was chosen to be slightly more liberal than the conventional 0.05, so as to hopefully strike a balance between possibly excluding any moderately important predictors and having too many significant variables in the models. The final models thus included all risk factors that were significant predictors of each respective outcome, mutually adjusted for one another. Where no dose-response was evident and similar trends were observed in different levels of a variable, categories were collapsed in the final models for simplicity. The Hosmer-Lemeshow goodness-of-fit test was used to assess the fit of the final multivariate logistic models (133).

After the construction of the cumulative outcome models, significant baseline variables from this analysis were used as a starting point for the repeated GEE models, under the assumption that baseline exposures would not be associated with incident coinfections if they were not associated with coinfections in general. In addition to the baseline exposures, all variables for which there was time-dependent data were also considered as possible predictors in the model. For a coinfection at time \( t \), a time-dependent variable referred to an exposure in the interval between visit \( t-1 \) and visit \( t \). Final multivariate GEE models were again selected using backward elimination at a significance level of 0.1.

In each analysis, Spearman correlation coefficients were calculated \textit{a priori} to assess correlations between all significant (p<0.25) candidate variables. When correlations of >0.7 were observed between variables, only one variable was selected to be entered into the multivariate models, usually the variable considered most important or “best” representing a given mechanism.
Similarly, in the repeated measures analysis, it was decided to keep number of lifetime sex partners as a time-fixed variable assessed at baseline, as otherwise it would be inherently correlated with number of new sex partners if both variables were allowed to vary over time at each visit. Finally, conventional collinearity diagnostics were run in SAS PROC REG on the final independent variables (with cut-offs of TOL <0.2, VIF >10 and/or conditional index >30 indicating collinearity) to assess any possible multicollinearity in the models.

3.5.3 Objective 3: Clinical implications of coinfections

Infection duration

Several approaches were used to assess the clinical implications of coinfections in this cohort. First, we attempted to study the relationship between coinfections and duration of infection, to investigate whether coinfections were associated with greater HPV persistence. In this analysis, the Kaplan-Meier technique was used to obtain actuarial estimates of the mean duration of type-specific HPV infections, considering specific types individually and also grouped by oncogenic risk. Importantly, the Kaplan-Meier method allowed us to account for censoring events, such as those women who had not cleared their infections by the end of the study. Infection duration was defined as the time to clearance of the longest type-specific infection, which was measured from the first visit in which the longest enduring HPV type was detected until the visit in which the same type was not detected, or for censored observations the last available visit. Durations were calculated separately for all infections and for incident infections only, with the latter being the longest type-specific infection newly acquired after enrollment. In the grouped durations, women with both HR-HPV and LR-HPV could contribute to the mean durations for both groups, using the infection duration corresponding to each respective type. In these groupings, HR-HPV referred to the following types: HPV16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, 59 and 68; LR-HPV referred to any other types that were detected (12). In the overall (any HPV) group, each woman could only contribute her longest enduring type-specific infection.
To assess the effect of coinfections, mean durations were stratified on whether infections were detected in isolation (monoinfections) or as cumulative coinfections with other HPV types during follow-up. Additionally, durations were also stratified on the cumulative number of types detected (1 type, 2-3 types, 4+ types) to assess whether there might be some threshold of effect among women with a higher number of types. Kaplan-Meier survival curves were generated for all comparisons and the log rank test was used to assess whether the time to clearance was significantly different between the individual strata.

**Occurrence of cervical lesions**

In addition to infection duration, we also wanted to investigate whether coinfections were associated with development of cervical squamous intraepithelial lesions (SILs) in this cohort. This was somewhat exploratory, however, as previous research in this cohort had revealed there was a very low number of SILs in this study population, many less than were originally expected (12). Nevertheless, several analytic approaches were used to study the relationship between HPV coinfections and risk of cervical lesions. Two outcomes were used in all analyses: occurrence of SIL cumulatively over follow-up and repeatedly at each study visit. If a woman had more than one cytologic outcome over the course of the study, the worst grade lesion outcome was used for the cumulative assessment.

To examine the effect of coinfections, we contrasted women with coinfections with both HPV-negative women and women with monoinfections separately, based on infection status over follow up and at individual study visits, with lesion outcomes assessed at the corresponding period of observation. Next, we stratified HPV infection status into number of types (1 type, 2-3 types, 4+ types) in contrast to women who were HPV-negative (0 types) both cumulatively and concurrently. In the concurrent models, lesions outcomes were evaluated at two time points, the same visit and the next visit (~6 months later) to explore temporal relationships between coinfections and lesion development. In the latter approach, when using the 6-month lag between
coinfection and lesion occurrence, the first visit was excluded for all women because coinfection status could not be assessed for the visit prior to baseline.

In all analyses, unconditional logistic regression was used to estimate ORs and 95% CIs for the associations between coinfections and lesion outcomes, while adjusting for age at enrollment and race as these have been observed to be important covariates of lesion risk (2). As occurrence of lesions was modeled dichotomously, ASCUS results were counted in the ‘normal’ cytologic category, and not as SIL outcomes. Because the concurrent models treated each visit as an independent observation, the GEE extension of logistic regression was used to account for the clustering within individuals due to repeated observations from the same women. A first-order auto-regressive working correlation matrix was also used again, as it was assumed that the occurrence of both coinfections and lesions would be more similar at closer time points. In the models that stratified on number of types, p-values for linear trend were obtained by fitting the categorical exposure variables as continuous variables with an ordinal scale corresponding to number of HPV types: 0 = 0 types, 1 = 1 type, 2 = 2-3 types, 3 = 4+ types.

3.5.4 Missing data

Treatment of missing data generally followed the same procedures as were taken for previous analyses in this cohort (12). Missing HPV results typically occurred when a sample had inadequate DNA for analysis due to being β-globin negative. This occurred in very few cases, however, as nearly all of the cervical specimens (97.6%) were suitable for HPV DNA testing (78). Women with a β-globin negative result were not excluded from the analysis as long as they had at least one other visit with an adequate HPV test result. In the prevalence, incidence, and logistic regression analyses, a β-globin negative result was treated as HPV-negative, whereas in the Kaplan-Meier analyses, a conservative approach was taken whereby a β-globin negative visit was deemed “unchanged” and the next visit with an adequate HPV result was used instead (78). Similarly, women who had a missing cytology result at a specific visit were classified as
cytologically normal for that visit. Of the 2675 visits used in the analyses, there were 15 visits where cytology status was missing.

Several approaches were taken to deal with missing data from the questionnaires. First, a deductive imputation strategy was used, whereby missing values were inferred from a subject’s responses to other questions, or to the same questions at other study visits (for time-fixed variables) (135). If this was not possible, and subjects were missing items from only one section of the questionnaire (e.g. contraceptive history), then the “last observation carried forward approach” was adopted and the value from the closest preceding visit was imputed for the missing item. If no prior data was present then information from the visit immediately subsequent was used instead. In situations where no previous or subsequent data were available, subjects were assigned the mode value or the median value based on the distribution of the entire study population, for variables that were categorical and continuous, respectively. Subjects who were missing multiple items from more than one section of the questionnaire after the first imputation approach (n=11) were excluded from the analysis in objective 2.

### 3.6 Ethical considerations

The original McGill-Concordia cohort study has received ongoing ethics approval from the Research Ethics Institutional Review Board at McGill University in the years since the initiation of the study. In addition, this thesis project sought expedited ethics approval from the Queen’s University Health Sciences Research Ethics Board (Appendix C). As mentioned, informed consent was obtained from all participants prior to enrollment in the study (Appendix D), and all data was kept confidential and analyzed without subject identifiers.
Chapter 4

Results

4.1 Objective 1: Occurrence of coinfections

4.1.1 Prevalence

More than half of the cohort became infected with HPV at least once in the study and over 30% of the cohort acquired multiple types of HPV over the course of follow-up. 174 women (28.0%) had multiple types of HPV detected in the same visit (concurrent coinfection) and 202 women (32.5%) had multiple types of HPV detected over follow-up (cumulative coinfection) (Table 4.1). Among the 328 women in which HPV was detected at one or more study visits, 53.1% had multiple HPV types detected concurrently and 61.6% had multiple HPV types detected cumulatively. These proportions indicate that the vast majority of coinfections that occurred in the cohort overlapped at one or more study visits.

<table>
<thead>
<tr>
<th>Coinfection Type</th>
<th>n</th>
<th>% of cohort (n=621)</th>
<th>% of HPV+ (n=328)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Concurrent</td>
<td>174</td>
<td>28.0</td>
<td>53.1</td>
</tr>
<tr>
<td>Cumulative</td>
<td>202</td>
<td>32.5</td>
<td>61.6</td>
</tr>
</tbody>
</table>

Table 4.1 Prevalence of concurrent and cumulative coinfections in entire cohort and in women who were ever HPV-positive

Table 4.2 shows the number of HPV types detected per woman at different points over the course of follow-up. At enrollment, 449 women (72.3%) were HPV-negative, 101 women (16.3%) were HPV-positive for a single type, and 71 women (11.4%) were HPV-positive for multiple types (coinfection). The maximum number of types detected in a single visit was 8 and 77 women (12.3%) had 3 or more HPV types detected in a single visit. The maximum number of types accumulated over time was 10 and 119 women (19.1%) accumulated 3 or more different
types over follow-up. 293 women (47.2%) remained HPV-negative throughout the duration of the study.

### Table 4.2

<table>
<thead>
<tr>
<th>Number of HPV types</th>
<th>Concurrent (at enrollment)</th>
<th>Concurrent (at a single visit)</th>
<th>Cumulative (over follow-up)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n  (%)</td>
<td>n    (%)</td>
<td>n    (%)</td>
</tr>
<tr>
<td>0</td>
<td>449 (72.3)</td>
<td>293 (47.2)</td>
<td>293 (47.2)</td>
</tr>
<tr>
<td>1</td>
<td>101 (16.3)</td>
<td>154 (24.8)</td>
<td>126 (20.3)</td>
</tr>
<tr>
<td>2</td>
<td>46 (7.4)</td>
<td>97 (15.6)</td>
<td>83 (13.4)</td>
</tr>
<tr>
<td>3</td>
<td>14 (2.3)</td>
<td>49 (7.9)</td>
<td>55 (8.9)</td>
</tr>
<tr>
<td>4</td>
<td>9 (1.5)</td>
<td>21 (3.4)</td>
<td>30 (4.8)</td>
</tr>
<tr>
<td>5</td>
<td>1 (0.2)</td>
<td>4 (0.6)</td>
<td>20 (3.2)</td>
</tr>
<tr>
<td>6</td>
<td>- -</td>
<td>2 (0.3)</td>
<td>4 (0.6)</td>
</tr>
<tr>
<td>7</td>
<td>- -</td>
<td>- -</td>
<td>5 (0.8)</td>
</tr>
<tr>
<td>8</td>
<td>1 (0.2)</td>
<td>1 (0.2)</td>
<td>3 (0.5)</td>
</tr>
<tr>
<td>9</td>
<td>- -</td>
<td>- -</td>
<td>1 (0.2)</td>
</tr>
<tr>
<td>10</td>
<td>- -</td>
<td>- -</td>
<td>1 (0.2)</td>
</tr>
</tbody>
</table>

Table 4.3 describes the type-specific occurrence of each HPV type detected in the cohort in terms of monoinfections and coinfections. Consistent with the definitions of coinfections, an infection was deemed a monoinfection if it was the only type of HPV acquired by a woman over the course of the study (i.e. no other types were detected over follow-up). HPV16 was the most common type detected with 103 women (16.6%) having an HPV16 infection at some point. Similarly, 31.4% of women who became infected with HPV had an infection with HPV16. Virtually all HPV types were more commonly detected as coinfections than as monoinfections. HPV45 had the largest proportion of monoinfections detected (26.3%) and four of the five HPV types with the largest proportions of monoinfections were oncogenic types (HPV45, HPV73, HPV16, HPV58). Relative to other types, HPV73 seemed to be less commonly found as a concurrent coinfection (59.1%), and it was among the types most frequently detected sequentially (18.2%) or as a monoinfection (22.7%). Interestingly, 3 types (HPV11, HPV26 and HPV35) were only detected as coinfections, although they were among the least frequently detected types.
## Table 4.3  Type-specific HPV occurrence in single- and multiple-type infections

<table>
<thead>
<tr>
<th>HPV type</th>
<th>Monoinfections</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td># of infections</td>
<td>% of infections*</td>
<td>Cumulative (total)</td>
<td>Coinfections</td>
<td>Sequential</td>
<td>Total</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td># of infections</td>
<td>% of infections*</td>
<td># of infections</td>
<td>% of infections*</td>
<td># of infections</td>
<td>% of infections*</td>
</tr>
<tr>
<td>16</td>
<td>23 (22.3)</td>
<td></td>
<td>80 (77.7)</td>
<td></td>
<td>72 (69.9)</td>
<td>8 (7.8)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>84</td>
<td>8 (11.4)</td>
<td></td>
<td>62 (88.6)</td>
<td></td>
<td>59 (84.3)</td>
<td>3 (4.3)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>51</td>
<td>12 (19.4)</td>
<td></td>
<td>50 (80.7)</td>
<td></td>
<td>46 (74.2)</td>
<td>4 (6.5)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>53</td>
<td>8 (14.3)</td>
<td></td>
<td>48 (85.7)</td>
<td></td>
<td>40 (71.4)</td>
<td>8 (14.3)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>54</td>
<td>9 (18.0)</td>
<td></td>
<td>41 (82.0)</td>
<td></td>
<td>36 (72.0)</td>
<td>5 (10.0)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>6 (13.3)</td>
<td></td>
<td>39 (86.7)</td>
<td></td>
<td>34 (75.6)</td>
<td>5 (11.1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>8 (18.6)</td>
<td></td>
<td>35 (81.4)</td>
<td></td>
<td>32 (74.4)</td>
<td>3 (7.0)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>31</td>
<td>7 (19.4)</td>
<td></td>
<td>29 (80.6)</td>
<td></td>
<td>27 (75.0)</td>
<td>2 (5.6)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>52</td>
<td>7 (19.4)</td>
<td></td>
<td>29 (80.6)</td>
<td></td>
<td>28 (77.8)</td>
<td>1 (2.8)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>56</td>
<td>3 (8.8)</td>
<td></td>
<td>31 (91.2)</td>
<td></td>
<td>30 (88.2)</td>
<td>1 (2.9)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>66</td>
<td>6 (19.4)</td>
<td></td>
<td>25 (80.7)</td>
<td></td>
<td>22 (71.0)</td>
<td>3 (9.7)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>39</td>
<td>3 (10.0)</td>
<td></td>
<td>27 (90.0)</td>
<td></td>
<td>26 (86.7)</td>
<td>1 (3.3)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>58</td>
<td>6 (22.2)</td>
<td></td>
<td>21 (77.8)</td>
<td></td>
<td>20 (74.1)</td>
<td>1 (3.7)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>73</td>
<td>5 (22.7)</td>
<td></td>
<td>17 (77.3)</td>
<td></td>
<td>13 (59.1)</td>
<td>4 (18.2)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>33</td>
<td>1 (5.3)</td>
<td></td>
<td>18 (94.7)</td>
<td></td>
<td>17 (89.5)</td>
<td>1 (5.3)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>45</td>
<td>5 (26.3)</td>
<td></td>
<td>14 (73.7)</td>
<td></td>
<td>13 (68.4)</td>
<td>1 (5.3)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>82</td>
<td>1 (5.6)</td>
<td></td>
<td>17 (94.4)</td>
<td></td>
<td>17 (94.4)</td>
<td>0 (0.0)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>59</td>
<td>1 (7.1)</td>
<td></td>
<td>13 (92.9)</td>
<td></td>
<td>12 (85.7)</td>
<td>1 (7.1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>55</td>
<td>1 (7.1)</td>
<td></td>
<td>13 (92.9)</td>
<td></td>
<td>13 (92.9)</td>
<td>0 (0.0)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>83</td>
<td>3 (23.1)</td>
<td></td>
<td>10 (76.9)</td>
<td></td>
<td>9 (69.2)</td>
<td>1 (7.7)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>68</td>
<td>1 (8.3)</td>
<td></td>
<td>11 (91.7)</td>
<td></td>
<td>11 (91.7)</td>
<td>0 (0.0)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>0 (0.0)</td>
<td></td>
<td>9 (100.0)</td>
<td></td>
<td>8 (88.9)</td>
<td>1 (11.1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>42</td>
<td>1 (14.3)</td>
<td></td>
<td>6 (85.7)</td>
<td></td>
<td>5 (71.4)</td>
<td>1 (14.3)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>26</td>
<td>0 (0.0)</td>
<td></td>
<td>5 (100.0)</td>
<td></td>
<td>4 (80.0)</td>
<td>1 (20.0)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>35</td>
<td>0 (0.0)</td>
<td></td>
<td>2 (100.0)</td>
<td></td>
<td>2 (100.0)</td>
<td>0 (0.0)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>40</td>
<td>1 (50.0)</td>
<td></td>
<td>1 (50.0)</td>
<td></td>
<td>1 (50.0)</td>
<td>0 (0.0)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>57</td>
<td>0 (0.0)</td>
<td></td>
<td>0 (0.0)</td>
<td></td>
<td>0 (0.0)</td>
<td>0 (0.0)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
a: Infections in which only 1 HPV type was present in isolation (no other types detected over follow-up);
b: Infections in which 2+ HPV types were detected over follow-up, at any point in time;
c: Infections in which 2+ HPV types were detected at the same time (same study visit);
d: Infections in which 2+ HPV types were detected at different, non-overlapping points in time;
◊ As cumulative coinfections include both concurrent and sequential coinfections, this column represents the total number of coinfections for each type;
* Proportion of infections for each HPV type
4.1.2 Incidence

Of the 174 women who acquired a concurrent coinfection over follow-up, 71 had a prevalent coinfection at enrollment. As such, there were 103 women with an incident concurrent coinfection in the cohort, yielding an incidence rate of 9.2 (7.5-11.2) per 1000 woman-months (Table 4.4). Stratifying on whether women were positive or negative for HPV at enrollment yielded statistically significant differences in the incidence of concurrent coinfections over time. Women with an HPV monoinfection at enrollment had almost 3 times the risk of contracting an incident coinfection compared to HPV-negative women at enrollment: RR=2.7 (95% CI: 1.9-3.8). Interestingly, however, among those who acquired concurrent coinfections, the crude mean time to acquisition was not significantly different between HPV-positive and HPV-negative women at enrollment. In contrast, the actuarial mean time to acquisition showed that HPV-positive women acquired coinfections in significantly less time than HPV-negative women. The actuarial estimates are longer than the crude estimates because the former include the follow-up time elapsed among censored women who had not acquired a coinfection by the end of the study.

Table 4.5 describes the rate of acquisition of new HPV types based on HPV infection status at enrollment. Compared to HPV-negative women, women with an HPV monoinfection at enrollment had almost 2 times the risk of acquiring a new type of HPV over follow-up: RR=1.8 (95% CI: 1.4-2.2). Similarly, women with a coinfection at enrollment had almost 2.5 times the risk of acquiring a new type of HPV compared to HPV-negative women: RR=2.3 (95% CI: 1.8-2.9). For those who acquired new HPV types, the mean number of additional types acquired was 2.0 (95% CI: 1.8-2.1) with no significant differences based on infection status at enrollment (data not shown). As would be expected from the incidence rates, both the crude and actuarial mean times to acquisition were significantly shorter for HPV-positive women compared to HPV-negative women, although there were no significant differences in time to acquisition between the two HPV-positive strata based on the number of types present at baseline.
### Table 4.4  Incidence rates of concurrent coinfection (n=550)

<table>
<thead>
<tr>
<th>Enrollment HPV status</th>
<th># with new HPV coinfection</th>
<th>Woman-months at risk</th>
<th>IR (95% CI) per 1000 woman-months</th>
<th>Rate Ratio (95% CI)</th>
<th>Mean time to acquisition (months)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Crude (95% CI)</td>
</tr>
<tr>
<td>All women◊</td>
<td>103</td>
<td>11164</td>
<td>9.2 (7.5-11.2)</td>
<td>-</td>
<td>20.3 (19.5-21.1)</td>
</tr>
<tr>
<td>HPV-negative§</td>
<td>67</td>
<td>9317</td>
<td>7.2 (5.6-9.1)</td>
<td>1.00</td>
<td>20.8</td>
</tr>
<tr>
<td>HPV-positive* (monoinfection)</td>
<td>36</td>
<td>1847</td>
<td>19.5 (13.7-27.0)</td>
<td>2.7 (1.9-3.8)</td>
<td>18.3</td>
</tr>
<tr>
<td>All women without a coinfection at enrollment (n=550)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Women who were HPV-negative at enrollment (n=449)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Women with an HPV monoinfection at enrollment (n=101)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Table 4.5  Incidence rates of new HPV infection (n=621)

<table>
<thead>
<tr>
<th>Enrollment HPV status</th>
<th># with new HPV infection</th>
<th>Woman-months at risk</th>
<th>IR (95% CI) per 1000 woman-months</th>
<th>Rate Ratio (95% CI)</th>
<th>Mean time to acquisition (months)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Crude (95% CI)</td>
</tr>
<tr>
<td>HPV-negative§</td>
<td>156</td>
<td>8347</td>
<td>18.7 (15.9-21.9)</td>
<td>1.00</td>
<td>18.6</td>
</tr>
<tr>
<td>HPV-positive (monoinfection)§</td>
<td>52</td>
<td>1581</td>
<td>32.9 (24.6-43.1)</td>
<td>1.8 (1.4-2.2)</td>
<td>15.7</td>
</tr>
<tr>
<td>HPV-positive (coinfection)*</td>
<td>42</td>
<td>971</td>
<td>43.3 (31.2-58.5)</td>
<td>2.3 (1.8-2.9)</td>
<td>13.7</td>
</tr>
<tr>
<td>Women who were HPV-negative at enrollment (n=449)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Women with an HPV monoinfection at enrollment (n=101)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Women with an HPV coinfection (2+ types) at enrollment (n=71)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Figure 4.1 shows the cumulative probability curves for acquiring (A) an incident concurrent coinfection and (B) an incident infection with a new HPV type, based on infection status at enrollment. Among HPV-negative women, the cumulative rate of new concurrent coinfections was 6.9% (95% CI: 6.7-7.1) at 1 year and 15.9% (95% CI: 15.3-16.5) at 2 years. By comparison, among women with an HPV monoinfection at enrollment, the cumulative rate of new concurrent coinfections was 24.7% (95% CI: 22.6-27.0) at 1 year and 34.2% (95% CI: 31.0-37.7) at 2 years (Figure 4.1A, log rank test: p<0.0001). Similarly, after 2 years of follow-up, the cumulative probability of acquiring a new type of HPV was 35.9% (95% CI: 34.1-37.7) among women who were HPV-negative at enrollment, 53.2% (95% CI: 47.7-59.2) among women with a monoinfection at enrollment and 65.1% (95% CI: 57.3-73.9) among women with a coinfection at enrollment (Figure 4.1B, log rank test: p<0.0001).
Figure 4.1 – Cumulative probabilities of (A) concurrent coinfection and (B) infection with new HPV types based on HPV status at baseline.
In the additional analyses to assess whether incidence rates varied according to length of follow up (Tables 4.6 and 4.7), no significant differences based on number of visits completed were observed, indicating that the incidence of coinfections and of new HPV infections was not related to detection opportunity. Based on these results, we can also be reasonably confident that loss to follow-up did not differ based on coinfection status.

**Table 4.6**

<table>
<thead>
<tr>
<th>Visits completed (n)</th>
<th>New HPV coinfections</th>
<th>Woman-months at risk</th>
<th>IR (95% CI) per 1000 woman-months</th>
</tr>
</thead>
<tbody>
<tr>
<td>2 or 3 (n=68)</td>
<td>10</td>
<td>915</td>
<td>10.9 (5.2-20.1)</td>
</tr>
<tr>
<td>4 (n=70)</td>
<td>12</td>
<td>1529</td>
<td>7.9 (4.1-13.7)</td>
</tr>
<tr>
<td>5 or 6 (n=374)</td>
<td>81</td>
<td>8719</td>
<td>9.3 (7.4-11.5)</td>
</tr>
</tbody>
</table>

Note: 38 women only completed 1 visit and thus contributed no follow up time and no events

**Table 4.7**

<table>
<thead>
<tr>
<th>Visits completed (n)</th>
<th>New HPV infections</th>
<th>Woman-months at risk</th>
<th>IR (95% CI) per 1000 woman-months</th>
</tr>
</thead>
<tbody>
<tr>
<td>2 or 3 (n=81)</td>
<td>20</td>
<td>1057</td>
<td>18.9 (11.6-29.2)</td>
</tr>
<tr>
<td>4 (n=76)</td>
<td>31</td>
<td>1483</td>
<td>20.9 (14.2-29.7)</td>
</tr>
<tr>
<td>5 or 6 (n=421)</td>
<td>199</td>
<td>8359</td>
<td>23.8 (20.6-27.4)</td>
</tr>
</tbody>
</table>

Note: 43 women only completed 1 visit and thus contributed no follow up time and no events
4.2 **Objective 2: Determinants of coinfections**

Baseline demographic and sexual behaviour variables for this cohort have been reported in detail previously (47, 78). At enrollment, the mean age of the cohort was approximately 23 years (median age: 21 years, age range: 17-42 years) and the vast majority (>80%) of women described themselves as white. HPV infection status did not differ with respect to age, ethnicity or socioeconomic status, though women who became HPV-positive tended to be younger at first intercourse, and to have had more lifetime sex partners. At enrollment, nearly half of the women (45%) had had 5 or more lifetime sex partners, a proportion that increased to 60% by the end of the study.

As mentioned earlier, due to the fact that HLA information was not available for all women in the cohort, the number of subjects available for this analysis was slightly reduced. After exclusions for missing data on HLA alleles and other covariates of interest, there were 537 women and 2499 individual study visits with complete data for analysis. Additionally, 65 visits were excluded from the concurrent models because they contained prevalent coinfections at enrollment (Figure 4.2).

**Figure 4.2 – Flow charts for Objective 2**

![Flow charts for Objective 2](image-url)
4.2.1 Cumulative coinfection status: bivariate results

Table 4.8 shows the bivariate associations between baseline exposures (significant at \( p<0.25 \)) and cumulative HPV infection status. These models examined predictors of cumulative coinfections in comparison to HPV-negative women (first model) and women with HPV monoinfection (second model). Of the 537 women for whom we had complete data, there were 115 women (21.4%) with a monoinfection and 192 women (35.8%) with a coinfection detected over follow-up. 230 women (42.8%) remained HPV-negative throughout the study.

As expected, several strong associations were observed in the coinfection v. HPV-negative model, as this was the more “extreme” comparison (Table 4.8). In particular, variables that were significantly associated with coinfections compared to being HPV-negative included: age at first intercourse, number of sex partners (both lifetime and new), history of anal sex, history of yeast infections, history of STIs (particularly *Chlamydia*), alcohol consumption and cigarette smoking. For the race variable, the decision was made to combine the categories of Black and Hispanic due to low frequencies (<5 in some cells) and also because these two groups have been found to have similar patterns of coinfection risk in the HPV literature (11, 112). Though no associations were found for this grouping, women of Asian origin or descent appeared be significantly protected compared to white women. In contrast, there were far fewer variables strongly associated with cumulative coinfection compared to having an HPV monoinfection: only number of new sex partners, frequency of sex per week and the HLA-G*01:01:01 allele were significantly associated with coinfections in this model.
Table 4.8  
Frequency distributions and bivariate ORs between baseline variables and cumulative coinfection status

<table>
<thead>
<tr>
<th>Variable</th>
<th>Frequency (%)</th>
<th>OR (95% CI)</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>HPV-negative</td>
<td>Monoinfection</td>
<td>Coinfection v. HPV-negative</td>
<td>Coinfection v. monoinfection</td>
</tr>
<tr>
<td></td>
<td>n=230 (42.8)</td>
<td>n=115 (21.4)</td>
<td>n=192 (35.8)</td>
<td></td>
</tr>
<tr>
<td>Age at enrollment</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>17-20</td>
<td>96 (45.1)</td>
<td>42 (19.7)</td>
<td>75 (35.2)</td>
<td>1.00</td>
</tr>
<tr>
<td>21-23</td>
<td>61 (37.9)</td>
<td>34 (21.1)</td>
<td>66 (41.0)</td>
<td>1.39 (0.87-2.00)</td>
</tr>
<tr>
<td>24-26</td>
<td>32 (37.2)</td>
<td>21 (24.4)</td>
<td>33 (38.4)</td>
<td>1.32 (0.75-2.34)</td>
</tr>
<tr>
<td>≥27</td>
<td>41 (53.2)</td>
<td>18 (23.4)</td>
<td>18 (23.4)</td>
<td>0.56 (0.30-1.06)</td>
</tr>
<tr>
<td>Race</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>178 (40.7)</td>
<td>93 (21.3)</td>
<td>166 (38.0)</td>
<td>1.00</td>
</tr>
<tr>
<td>Black/Hispanic‡</td>
<td>19 (45.2)</td>
<td>8 (19.1)</td>
<td>15 (35.7)</td>
<td>0.85 (0.42-1.72)</td>
</tr>
<tr>
<td>Asian</td>
<td>33 (56.9)</td>
<td>14 (24.1)</td>
<td>11 (19.0)</td>
<td><strong>0.36 (0.18-0.73)</strong></td>
</tr>
<tr>
<td>Age at first intercourse</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;16</td>
<td>40 (33.3)</td>
<td>26 (21.7)</td>
<td>54 (45.0)</td>
<td>1.00</td>
</tr>
<tr>
<td>16-18</td>
<td>121 (43.4)</td>
<td>59 (21.1)</td>
<td>99 (35.5)</td>
<td>0.61 (0.37-0.99)</td>
</tr>
<tr>
<td>≥19</td>
<td>69 (50.0)</td>
<td>30 (21.7)</td>
<td>39 (28.3)</td>
<td><strong>0.42 (0.24-0.74)</strong></td>
</tr>
<tr>
<td>Number of lifetime sex partners</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0-1</td>
<td>75 (65.8)</td>
<td>19 (16.7)</td>
<td>20 (17.5)</td>
<td>1.00</td>
</tr>
<tr>
<td>2-4</td>
<td>76 (44.2)</td>
<td>36 (20.9)</td>
<td>60 (34.9)</td>
<td><strong>2.96 (1.63-5.39)</strong></td>
</tr>
<tr>
<td>5-9</td>
<td>48 (33.6)</td>
<td>30 (21.0)</td>
<td>65 (45.4)</td>
<td><strong>5.08 (2.74-9.43)</strong></td>
</tr>
<tr>
<td>≥10</td>
<td>31 (28.7)</td>
<td>30 (27.8)</td>
<td>47 (43.5)</td>
<td><strong>5.69 (2.91-11.11)</strong></td>
</tr>
<tr>
<td>Number of new sex partners§</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>120 (52.9)</td>
<td>54 (23.8)</td>
<td>53 (23.3)</td>
<td>1.00</td>
</tr>
<tr>
<td>1</td>
<td>78 (47.9)</td>
<td>31 (19.0)</td>
<td>54 (33.1)</td>
<td>1.57 (0.98-2.52)</td>
</tr>
<tr>
<td>≥2</td>
<td>32 (21.8)</td>
<td>30 (20.4)</td>
<td>85 (57.8)</td>
<td><strong>6.01 (3.58-10.11)</strong></td>
</tr>
<tr>
<td>Anal sex§</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Never</td>
<td>182 (45.7)</td>
<td>83 (20.8)</td>
<td>133 (33.4)</td>
<td>1.00</td>
</tr>
<tr>
<td>Ever</td>
<td>43 (34.5)</td>
<td>32 (23.0)</td>
<td>59 (42.5)</td>
<td><strong>1.68 (1.08-2.62)</strong></td>
</tr>
<tr>
<td>Oral contraceptive use§</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Regularly</td>
<td>156 (42.9)</td>
<td>83 (22.8)</td>
<td>125 (34.3)</td>
<td>1.00</td>
</tr>
<tr>
<td>Sometimes</td>
<td>21 (36.2)</td>
<td>10 (17.2)</td>
<td>27 (46.6)</td>
<td>1.61 (0.87-2.97)</td>
</tr>
<tr>
<td>Never</td>
<td>53 (46.1)</td>
<td>22 (19.1)</td>
<td>40 (34.8)</td>
<td>0.94 (0.59-1.51)</td>
</tr>
<tr>
<td>Variable</td>
<td>HPV-negative</td>
<td>Monoinfection</td>
<td>Coinfection</td>
<td>Coinfection v. HPV-negative</td>
</tr>
<tr>
<td>-----------------------------------------------</td>
<td>-------------------</td>
<td>------------------</td>
<td>------------------</td>
<td>----------------------------</td>
</tr>
<tr>
<td><strong>Frequency of sex/week</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;1 time</td>
<td>52 (39.1)</td>
<td>24 (18.1)</td>
<td>57 (42.9)</td>
<td>1.00</td>
</tr>
<tr>
<td>1-3 times</td>
<td>127 (44.4)</td>
<td>70 (24.5)</td>
<td>89 (31.1)</td>
<td>0.64 (0.40-1.02)</td>
</tr>
<tr>
<td>4-6 times</td>
<td>40 (40.8)</td>
<td>18 (18.3)</td>
<td>40 (40.8)</td>
<td>0.91 (0.51-1.63)</td>
</tr>
<tr>
<td>≥7 times</td>
<td>11 (55.0)</td>
<td>3 (15.0)</td>
<td>6 (30.0)</td>
<td>0.50 (0.17-1.44)</td>
</tr>
<tr>
<td><strong>History of yeast infections</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>126 (47.9)</td>
<td>53 (20.2)</td>
<td>84 (31.9)</td>
<td>1.00</td>
</tr>
<tr>
<td>Yes</td>
<td>104 (38.0)</td>
<td>62 (22.6)</td>
<td>108 (39.4)</td>
<td><strong>1.56 (1.06-2.29)</strong></td>
</tr>
<tr>
<td><strong>History of STIs</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>219 (44.9)</td>
<td>101 (20.7)</td>
<td>168 (34.4)</td>
<td>1.00</td>
</tr>
<tr>
<td>Yes</td>
<td>11 (22.4)</td>
<td>14 (28.6)</td>
<td>24 (49.0)</td>
<td><strong>2.84 (1.36-5.97)</strong></td>
</tr>
<tr>
<td><strong>Chlamydia</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>226 (44.3)</td>
<td>106 (20.8)</td>
<td>176 (34.9)</td>
<td>1.00</td>
</tr>
<tr>
<td>Yes</td>
<td>4 (14.8)</td>
<td>3 (33.3)</td>
<td>14 (51.9)</td>
<td><strong>4.44 (1.44-13.74)</strong></td>
</tr>
<tr>
<td><strong>Herpes (HSV-2)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>226 (43.4)</td>
<td>112 (21.5)</td>
<td>183 (35.1)</td>
<td>1.00</td>
</tr>
<tr>
<td>Yes</td>
<td>4 (25.0)</td>
<td>3 (18.8)</td>
<td>9 (56.3)</td>
<td>2.78 (0.84-9.17)</td>
</tr>
<tr>
<td><strong>Vegetable consumption</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(Servings)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>≥1 per day</td>
<td>47 (45.2)</td>
<td>21 (20.2)</td>
<td>36 (34.6)</td>
<td>1.00</td>
</tr>
<tr>
<td>&gt;1 per week</td>
<td>103 (40.7)</td>
<td>51 (20.2)</td>
<td>99 (39.1)</td>
<td>1.26 (0.75-2.10)</td>
</tr>
<tr>
<td>1 per week</td>
<td>64 (44.1)</td>
<td>33 (22.8)</td>
<td>48 (33.1)</td>
<td>0.98 (0.55-1.74)</td>
</tr>
<tr>
<td>Rarely</td>
<td>16 (45.7)</td>
<td>10 (28.6)</td>
<td>9 (25.7)</td>
<td>0.73 (0.29-1.85)</td>
</tr>
<tr>
<td><strong>Alcohol consumption</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(Drinks per week)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>95 (48.7)</td>
<td>35 (18.0)</td>
<td>65 (33.3)</td>
<td>1.00</td>
</tr>
<tr>
<td>1-3</td>
<td>77 (49.0)</td>
<td>31 (19.8)</td>
<td>49 (31.2)</td>
<td>0.93 (0.58-1.50)</td>
</tr>
<tr>
<td>≥4</td>
<td>58 (31.4)</td>
<td>49 (26.5)</td>
<td>78 (42.2)</td>
<td><strong>1.97 (1.24-3.13)</strong></td>
</tr>
<tr>
<td><strong>Average cigarettes smoked/day</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non smoker</td>
<td>152 (48.0)</td>
<td>66 (20.8)</td>
<td>99 (31.2)</td>
<td>1.00</td>
</tr>
<tr>
<td>&lt;1</td>
<td>10 (43.5)</td>
<td>6 (26.1)</td>
<td>7 (30.4)</td>
<td>1.08 (0.40-2.92)</td>
</tr>
<tr>
<td>1-5</td>
<td>34 (37.8)</td>
<td>17 (18.9)</td>
<td>39 (43.3)</td>
<td><strong>1.76 (1.04-2.98)</strong></td>
</tr>
<tr>
<td>&gt;5</td>
<td>34 (37.8)</td>
<td>26 (24.3)</td>
<td>47 (43.9)</td>
<td><strong>2.12 (1.28-3.53)</strong></td>
</tr>
<tr>
<td>Any</td>
<td>78 (35.5)</td>
<td>49 (22.3)</td>
<td>93 (42.3)</td>
<td><strong>1.83 (1.24-2.71)</strong></td>
</tr>
<tr>
<td>Variable</td>
<td>HPV-negative n=230 (42.8)</td>
<td>Monoinfection n=115 (21.4)</td>
<td>Coinfection n=192 (35.8)</td>
<td>Coinfection v. HPV-negative</td>
</tr>
<tr>
<td>---------------------</td>
<td>---------------------------</td>
<td>-----------------------------</td>
<td>--------------------------</td>
<td>-----------------------------</td>
</tr>
<tr>
<td>HLA-DQB1*03</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>70 (37.0)</td>
<td>44 (23.3)</td>
<td>75 (39.7)</td>
<td>1.00</td>
</tr>
<tr>
<td>Yes</td>
<td>160 (46.0)</td>
<td>71 (20.4)</td>
<td>117 (33.6)</td>
<td>0.68 (0.46-1.02)</td>
</tr>
<tr>
<td>HLA-DQB1*06:02</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>187 (44.3)</td>
<td>92 (21.8)</td>
<td>143 (33.9)</td>
<td>1.00</td>
</tr>
<tr>
<td>Yes</td>
<td>43 (37.4)</td>
<td>23 (20.0)</td>
<td>49 (42.6)</td>
<td>1.49 (0.94-2.37)</td>
</tr>
<tr>
<td>HLA-DRB1*15:01</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>176 (44.2)</td>
<td>87 (21.9)</td>
<td>135 (33.9)</td>
<td>1.00</td>
</tr>
<tr>
<td>Yes</td>
<td>54 (38.9)</td>
<td>28 (20.1)</td>
<td>57 (41.0)</td>
<td>1.38 (0.89-2.13)</td>
</tr>
<tr>
<td>HLA-G*01:01:01</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>71 (41.3)</td>
<td>46 (26.7)</td>
<td>55 (32.0)</td>
<td>1.00</td>
</tr>
<tr>
<td>Yes</td>
<td>159 (43.6)</td>
<td>69 (18.9)</td>
<td>137 (37.5)</td>
<td>1.11 (0.73-1.69)</td>
</tr>
<tr>
<td>HLA-G*01:01:03</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>205 (43.8)</td>
<td>94 (20.1)</td>
<td>169 (36.1)</td>
<td>1.00</td>
</tr>
<tr>
<td>Yes</td>
<td>25 (36.2)</td>
<td>21 (30.4)</td>
<td>23 (33.3)</td>
<td>1.12 (0.61-2.04)</td>
</tr>
<tr>
<td>HLA-G*01:01:05</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>216 (42.9)</td>
<td>110 (21.9)</td>
<td>177 (35.2)</td>
<td>1.00</td>
</tr>
<tr>
<td>Yes</td>
<td>14 (41.2)</td>
<td>5 (14.7)</td>
<td>15 (44.1)</td>
<td>1.31 (0.62-2.78)</td>
</tr>
<tr>
<td>HLA-G*01:01:07</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>221 (42.9)</td>
<td>108 (21.0)</td>
<td>186 (36.1)</td>
<td>1.00</td>
</tr>
<tr>
<td>Yes</td>
<td>9 (40.9)</td>
<td>7 (31.8)</td>
<td>6 (27.3)</td>
<td>0.79 (0.28-2.27)</td>
</tr>
</tbody>
</table>

† These categories combined due to small numbers
◊ In the last year prior to enrollment
§ In lifetime
* Includes trichomoniasis, chlamydia, herpes, syphilis, gonorrhea, genital sores
♦ In the past 5 years
4.2.2 Cumulative coinfection status: multivariate results

Table 4.9 shows the multivariate associations for all variables retained in the final regression models, mutually adjusted for one another. In the first model, compared to women who remained HPV-negative, significant baseline predictors of developing a cumulative coinfection during the study included having \( \geq 2 \) lifetime sex partners (OR=3.62; 95% CI: 1.88-6.95 for \( \geq 5 \) partners), having \( \geq 2 \) new sex partners in the year prior to enrollment (OR=3.87; 95% CI: 2.16-6.93), having a previous history of STIs (OR=2.96; 95% CI: 1.31-6.69) and having the HLA-DQB1*06:02 allele (OR=1.83; 95% CI: 1.08-3.09). Being age 27 or older at enrollment was also significantly protective (OR=0.34; 95% CI: 0.17-0.67) against developing a cumulative coinfection over follow-up in this model. In the second model, in which all women were HPV-positive, only 3 predictors were significantly associated with development of cumulative coinfections over follow-up: having \( \geq 2 \) new sex partners in the last year (OR=2.74; 95% CI: 1.48-5.08) and having the HLA-G*01:01:03 (OR=0.34; 95% CI: 0.14-0.81) and HLA-G*01:01:05 (OR=6.44; 95% CI: 1.66-24.95) alleles. Though a very strong association was found for the latter allele in particular, it should be noted that this association was based on relatively small frequencies of women (Table 4.8), as evidenced by the width of the confidence interval. According to the Hosmer-Lemeshow goodness-of-fit tests, model fit was adequate for both models (p=0.9552 and p=0.3387, respectively, each with 7 degrees of freedom).
Table 4.9
Multivariate ORs* for associations between baseline exposures and cumulative coinfection status

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coinfection v. HPV-negative (n=422)</th>
<th>Coinfection v. monoinfection (n=307)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>OR (95% CI)</td>
<td>OR (95% CI)</td>
</tr>
<tr>
<td><strong>Age at enrollment</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>≤26</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>≥27</td>
<td>0.34 (0.17-0.67)</td>
<td>0.54 (0.25-1.17)</td>
</tr>
<tr>
<td><strong>Number of lifetime sex partners</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0-1</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>2-4</td>
<td>2.15 (1.14-4.08)</td>
<td>1.19 (0.54-2.65)</td>
</tr>
<tr>
<td>≥5</td>
<td>3.62 (1.88-6.95)</td>
<td>1.22 (0.56-2.66)</td>
</tr>
<tr>
<td><strong>Number of new sex partners◊</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>1</td>
<td>1.33 (0.80-2.21)</td>
<td>1.57 (0.85-2.89)</td>
</tr>
<tr>
<td>≥2</td>
<td>3.87 (2.16-6.93)</td>
<td>2.74 (1.48-5.08)</td>
</tr>
<tr>
<td><strong>History of STIs§</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>Yes</td>
<td>2.96 (1.31-6.69)</td>
<td>1.06 (0.50-2.24)</td>
</tr>
<tr>
<td><strong>HLA-DQB1*06:02</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>Yes</td>
<td>1.83 (1.08-3.09)</td>
<td>1.62 (0.90-2.93)</td>
</tr>
<tr>
<td><strong>HLA-G*01:01:03</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>Yes</td>
<td>1.12 (0.47-2.71)</td>
<td>0.34 (0.14-0.81)</td>
</tr>
<tr>
<td><strong>HLA-G*01:01:05</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>Yes</td>
<td>1.06 (0.35-3.21)</td>
<td>6.44 (1.66-24.95)</td>
</tr>
</tbody>
</table>

* Mutually adjusted for all other variables in a column
◊ In the year prior to enrollment
§ In lifetime; includes trichomoniasis, chlamydia, herpes, syphilis, gonorrhea, genital sores
4.2.3 Concurrent coinfection status: multivariate results

In the repeated measures analysis, the 537 subjects from the cumulative analysis generated 2434 visits where incident coinfection status could be evaluated. Of these, there were 746 HPV-positive visits (30.6%) and 266 visits (10.9%) where an incident concurrent coinfection was detected. Table 4.10 presents the results of the multivariate repeated measures models for associations between baseline and time-dependent predictors of coinfections at individual study visits. These results were somewhat different from the cumulative models, though age and number of sexual partners still appeared to be the most important determinants of developing coinfections, particularly when visits with coinfections were contrasted with HPV-negative visits.

In terms of the time-dependent exposures, aspects of sexual behaviour appeared to be the most important factors involved in coinfection occurrence in both models. Where information from the most recent visits was concerned, number of new sex partners, irregular use of oral contraceptives and never using condoms were all significantly associated with developing coinfections in at least one of the models. Number of new sex partners was the only exposure found to be significant in both models; however, in contrast to the cumulative model where an effect was only seen for ≥2 new sex partners, in the repeated measures models, having any new sex partners in the last 12 months was positively associated with developing coinfections. In contrast to the other time-dependent variables, the exposure time-window for the new sex partner variable was extended to 12 months because when we initially looked at number of new sex partners since the last visit (i.e. last 6 months), no associations were observed. Women who reported never using condoms since their previous visit appeared more likely to have coinfections detected in both models, although the associations were only borderline significant. Interestingly, irregular use of oral contraceptives was also positively associated with coinfection occurrence (OR=1.88; 95% CI: 1.08-3.26) but only in comparison to women with monoinfections. Though age at first intercourse and frequency of sex per week were both significant at p<0.1, none of the
confidence intervals were significantly different from unity and no clear trends seemed to be evident for these variables.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coinfection v. HPV-negative (n=1954 visits)</th>
<th>Coinfection v. monoinfection (n=746 visits)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>OR (95% CI)</td>
<td>OR (95% CI)</td>
</tr>
<tr>
<td><strong>Baseline exposures</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age at enrollment</td>
<td></td>
<td></td>
</tr>
<tr>
<td>≤26</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>≥27</td>
<td><strong>0.38 (0.17-0.87)</strong></td>
<td><strong>0.54 (0.24-1.25)</strong></td>
</tr>
<tr>
<td>Age at first intercourse</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;16</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>16-18</td>
<td>1.21 (0.76-1.94)</td>
<td>1.60 (0.97-2.63)</td>
</tr>
<tr>
<td>≥19</td>
<td>1.10 (0.58-2.06)</td>
<td>1.18 (0.61-2.27)</td>
</tr>
<tr>
<td>Number of lifetime sex partners</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0-1</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>2-4</td>
<td><strong>3.13 (1.60-6.14)</strong></td>
<td>1.12 (0.55-2.27)</td>
</tr>
<tr>
<td>≥5</td>
<td><strong>4.90 (2.38-10.10)</strong></td>
<td>1.07 (0.51-2.25)</td>
</tr>
<tr>
<td><strong>Time-dependent exposures</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of new sex partners</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>1</td>
<td>1.64 (1.22-2.20)</td>
<td><strong>1.59 (1.09-1.34)</strong></td>
</tr>
<tr>
<td>≥2</td>
<td><strong>2.07 (1.53-2.81)</strong></td>
<td><strong>1.52 (1.03-2.24)</strong></td>
</tr>
<tr>
<td>Frequency of sex per week</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;1 time</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>1-3 times</td>
<td>1.00 (0.75-1.34)</td>
<td>0.93 (0.62-1.38)</td>
</tr>
<tr>
<td>4-6 times</td>
<td>1.34 (0.95-1.88)</td>
<td>0.85 (0.50-1.44)</td>
</tr>
<tr>
<td>7+ times</td>
<td>0.96 (0.54-1.67)</td>
<td>0.43 (0.15-1.21)</td>
</tr>
<tr>
<td>Oral contraceptive use</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Regularly</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>Sometimes</td>
<td>1.26 (0.83-1.90)</td>
<td><strong>1.88 (1.08-3.26)</strong></td>
</tr>
<tr>
<td>Never</td>
<td>0.98 (0.73-1.32)</td>
<td>1.23 (0.83-1.82)</td>
</tr>
<tr>
<td>Condom use</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Regularly</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>Sometimes</td>
<td>1.19 (0.91-1.56)</td>
<td>1.24 (0.83-1.83)</td>
</tr>
<tr>
<td>Never</td>
<td><strong>1.33 (1.01-1.74)</strong></td>
<td>1.56 (1.00-2.44)</td>
</tr>
</tbody>
</table>

* Mutually adjusted for all other variables in a column
◊ Since the last 2 visits (~12 months)
§ Since the last visit (~6 months)
4.3 Objective 3: Clinical implications of coinfections

4.3.1 Infection duration

Mean durations of incident type-specific HPV infections are presented in Table 4.11. In general, coinfection with additional types over follow-up seemed to increase the duration of incident type-specific infections. For example, the mean duration of HPV16 monoinfections was 13.4 months (95% CI: 9.1-17.8), whereas the mean duration of HPV16 infections in the presence of other types was 18.3 months (95% CI: 12.6-24.0). Interestingly, when type-specific infections were grouped by oncogenic risk, coinfections seemed to increase the durations for any HPV type overall and for HR-HPV types, but not for LR-HPV types, a trend that was also evident in the individual type-specific durations (e.g. HR: HPV16, 18, 51, 53 v. LR: HPV6, 84). When infections were grouped in terms of the cumulative number of types acquired, women who had 4+ HPV types detected over follow up had type-specific infections of nearly double the mean duration compared to women who accumulated less than 4 types of HPV. Though we have chosen to present data for incident infections only, when the mean durations of all type-specific infections were examined, the trends were virtually the same as for incident infections only, though as expected the mean durations tended to be somewhat longer and the estimates were slightly more precise due to the inclusion of prevalent infections. In some cases, however, more extreme differences were observed when all infections were included; for example, the mean duration of HPV31 monoinfections was 8.9 months (95% CI: 4.9-12.8) and the mean duration of HPV31 in the presence of other types was virtually double—16.7 months (95% CI: 12.0-21.3) (data not shown).
Table 4.11
Mean duration of type-specific HPV infections based on cumulative HPV infection status and stratified by cumulative number of types

<table>
<thead>
<tr>
<th>HPV type</th>
<th>Number of subjects</th>
<th>Mean duration, months (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mono infection</td>
<td>Coinfection</td>
</tr>
<tr>
<td>6</td>
<td>4</td>
<td>25</td>
</tr>
<tr>
<td>16*</td>
<td>10</td>
<td>52</td>
</tr>
<tr>
<td>18*</td>
<td>5</td>
<td>20</td>
</tr>
<tr>
<td>31*</td>
<td>2</td>
<td>19</td>
</tr>
<tr>
<td>51*</td>
<td>9</td>
<td>36</td>
</tr>
<tr>
<td>53*</td>
<td>3</td>
<td>28</td>
</tr>
<tr>
<td>54</td>
<td>7</td>
<td>27</td>
</tr>
<tr>
<td>84</td>
<td>8</td>
<td>39</td>
</tr>
<tr>
<td>Any HPV(^a)</td>
<td>77</td>
<td>80</td>
</tr>
<tr>
<td>HR-HPV(^b)</td>
<td>39</td>
<td>73</td>
</tr>
<tr>
<td>LR-HPV(^c)</td>
<td>38</td>
<td>64</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Number of types</th>
<th>Number of subjects</th>
<th>Mean duration, months (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 type</td>
<td>77</td>
<td>13.7 (11.3-16.0)</td>
</tr>
<tr>
<td>2-3 types</td>
<td>58</td>
<td>12.8 (11.0-14.6)</td>
</tr>
<tr>
<td>4+ types</td>
<td>22</td>
<td>20.3 (14.1-26.5)</td>
</tr>
</tbody>
</table>

* Types considered high-risk or probably high-risk
\(^a\): Refers to the longest type-specific infection with any HPV type in women with incident HPV
\(^b\): Refers to the longest type-specific HR-HPV infection in women with incident HR-HPV
\(^c\): Refers to the longest type-specific LR-HPV infection in women with incident LR-HPV

Figures 4.2 and 4.3 show the time to clearance of the longest type-specific infection based on cumulative HPV infection status for both (A) incident infections only and (B) all infections. In most cases, cumulative coinfection with ≥2 HPV types increased the time to clearance, however, differences were more prominent when all infections were included (Figure 4.2B, log rank test: p=0.002; Figure 4.3B, log rank test: p<0.0001 for 4+ types v. 1 type and p=0.021 for 4+ types v. 2-3 types). Worth noting however, the time to clearance of the longest *incident* type-specific infection was also notably longer among women with 4+ HPV types compared to 2-3 types detected over follow-up (Figure 4.3A, log rank test: p=0.083).
Figure 4.2 – Time to clearance of longest enduring type-specific infection, stratified by coinfection status for (A) incident infections and (B) all infections.
Figure 4.3 - Time to clearance of longest enduring type-specific infection, stratified by cumulative number of types for (A) incident infections and (B) all infections
4.3.2 Cervical lesions

Cytological outcomes were relatively uncommon in this group of young women (Table 4.12). The most prevalent lesion outcome was LSIL, which was the worst cytological outcome in 42 women (6.8%). When each visit was considered individually, there were slightly more lesion outcomes because some women had multiple outcomes detected or had more than one visit with the same lesion outcome. In general, women with SIL (LSIL or HSIL) had a greater proportion of lifetime sex partners and a greater occurrence of HPV infections compared to women with normal cytology (12). As well, HPV positivity increased in accordance with the severity of lesion grades, such that all four women with HSIL had concurrent coinfections detected (12).

Table 4.12

<table>
<thead>
<tr>
<th>Lesion Grade</th>
<th>Each woman (n=621)</th>
<th>At each visit (n=2675)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal</td>
<td>568 (91.5)</td>
<td>2614 (97.7)</td>
</tr>
<tr>
<td>ASCUS</td>
<td>7 (1.1)</td>
<td>8 (0.3)</td>
</tr>
<tr>
<td>LSIL</td>
<td>42 (6.8)</td>
<td>49 (1.8)</td>
</tr>
<tr>
<td>HSIL</td>
<td>4 (0.6)</td>
<td>4 (0.2)</td>
</tr>
</tbody>
</table>

* Among women with more than one lesion outcome, refers to the worst grade lesion detected

Table 4.13 shows the adjusted ORs and 95% CIs for associations between different types of coinfections and detection of SIL over follow-up (cumulative) and at each visit (concurrent). As expected, the strongest associations were observed for women with coinfections compared to HPV-negative women: in particular, women with concurrent coinfections had a 25-fold increase in odds of SIL (OR=25.08; 95% CI: 10.38-60.60), compared to HPV-negative women at the same visit. Women with a coinfection over follow-up also appeared more likely to have SIL detected than women with a monoinfection over follow-up, however the association was not statistically significant (OR=1.61; 95% CI: 0.80-3.26).
Table 4.13
Associations between HPV coinfection status and SIL cumulatively over follow-up and concurrently at the same visit

<table>
<thead>
<tr>
<th>HPV infection status</th>
<th>Cumulative detection</th>
<th>Concurrent detection</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Adjusted* OR of SIL</td>
<td>Adjusted* OR of SIL</td>
</tr>
<tr>
<td>HPV-negative Coinfection</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td>17.50 (5.21-58.78)</td>
<td>25.08 (10.38-60.60)</td>
</tr>
<tr>
<td></td>
<td>(n=495)</td>
<td>(n= 2180)</td>
</tr>
<tr>
<td>Monoinfection Coinfection</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td>1.61 (0.80-3.26)</td>
<td>2.04 (1.08-3.85)</td>
</tr>
<tr>
<td></td>
<td>(n=328)</td>
<td>(n=837)</td>
</tr>
</tbody>
</table>

* Adjusted for age and race

In contrast, when we stratified HPV status by number of types, the results were even more pronounced (Table 4.14). In all models, associations were strongest for the detection of 4+ HPV types (v. HPV-negative) and a clear dose-response pattern was evident. Again, the strongest associations were seen for concurrent detection of multiple types and SIL at the same study visit. When we investigated the temporal relationship between concurrent coinfections and detection of lesions at the next visit, the associations were substantially attenuated compared to the other 2 models and the linear trend was no longer significant (p=0.1210). Though this result is plausible given the natural history of HPV, it should be noted that this model had considerably less power than the other models because a substantial number of lesions (>30%) were lost when we excluded prevalent lesion outcomes.
### Table 4.14
Associations between number of HPV types and SIL at different time points

<table>
<thead>
<tr>
<th>Number of HPV types</th>
<th>Cumulative detection (n=621) Adjusted* OR of SIL</th>
<th>Concurrent detection; lesion at same visit (n=2675) Adjusted* OR of SIL</th>
<th>Concurrent detection; lesion at next visit (n=2054(a)) Adjusted* OR of SIL</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>1 type</td>
<td>11.29 (3.15-40.46)</td>
<td>14.10 (5.37-37.02)</td>
<td>2.76 (1.23-6.18)</td>
</tr>
<tr>
<td>2-3 types</td>
<td>15.65 (4.49-54.57)</td>
<td>24.87 (10.20-60.65)</td>
<td>4.86 (2.14-11.07)</td>
</tr>
<tr>
<td>4+ types</td>
<td>26.44 (7.05-99.09)</td>
<td>52.19 (12.76-213.32)</td>
<td>8.82 (1.18-65.80)</td>
</tr>
<tr>
<td>P trend</td>
<td>&lt; .0001</td>
<td>&lt; .0001</td>
<td>0.1210</td>
</tr>
</tbody>
</table>

*a*: The first visit was excluded from this model because coinfection status was unknown prior to the baseline visit
*Adjusted for age and race

### 4.4 Brief summary

In summary, we observed that HPV coinfections occurred in a substantial proportion of the cohort, with significantly different rates of acquisition based on infection status at enrollment. In our second objective, where we investigated the risk factors underlying these differences, several significant predictors of coinfections were identified, mainly relating to aspects of sexual activity. In addition, however, we also observed significant associations for several HLA alleles, possibly indicating the presence of differential immune susceptibility. Finally, we observed that women with coinfections had both longer infection durations and greater occurrence of cervical lesions, with an apparent threshold of effect such that women with 4+ HPV types were at highest risk of both outcomes. The next chapter will elaborate on the meaning and implications of these findings.
Chapter 5

Discussion

To our knowledge, this is the first study to examine the occurrence, determinants and
dynamics of HPV coinfections in a longitudinal cohort of young women using multiple
coinfection definitions. It is also the first study in a Canadian population to explicitly examine
multiple-type HPV infections. The aims of this project were threefold: first, to document the
number of women who acquire coinfections over time and the multiplicity of HPV types they
acquire; second, to examine both behavioural and biological predictors of coinfection occurrence;
and third, to explore the clinical implications of coinfections in the continuum of cervical
carcinogenesis. This chapter will summarize and interpret the major findings pertaining to each of
these objectives as well as discuss the strengths, limitations and significance of this project.

5.1 Summary of results and interpretations of findings

5.1.1 Prevalence and incidence of coinfections

As expected, coinfections were very common in the cohort, occurring in approximately
one-third of all women and in the majority of HPV-positive women. What was somewhat
surprising, however, was the fact that nearly all of the coinfections overlapped at one or more
study visits (i.e. were concurrent), as we had initially thought that there would be many more
cumulative coinfections than concurrent, as the former was the more liberal coinfection
definition. This finding likely reflects two facts about coinfection dynamics in the cohort: firstly,
that women largely acquired their coinfections at approximately the same time or within short
duration of each other, and secondly, that infections tended to persist over several study visits,
such that they overlapped with newly acquired types. It was also interesting that there was such a
high multiplicity of HPV types among individual women. For example, in the cumulative assessment, a larger proportion of women acquired 3+ types (19.1%) than acquired 2 types (13.4%) over the course of the study. Put another way, nearly 60% of women with a cumulative coinfection acquired 3+ types over time.

In general, these prevalence estimates are comparable to previous estimates from the literature (106, 107), though they are higher than several other estimates that have been reported (108-110). This difference is likely owing to the young age of this study population, as well as the longitudinal nature of the study, a design which allowed us to calculate estimates of period prevalence rather than point prevalence. Many other estimates in the literature have come from cross-sectional studies, which only report a “snapshot” of coinfection prevalence at a single point in time. Our study improves upon this by presenting the cumulative prevalence of coinfections over 2 to 3 years of follow-up, presenting a much richer picture of the true proportion of coinfections that occur among young, sexually active women. Highlighting the differences between these two estimates, only 11.4% of the cohort had a coinfection cross-sectionally at enrollment, whereas this proportion increased 3-fold to 32.5% over follow-up.

In our assessment of cumulative type-specific prevalence, it was noteworthy that the types with the largest proportion of monoinfections were nearly all oncogenic types, though this finding may have been an artifact of the method used to detect HPV DNA. For example, it is possible that any coexisting LR-HPV types may have been systematically under-detected, since the PCR protocol that was used in this study has been shown to favour the detection of HR-HPV (64) and HR-HPV types are typically present at higher viral loads. Similar to previous reports (78), HPV16 was the most frequently detected type in this cohort, whereas other types covered by current prophylactic vaccines (HPV6, 11, and 18) were considerably less common.
In terms of incidence, we consistently observed that HPV-infected women acquired additional types at a greater rate and in significantly less time than HPV-negative women when we stratified on infection status at enrollment. Most strikingly, women with coinfections at enrollment seemed to acquire new types of HPV at the highest rate of all. Though these results are consistent with the trends present in the literature, there are few studies with estimates of coinfection incidence available for comparison. One study which reported the incidence of concurrent coinfections in an older population (mean age = 33y) of Brazilian women (105), found that the cumulative probability of concurrent coinfection among HPV-negative women was 3% (95% CI: 2-4) at 2 years, whereas in this cohort the corresponding estimate was 16% (95% CI: 15-17)—more than 5 times greater. Though one explanation for this discrepancy may be differences in age distribution, our IR of 7.2 per 1000 woman-months was still nearly two times higher than the IR for women aged 18-24 in the Brazilian cohort. As such, another explanation for these differences may be the higher numbers of sex partners had by women in this study population compared to the Brazilian cohort, in which 44% of women had only one lifetime sex partner. Another study, which looked at the cumulative incidence of repeated HPV infections after an initial infection (11), found that 63% of women with a sequential coinfection acquired a third HPV infection within 3 years. Though the shorter duration of our study did not permit 3 years of follow-up in all women, in less than 3 years, up to 80% of women with a coinfection at baseline acquired a new HPV infection—considerably more than proportions reported elsewhere.

5.1.2 Risk factors for coinfection occurrence

In our second objective we investigated predictors of HPV coinfection occurrence. In order to gain the greatest amount of information from this analysis we used two different comparison groups and two definitions of coinfections, and examined both baseline and time-varying risk factors. In these analyses, HPV status was defined in one of three ways: HPV-
negative, HPV monoinfection or HPV coinfection, cumulatively over the course of the study and concurrently at each visit, depending on the outcome. Of our two comparisons, we were most interested in the coinfection v. monoinfection models as it was hypothesized that they might help elucidate one of the important questions in coinfection research: namely, why some HPV-positive women acquire coinfections and others do not. Several studies (112, 120) have reported increased sexual activity to be the overwhelming cause of coinfection occurrence, but there has been little research into immunological factors that may be involved in modulating individual susceptibilities. By including several exposures purported to affect the immune response (i.e. HLA alleles, smoking status, vegetable consumption), we hoped to tease out some of these more subtle predictors of coinfection occurrence. By contrast, in the models that included HPV-negative women, we expected the most important predictors of coinfections to be predominantly sexual in nature, and so these models served as a good comparison to the exclusively HPV-positive models, in which we hypothesized that sexual behaviour(s) might be less important.

As expected, in the final multivariate models examining baseline predictors of cumulative coinfections, we found several differences in the two comparison models, and each model appeared to reveal slightly different aspects of coinfection etiology. As hypothesized, the coinfection v. HPV-negative model mainly reflected traditional risk factors of HPV infection (e.g. age) as well as markers of increased sexual activity (e.g. more lifetime sex partners, history of STIs). Number of new sex partners was the only significant predictor of coinfections that appeared in both models, which is likely explained by the fact that it is a proxy for exposure to new types of HPV. Notably, three HLA alleles were retained as predictors in the final models, and as hypothesized, seemed to be particularly associated with coinfections among HPV-positive women. In the repeated measures models examining predictors of incident concurrent coinfections at individual study visits, the results were generally similar to the cumulative models, though we had the added benefit of assessing predictors that could vary over time. As
such, we identified several significant time-varying predictors of concurrent coinfections, namely, new sex partners, recent condom use and recent oral contraceptive use.

As age is one of the most consistent risk factors for HPV infection, it was unsurprising that it was an important predictor of coinfections in the models that contained HPV-negative women. In both of these models, women aged 27 or older had significantly decreased odds of developing a coinfection in comparison to younger women, an effect that is most likely explained by an age-related increase in HPV-acquired immunity. Age was not a significant predictor of coinfections in either of the HPV-positive models, however, in contrast to the findings of several other studies (11, 104, 122). One interpretation for this difference could be the lack of age variation found in this cohort, as >90% of the women in our study were less than 30 years old. In support of this explanation, all of the studies mentioned above were conducted in cohorts with diverse age distributions, whereas a similar study conducted among female university students did not find an association between age and coinfections among HPV-positive women (112).

Number of sex partners, perhaps the best marker of HPV exposure, appeared to be the most important risk factor for coinfections in all of the models. As expected, women who had the highest number of sex partners (both lifetime and new) had the greatest odds of coinfections, though number of lifetime partners was only a significant predictor in comparison to HPV-negative women. This finding is likely explained by the fact that this model contained women with the most extreme differences in terms of sexual activity, whereas the exclusively HPV-positive model contained women that were much more similar in this regard. This interpretation probably also explains the finding for history of STIs, as this is another marker of increased sexual activity. Though it is believed that concurrent STIs may facilitate HPV susceptibility by increasing irritation and/or inflammation in the cervix (43), it is difficult to draw this conclusion based on these findings since were not able to assess when STIs prior to baseline actually
occurred. Although a previous study in this cohort (12) found a positive association between recent *Chlamydia* infections and HPV acquisition, we did not observe any effect(s) with recent STIs in the repeated measures models.

As the most likely source of new HPV types, number of new sex partners has been consistently associated with coinfections in several other longitudinal studies (11, 104, 112), so it was not surprising that this was also an important risk factor in this cohort. Moreover, the fact that it was the most consistent predictor of coinfections among HPV-positive women strongly suggests that coinfection occurrence is driven by increased sexual activity. Notably, however, we found that slightly different quantities of new sex partners were associated with increased odds of coinfections depending on which coinfection outcome was assessed. For example, only 2 or more new sex partners in the year prior to baseline increased the odds of cumulative coinfections, whereas any new sex partner in the last 12 months was associated with increased odds of incident concurrent coinfections. This difference, though slight, is interesting because it reflects the subtleties in risk that may be gained by using more recent exposure information and more accurate outcome assessment. Though both exposures reflected the number of new sex partners in the previous 12 months, the concurrent model enabled a more precise assessment of when coinfections actually occurred.

Furthermore, our examination of new sex partners in the repeated measures models also revealed important information about the time it takes for newly acquired HPV infections to replicate to detectable levels. In the original conceptualization of the new sex partner variable, we assessed the number of new sex partners reported since the last visit (i.e. last 6 months) but found that the variable was not associated with coinfections at all. However, when the exposure time-window was extended to include the interval incorporating the last 2 visits (i.e. last 12 months), the new sex partner variable became highly significant. As it has been previously suggested that it
may take up to 8 months for a newly transmitted HPV infection to become detectable (11), it is worth noting that these results suggest that it may have taken up to 12 months in our study, though we had less frequent HPV testing intervals.

In addition to age and number of sex partners, we identified two significant time-dependent predictors in the repeated measures models: condom use and OC use. Women who reported ‘never’ using condoms since their previous visit appeared to have increased odds of coinfections in both models, whereas irregular (i.e. ‘sometimes’) use of OCs since the last visit was only a significant predictor of coinfections among HPV-positive women. These results are noteworthy because conflicting results have typically been found for these factors in several other studies of coinfections. For example, none of the longitudinal studies of coinfections have identified recent condom use as a significant predictor; however, it was associated with decreased occurrence of coinfections in two cross-sectional studies (70), though one was conducted among men (111). Indeed, though condom use has not been found to have the consistent protective effect with HPV that it has with other STIs (84, 136), it is likely that when properly used condoms convey at least some degree of protection, either by acting as a protective barrier against skin-to-skin transmission of the virus or by decreasing HPV viral load (12).

OC use has also been found to have an equivocal effect on HPV coinfection risk (112, 122); however, part of the challenge in studying this factor is that OC use is strongly associated with sexual activity, and residual confounding may exist even after adjusting for number of sexual partners (96, 137). As such, and because the effect(s) of OC use are largely thought to operate “downstream” of HPV infection (96), the observed association for OC use in this study is most likely explained by the fact that irregular OC users were also more likely to have had more new and/or recent sex partners in this cohort (data not shown).
In terms of immunological factors, in our exploratory examination of HLA alleles, we found three alleles to be significantly associated with coinfections in the cumulative models. It was particularly interesting that as we had hypothesized, two of the three alleles appeared in the HPV-positive model (where HPV infection status was effectively ‘controlled for’), and only after adjustment for number of new sexual partners. As these alleles are involved in cell-mediated antigen presentation, it is conceivable that they could be involved in immune recognition and clearance of HPV among women exposed to the virus. However, as precise biologic mechanisms between HLA molecules and HPV remain undetermined (46), we can only make tentative interpretations of what these associations may represent at the molecular level. For example, based on current understanding, the positive associations we observed may be evidence of inefficient recognition between HPV antigen and the HLA molecules produced by these alleles, such that there is increased immunological susceptibility to the virus. Additionally, as genes rarely operate in isolation, we cannot rule out the possibility that the observed effects could also be due—at least in part—to other genetic loci of immune recognition that were not assessed in this study.

As this is the first study to explore the role of HLA alleles in the development of HPV coinfections, it is somewhat difficult to contextualize these results with other studies in the literature. Nevertheless, the HLA-DQB1*06:02 allele has been fairly consistently linked to both cervical cancer and malignant precursor lesions in several studies (3, 46, 138), and in a previous study in this cohort (47), this allele was associated with both HR-HPV and HPV16 positivity, though the associations did not reach statistical significance. Less evidence is available for the HLA-G alleles, as they have been less frequently studied to date. Somewhat paradoxically, another previous study in this cohort using a different analysis approach found both the HLA-G*01:01:03 and HLA-G*01:01:05 alleles to be associated with increased odds of certain LR-HPV infections, but not among women purported to be highly exposed to HPV (95). Another
recent study in a Brazilian population (48) found the HLA-G*01:03 allele to be protective against both HPV infection and SIL but they did not observe any effect for the HLA-G*01:01:03/05 alleles that we found in this study. As such, though our HLA findings are biologically plausible, due to their highly exploratory nature and the fact that they are based on small numbers of women, we advocate that these results be interpreted with caution until they can be confirmed in additional studies.

5.1.3 Coinfections, infection duration and lesions

Infection duration

When we compared the mean durations of type-specific infections in isolation and in the presence of other HPV types (coinfection) we typically observed that coinfections were associated with longer infection durations. Interestingly, this trend was observed primarily with HR-HPV types (both grouped and individually), as opposed to LR-HPV types. Although we were not able to look at all HPV types individually due to the small frequencies that resulted from multiple stratifications (i.e. incident infections only, coinfections v. monoinfections), we were able to stratify mean durations on the cumulative number of HPV types acquired, something novel that no other study has done to date. In this analysis, we found that women who acquired the most HPV types (4+) had much longer infection durations than women with fewer types, though the confidence intervals slightly overlapped due to small frequencies in the ‘4+’ category. Women who cumulatively acquired 4+ types also took substantially longer to clear their type-specific infections in the Kaplan-Meier time to clearance analysis for both all infections and incident infections only.

To our knowledge, there is only one other study (10) that has examined the duration of type-specific HPV infections stratified on the presence of coinfections. Despite having much longer follow-up than this cohort, that study, conducted among a large cohort of Brazilian
women, found similar results to what we observed including the general trend of coinfections increasing the duration of HR-HPV infections to a greater degree than LR-HPV infections. Reasons for this difference are largely unclear, but it emphasizes the potential clinical significance of coinfections and reinforces the idea that coinfections may be a marker of pathogenicity or of susceptibility to HPV.

Other studies that have investigated coinfections and persistence have typically not been type-specific, or have not looked at exclusively incident infections, which may explain their conflicting results (75, 77, 107, 114). Indeed, two strengths of our analysis were our focus on type-specific infections and on incident infections as a separate group, as failing to do this can have detrimental effects on a study’s validity. For example, when studying coinfections, not restricting to type-specific durations will result in overestimates of infection duration, as women who have multiple overlapping infections will naturally have a longer duration of HPV positivity than women with only one HPV infection. Similarly, the inclusion of prevalent infections will tend to over-represent more persistent HPV infections as well as preclude the definition of the actual boundaries of duration episodes, thereby biasing estimates of ‘true’ infection duration. Though we did look at durations of all HPV infections (i.e. including prevalent) in our analysis, we presented data for incident infections only wherever possible, as these are the more accurate estimates. It was particularly important that we take precautions against overestimation in our study, as the relatively short follow-up period resulted in a high number of censored observations in the clearance analysis, which may have slightly inflated our duration estimates.

Nevertheless, this analysis provides important contributions to understanding the relationship between coinfections and persistence, by showing that young women who acquire more HPV types are also more likely to have longer infection durations. These results are interesting because they suggest the existence of a possible ‘risk threshold’ among women with
coinfections, whereby women who acquire 4+ HPV types are at particularly increased risk for HPV persistence. This finding makes sense in the context of the ‘immune dysfunction hypothesis,’ which proposes that women whose immune systems respond poorly to HPV are more likely to have both coinfections and infection persistence (11). Importantly, this analysis expands on this hypothesis by potentially identifying a subset of women among those with coinfections who may indeed have greater immunological susceptibility to HPV.

Cervical lesions

As mentioned previously, our investigation of coinfections and cervical lesions was fairly exploratory, as it was known that very few lesions had occurred in the cohort and that the majority were LSILs, which are often reflective of productive HPV infections rather than true oncogenic progression. Nevertheless, despite the relatively low prevalence of lesions in this population, we observed highly significant associations between HPV coinfections and development of SIL using several complementary analytic approaches. In all analyses, the strongest associations were observed for concurrent coinfections and SIL in the same study visit whether we considered all coinfections together or stratified them based on the multiplicity of coinfecting types. Unsurprisingly, associations were also strongest when women with coinfections were compared to HPV-negative women; however, a significantly increased risk, albeit greatly attenuated, was still observed for concurrent coinfections and SIL in comparison to women with HPV monoinfections (OR=2.04; 95% CI: 1.08-3.85). In all of the stratified models, women with 4+ types of HPV (cumulatively or concurrently) had greatly increased odds of SIL, up to double or triple the odds of women with 2-3 types and 1 type, respectively. However, when we looked at the occurrence of SIL approximately 6 months after the detection of a concurrent coinfection, the associations decreased markedly in comparison to when the exposure (number of concurrent types) and outcome (SIL) were assessed at the same visit.
Many of these results can be explained by the fact that virtually all of the SILs in the cohort were LSILs, a fairly typical finding in studies of very young women as it often takes several years after HPV infection for HSILs to develop (2, 103). For example, that the strongest associations were observed for concurrent coinfections and SIL in the same visit likely represents the fact that the majority of the lesions in the cohort were caused by the cytopathic effects of productive viral infections, as opposed to actual oncogenic progression. This interpretation is strengthened by the fact that associations between coinfections and SIL decreased substantially when a 6-month time lag was introduced, indicating that many of the LSILs regressed as women cleared their HPV infections over time.

Several aspects of our analysis were similar to a previous study (123) conducted in the same Brazilian cohort mentioned earlier, which had an older study population, a longer follow-up period and a much higher prevalence of SIL (particularly HSIL) than the Montreal cohort. Despite these important differences, it is noteworthy that we found virtually the same trends in our cohort as were found in that study, although our effect estimates were not as strong. In particular, women with 4+ types of HPV also had greatly increased odds of SIL in the Brazilian study, even when HPV16 infections were excluded. Unlike our study, however, in the Brazilian cohort, women with 4+ HPV types had substantially increased odds of SIL six months later, particularly where HSIL was concerned (OR>1000). Though our cohort had a higher prevalence of coinfections than the Brazilian cohort, our low rate of lesion occurrence and virtual lack of HSILs precluded our ability to make similar stratifications in our dataset.

Despite these limitations, however, our analysis is still informative. It suggests that even among young women early on in the natural history of HPV, HPV coinfections are associated with the detection and/or development of cervical lesions to a greater degree than HPV monoinfections alone. It also makes evident that there is a clear dose-response corresponding to
the number of HPV types detected, such that women with 4+ types have much greater odds of lesion detection than women with less than 4 types. Taken together with the findings from our duration analysis, these results strongly suggest that women who acquire 4+ types of HPV may represent a subset of women with increased HPV susceptibility and greater risk of further HPV-related disease. Though there is some evidence to suggest that this effect also exists in other study populations and for high-grade lesions (HSIL) (123), further investigation of this effect should be the focus of additional studies in larger cohorts of women.

5.2 Methodological issues and limitations

As with all research, this study had several limitations. This section will discuss the most important methodological issues and attempt to assess their impact(s), if any, on our findings.

5.2.1 Loss to follow up

In prospective cohort studies, loss to follow up is always one of the primary concerns. Though this study was of relatively short duration, it required regular follow-up visits of a somewhat invasive nature, which may explain why approximately one-third of the cohort was eventually lost to follow up. If women with a certain exposure profile were less likely to remain in the study and they were also more (or less) likely to acquire coinfections then it is possible that some of our associations would be invalid, and it would be difficult to predict in which direction the estimated odds ratios would be biased.

However, as none of the participants were informed of their HPV status until the end of the study, it is unlikely that loss to follow up would have been influenced by our main study outcome. In support of this contention, when the incidence of coinfections was stratified by the number of visits completed, no significant differences between women who were lost to follow
up and women who completed the full duration of the study were observed (Tables 4.6 and 4.7). Similarly, when the distribution of exposure variables measured at baseline was compared between women lost to follow up and women that completed all five study visits, no substantial differences were observed (Appendix A) (12). As such, we can be reasonably confident that loss to follow up in the cohort was not related to either the exposures or outcomes of interest.

5.2.2 Measurement error and misclassification

5.2.2.1 HPV status

In terms of measurement, one of the most important limitations was the fact that our ability to measure HPV coinfections was dependent on the detection of multiple HPV genotypes in cervical cell specimens. As previously mentioned, the MY09/11 protocol that was used for the PCR amplification in this study may not be as sensitive for the detection of multiple HPV types as other primers that have since become available (e.g. PGMY09/11) (66, 67). As this was particularly a concern with concurrent coinfections, where the aim was to detect multiple types in the same cervical specimen, one of the ways we hoped to minimize this limitation was through the examination of both cumulative and concurrent coinfections, since the former should have been less prone to this type of outcome misclassification. Nevertheless, if such misclassification did occur, it would have led to a systematic underestimation of coinfection detection, because women with multiple types may have been classified as having fewer types than they actually did. Similarly, it is also important to acknowledge that any instance(s) of negative results for a given HPV type may have been a consequence of low viral load such that it was below the threshold of detection by the assay, or because of insufficient sampling of cervical cells in the specimen. Though this could have decreased our occurrence estimates, there is no reason to believe that the sensitivity of HPV detection would have differed based on the distribution of certain exposure variables, therefore, any misclassification would most likely have been non-differential, and in
the direction of the null. Moreover, the fact that our estimates of coinfection occurrence were comparable (and often greater) than others in the literature, suggests that coinfection under-detection was not a major problem in this study.

5.2.2.2 Other variables

As the HLA alleles in this study were genotyped using highly sensitive and accurate laboratory techniques, other sources of misclassification would most likely have come from the questionnaire data. First, due to aspects of social desirability, there is the possibility that women may have under-reported certain risk behaviours, especially as the questionnaire dealt with several private, sensitive subjects (e.g. history of STIs, number of sexual partners, number of drinks per week). The fact that the questionnaire was self-administered and filled out privately should have helped minimize this effect, but if it did occur, there is no reason to believe that such under-reporting would have happened to a greater extent in women with or without coinfections since subjects were not aware of their HPV status. Similarly, it is possible that misclassification could also have been introduced among women with missing data, if the values that were imputed (i.e. values from previous visits) were not correct, and women had, in fact, changed their behaviour(s). Attempts were made to minimize this effect by excluding the few women who had large numbers of missing items, but again, if it did occur, it is unlikely that women who failed to respond to certain questions were also systematically more likely to have had coinfections detected, and thus any misclassification would also most likely have been non-differential. As such, it may be important to interpret our null associations in objective 2 with due caution.

5.2.3 Generalizability

Another important limitation may be the generalizability of these findings. Because this cohort represents a convenient sample of women who attended university health clinics and agreed to participate in this study, it may not be representative of female students at other
universities or even of all female students at McGill or Concordia at the time the study was conducted. As such, there is a possibility that women who agreed to participate may have been systematically different from those women who chose not to participate in the study. Though we were able to compare rates of cervical abnormalities among women who participated and women who refused (and found them to be very similar), we were not able to obtain information on risk behaviours or lifestyle characteristics among women who refused to participate. Though this potential lack of generalizability may have resulted in biased estimates of coinfection occurrence (e.g. if women in the study were more/less sexually active), it would likely not affect the internal validity of the study or the effect estimates, particularly as most associations involved biological mechanisms that should not vary between populations. In fact, it is worth noting that virtually all of the current body of knowledge on the epidemiology of HPV has been derived from non-population-based studies similar to this one (12). Nevertheless, when the original study was conducted, attempts were made to assess the representativeness of the cohort by comparing various study characteristics with two other surveys conducted among university students around the same time, with few to no substantial differences observed (Appendix B) (12).

Also worthy of consideration, however, is the time that has passed since this study was conducted. Because the majority of data collection occurred over 10 years ago, it is possible that our estimates of coinfection prevalence and incidence may not be generalizable to populations of female university students at the present time. For example, one of the most major changes to have occurred in recent years has been the introduction of the aforementioned prophylactic HPV vaccines, the first of which was approved for use in Canada in 2006 (139). However, despite the availability of these vaccines, current vaccine coverage among female university students is not likely to be extremely high, as women in this age group are not covered under publicly-funded HPV vaccination programs and must therefore pay for the vaccines if they wish to be vaccinated. In fact, this assertion is supported by data from a more recent study of HPV infection conducted
among a similar target population in Montreal, which found that only 11% of women had been vaccinated and that vaccine-covered types were still common in the study population (82). As HPV vaccination continues to become more common, however, and young girls vaccinated through school-based programs reach university age, it is likely that our coinfection estimates will become less and less generalizable, as all four vaccine-covered types were frequently found as coinfections in this study (Table 4.3).

5.3 Strengths and significance

Despite these limitations, this study also has several important strengths. Perhaps most importantly, we were able to use a longitudinal repeated measures design to study coinfections, which enabled us to examine several important aspects of coinfection dynamics (e.g. incidence, duration) that have been rarely reported in previous studies. The repeated measures design also allowed the assessment of determinants of coinfections that vary over time, which may better represent temporal etiologic associations. Additionally, having longitudinal data permitted us to report cumulative prevalence estimates, which are less likely to result in misclassification of coinfection status by capturing women who acquire coinfections over time. Another strength of this study was the relatively young age of the study cohort, as younger populations are both more efficient and most relevant for the study of HPV coinfections, as young women are at greatest risk of coinfection occurrence. The young age of the study population was also a determining factor in our ability to stratify our analyses by number of coinfecting types, as such a high multiplicity of types would be an unlikely occurrence among populations of older women.

This research is also significant for a number of reasons. More generally, it represents both the first study of coinfections in a Canadian population and one of the most comprehensive examinations of coinfection epidemiology among young women in the literature. As well, from a
public health perspective, this study contributes to our understanding of coinfection risk and mechanisms of risk modification. In particular, we observed that women with coinfections are a high-risk group for further HPV acquisition, and we identified several markers of increased sexual activity and poor contraceptive practices as modifiable risk factors of coinfection occurrence. Additionally, these findings are also significant from a clinical standpoint, as we showed that young women with coinfections may be at increased risk of both HPV persistence and cervical lesions, two outcomes on the pathway to oncogenic progression. Importantly, if these results can be replicated and expanded upon in additional studies, there may be a rationale for considering whether testing for coinfections might be a worthy approach of HPV risk management in clinical settings. Finally, this research is significant scientifically, as it has contributed to the ongoing research effort to identify cofactors and/or markers of increased cervical cancer risk. In particular, we have helped advance the theory of coinfections and immune dysfunction, by identifying several possible genetic markers of differential immune susceptibility as well as evidence that may represent a gradient of susceptibility among women with coinfections.

5.4 Conclusions and future directions

In summary, this thesis examined the occurrence, determinants and dynamics of HPV coinfections in a cohort of young women in Montreal, Quebec. Our results showed that in this study population, women acquire coinfections frequently, with many women accumulating three or more HPV types in a relatively short period of time. Moreover, we showed that women with coinfections continue to acquire additional HPV types at a greater rate than women without coinfections, and tend to have increased persistence and lesion occurrence relative to women with HPV monoinfections. As hypothesized, our determinants analysis revealed that coinfections are
caused by a complex interplay of factors affecting both exposure and susceptibility to HPV. Notably, we observed several HLA alleles that were associated with coinfection occurrence, even after controlling for more traditional risk factors such as age and sexual activity. Taken together, these results suggest that among young women with coinfections, there is a subset of women who likely have increased immunological susceptibility to HPV. Most intriguingly, we appeared to identify a threshold of risk among women who accumulated the most HPV types; however, this effect needs further study in other populations before any definitive conclusions can be made.

Given these findings, there are several important directions for further research. First, there is a need to investigate whether comparable trends in coinfection dynamics exist in other cohorts of women, including those with more diverse age distributions as risk of HPV persistence and lesion development increases over the life course. Similarly, there is a need for further research into immunological biomarkers of HPV susceptibility, particularly among women who acquire many types of HPV over short periods of time, as these women may represent an important risk subgroup. As it is only a small proportion of the vast numbers of women infected with HPV that experience advanced cervical disease, it remains important to identify factors that make this group of women unique. As women with coinfections—or at least a subset of them—likely represent women with some degree of HPV susceptibility, it makes sense to direct our attention towards this group of women, as they may embody a promising lead in the ongoing search for cofactors of HPV-mediated cervical carcinogenesis.
References


Appendix A: Assessment of Differential Loss to Follow up

As can be seen in the table below (adapted from (12) p. 75-76), loss to follow up did not appear to differ markedly by any of our key exposure or outcome variables. Though there was some differential loss to follow up related to age at enrollment, this was most likely due to graduation on the part of students in the ‘21-23’ age category. As a result, the final cohort contained a slightly larger proportion of older women (age 27+) than the original cohort.

<table>
<thead>
<tr>
<th>Distribution of selected variables at baseline among women lost to follow up and women who completed follow up</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline characteristic</td>
</tr>
<tr>
<td>--------------------------</td>
</tr>
<tr>
<td>Age at enrollment</td>
</tr>
<tr>
<td>17-20</td>
</tr>
<tr>
<td>21-23</td>
</tr>
<tr>
<td>24-26</td>
</tr>
<tr>
<td>27+</td>
</tr>
<tr>
<td>Number of new sex partners</td>
</tr>
<tr>
<td>0</td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>2</td>
</tr>
<tr>
<td>HPV status (visit 1)</td>
</tr>
<tr>
<td>Negative</td>
</tr>
<tr>
<td>Positive</td>
</tr>
<tr>
<td>β-globin negative</td>
</tr>
<tr>
<td>SIL status (visit 1)</td>
</tr>
<tr>
<td>Normal</td>
</tr>
<tr>
<td>ASCUS</td>
</tr>
<tr>
<td>LSIL</td>
</tr>
<tr>
<td>HSIL</td>
</tr>
</tbody>
</table>
Appendix B: Assessment of Generalizability

As can be seen in the table below (adapted from (12) p. 54), external comparisons were conducted to assess the representativeness of the cohort, and subsequently, the generalizability of the coinfection estimates. Most notably, in comparison to the NPHS (restricted to women currently in school, or to women age 20-45 who had already received a post-secondary degree), women in the Montreal cohort appeared to be less sexually active (e.g. fewer sex partners in the last year). This may indicate that women at McGill/Concordia with more sex partners may have chosen not to participate in the study, which suggests that our coinfection estimates may be an underestimate compared to other populations of female students. Other minor differences were observed for condom use and proportion of daily smokers, however these were not substantial.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Montreal cohort (n=621) (%)</th>
<th>NPHS (n=2224) (%)</th>
<th>CCS (n=7800) (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of sex partners (last year)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0-1</td>
<td>87.5</td>
<td>61.9</td>
<td>NA</td>
</tr>
<tr>
<td>2</td>
<td>7.5</td>
<td>19.8</td>
<td>NA</td>
</tr>
<tr>
<td>3+</td>
<td>5.0</td>
<td>18.3</td>
<td>NA</td>
</tr>
<tr>
<td>Condom use (last year)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Never</td>
<td>26.9</td>
<td>15.0</td>
<td>27.9</td>
</tr>
<tr>
<td>Sometimes</td>
<td>27.5</td>
<td>30.4</td>
<td>NA</td>
</tr>
<tr>
<td>Always</td>
<td>45.6</td>
<td>54.6</td>
<td>28.9</td>
</tr>
<tr>
<td>History of STIs (last year)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Never</td>
<td>94.5</td>
<td>95.7</td>
<td>NA</td>
</tr>
<tr>
<td>Ever</td>
<td>5.5</td>
<td>4.3</td>
<td>NA</td>
</tr>
<tr>
<td>Daily smokers</td>
<td>25.0</td>
<td>19.4</td>
<td>12.1</td>
</tr>
<tr>
<td>Alcoholic drinks (weekly)</td>
<td>3.1</td>
<td>3.3</td>
<td>3.9</td>
</tr>
<tr>
<td>SIL status (visit 1)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Normal</td>
<td>97.1</td>
<td></td>
<td>97.3</td>
</tr>
<tr>
<td>LSIL</td>
<td>2.7</td>
<td>NA</td>
<td>2.7</td>
</tr>
<tr>
<td>HSIL</td>
<td>0.2</td>
<td></td>
<td>0.2</td>
</tr>
</tbody>
</table>
Appendix C: Ethics Approval

QUEEN'S UNIVERSITY HEALTH SCIENCES & AFFILIATED TEACHING HOSPITALS RESEARCH ETHICS BOARD

September 8, 2010

Ms. Michaela Smith
Department of Community Health and Epidemiology
Carruthers Hall
Queen's University

Dear Ms. Smith,

Study Title: Examining the occurrence and determinants of HPV coinfections in a cohort of Montreal University students

Co-Investigators: Dr. H. Richardson and Dr. E. Franco

I am writing to acknowledge receipt of your recent ethics submission. We have examined the protocol for your project (as stated above) and consider it to be ethically acceptable. This approval is valid for one year from the date of the Chair’s signature below. This approval will be reported to the Research Ethics Board. Please attend carefully to the following list of ethics requirements you must fulfill over the course of your study:

➢ Reporting of Amendments: If there are any changes to your study (e.g. consent, protocol, study procedures, etc.), you must submit an amendment to the Research Ethics Board for approval. (see http://www.queensu.ca/vpr/reb.htm).

➢ Reporting of Serious Adverse Events: Any unexpected serious adverse event occurring locally must be reported within 2 working days or earlier if required by the study sponsor. All other serious adverse events must be reported within 15 days after becoming aware of the information.

➢ Reporting of Complaints: Any complaints made by participants or persons acting on behalf of participants must be reported to the Research Ethics Board within 7 days of becoming aware of the complaint. Note: All documents supplied to participants must have the contact information for the Research Ethics Board.

➢ Annual Renewal: Prior to the expiration of your approval (which is one year from the date of the Chair’s signature below), you will be reminded to submit your renewal form along with any new changes or amendments you wish to make to your study. If there have been no major changes to your protocol, your approval may be renewed for another year.

Yours sincerely,

[Signature]
Chair, Research Ethics Board

DATE

Study Code: EPID-322-10

➢ Investigators please note that if your trial is registered by the sponsor, you must take responsibility to ensure that the registration information is accurate and complete
Appendix D: Consent Form

Study on Persistent HPV Infection and Cervical Intraepithelial Neoplasia

Consent Form

I ___________________________ freely consent to participate in the research project of the above title. The following aspects of the study have been explained to me.

A. Purpose of this study

This project conducted by McGill scientists seeks to investigate the occurrence of transient and persistent Human Papillomavirus (HPV) infection and cervical lesions detected by the Pap test in female university students and the related life style risk factors. HPV is a virus (the human papilloma virus) which is normally associated with asymptomatic infections of the genital area. HPV infection is detected by collecting samples of cells from the cervix of the uterus. The sample is then examined to determine the presence of HPV. If HPV is detected, further analysis is conducted to classify the type of HPV.

B. Procedure

If I agree to participate, I will be asked to complete eight self-administered questionnaires, 2 interviews and undergo five Pap smears. The pap specimens will be sent to a lab to be tested for evidence of HPV infection or any alterations that may much later in my life develop into cancer.

C. Personal Inconvenience accompanying participation in this study

It is recommended by Canadian health officials that all sexually active women undergo a Pap smear at least once a year to detect gynecological malignancy. For purposes of this study I will be required to have 5 Pap Smears over a period of 24 months at 6 month intervals. For example.

Jan 97  June 97  Dec. 97  June 98  Dec. 98
1st Pap  2nd Pap  3rd Pap  4th Pap  5th Pap
1st mth  6 mth    12 mth   18 mth   24 mth

The interviews and self-report questionnaires will be scheduled to coincide with my final two visits. A member of the research team will contact me by phone prior to each scheduled visit to remind me of my appointment and obtain the necessary information for each visit.

D. Compliance

It is imperative for statistical purposes to obtain the lab results of all five Pap Smears as well as responses to the questionnaires. The researchers are aware of the inconveniences of returning to the clinic every six months and will pay you $20 per follow-up return visits. If you complete the entire study you will receive a total of $80.00

E. Risks and Benefits

The risks in this study are minimal as the Pap smear is a safe examination. As with any gynecological examination, there is a possibility that a slight discomfort might be felt during the insertion of the cervical sampler to collect the Pap smear. The benefits of this project include improved characterization of the suspected virus and increase our knowledge of a potentially hazardous disease. As well if any lesions are detected we will notify your doctor at the clinic so that you can be treated if necessary.
F. Confidentiality

In order to ensure my privacy and confidentiality my name will not appear on any record or results. Instead the patient identification number will be assigned to me and will appear on all my records. My patient number will be kept on file at the McGill University Students Health Services clinic and only the investigator and the assistants will have access to the study number. I understand that all information about me or my Pap smear results will be treated in the same confidential manner as other medical records and I will not be identified in any subsequent reporting of results.

G. Voluntary consent

I understand that my participation is entirely voluntary and that I may withdraw at any time without affecting my status at the McGill University Student Health Services Clinic. By signing this consent form I acknowledge that this research study has been thoroughly explained to me and I fully understand that I will have to commit to making four additional visits once every six months for 24 months. I have the opportunity to ask questions and to seek further information about the procedure and the results of the study. I understand that I am free to ask additional questions in the future and that my identity will remain confidential.

Participant (signature)                                      Clinical Nurse coordinator (signature)

Print Name                                                  Print Name

Date                                                        Date
Appendix E: Enrollment Questionnaire

STUDY NO: _____

WOMEN'S HEALTH STUDY
INITIAL QUESTIONNAIRE

McGill University Student Health Services
Departments of Oncology and Epidemiology & Biostatistics

INSTRUCTIONS FOR THE QUESTIONNAIRE

This questionnaire is composed of the following sections:

- General information
- Diet History
- Smoking history and alcohol consumption
- Reproductive history
- Sexual history
- Contraceptive history
- Personal hygiene habits
- Medical history

Most questions require that you simply check a box ☐ with an "X" to indicate your choice. Other questions require a specific answer, such as age, date, or another number. Depending on your answer for some questions, you will be told to skip the next question and go to a different part of the questionnaire. This is to save you time, so that you won’t have to go over questions that do not apply to you.

There are no right or wrong answers to any question. Many questions require that you think back over your adult years, particularly over the past year, to recall specific information. Please take the time to reflect. If you prefer, you can answer sections of the questionnaire on different days. If you choose to do so, check your answers from previous days to make sure you agree with them before mailing the questionnaire back to us. You will be surprised that by being "forced" to recall specific information of one type, some of the answers for other questions may come more naturally to you later on. If you can’t possibly remember the information skip the question, but we would like to encourage you to try to answer all questions. A good guess is always better than no information at all. If you’d like to tell us more about any specific items please use the available space at the end of the questionnaire.

WE APPRECIATE YOUR COOPERATION WITH THE STUDY
GENERAL INFORMATION

This portion of the questionnaire concerns general information about you and where you live.

1. What is your date of birth? ______ / ______ / ______ (very important)
   D   M   Y

2. In what country were you born? ____________________________
   If born in Canada: indicate province: _______________________

3. What is your current marital status?
   ❑ Married  ❑ Single
   ❑ Unmarried, but living with a partner  ❑ Divorced/separated
   ❑ Widowed

4. The Montreal area is made up of many ethnic groups. We would like to know in which group you would place yourself. Check the most appropriate category:
   ❑ French Canadian  ❑ Hispanic/Portuguese
   ❑ English Canadian  ❑ Greek
   ❑ Black Canadian  ❑ Italian
   ❑ Native Indian  ❑ Asian/Oriental
   ❑ Jewish  Other: _______________________

5. a) What is/was your father’s occupation? _______________________
   b) What is/was your mother’s occupation? _______________________
   c) Would you say that your family’s financial situation while growing up was:
      ❑ Difficult  ❑ Moderate  ❑ Very comfortable

6. How are you presently enrolled at McGill?
   ❑ Undergraduate - Regular Student (State year: ________)
   ❑ Graduate studies - Diploma, Master’s or Doctoral Program
   ❑ Other (e.g. Trainee, Postdoctoral Studies, Sabbatical)
This section of the questionnaire concerns some specific food and beverage items. We want to know about your usual adult diet, that is, your usual eating habits during all your adult years.

7. For each item check the category that best reflects your average consumption pattern. Try a good guess considering a typical serving and any cooking method:

<table>
<thead>
<tr>
<th>Item</th>
<th>At least once a day</th>
<th>&lt;1 per day &gt;1 per week</th>
<th>At least once a week</th>
<th>At least once a month</th>
<th>Never</th>
</tr>
</thead>
<tbody>
<tr>
<td>a) carrots</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>b) spinach</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c) broccoli</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>d) lettuce</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>e) cabbage</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>f) cheese or cream</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>g) milk or yoghurt</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>h) liver</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>i) pure orange juice</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>j) fresh orange or grapefruit</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>k) vitamin C-fortified fruit drink</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>l) vitamin-C supplements</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>m) mixed vegetable juices (V-8, garden cocktail, tomato juice)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
SMOKING HISTORY AND ALCOHOL CONSUMPTION

The following questions are about your tobacco smoking and alcoholic beverage consumption habits. Please try to be as specific as possible in your answers.

8. Have you ever smoked cigarettes regularly, that is, one cigarette or more each day for a year or more?
   - Yes
   - No

9. Have you smoked a total of at least 100 cigarettes in your lifetime?
   - Yes
   - No
   - If No, go to question 13

10. At what age did you start to smoke? ______ years

11. Do you still smoke?
    - Yes
    - No
    - If No, at what age did you stop? ______ years

12. On average, how many cigarettes do/did you smoke a day? ______ cigarettes per day

13. Has there ever been a period in your life when you drank beer, wine or liquor AT LEAST ONCE A MONTH?
    - Yes
    - No
    - If No, go to question 15

14. Has there ever been a period in your life when you drank beer, wine or liquor AT LEAST ONCE A WEEK?
    - Yes
    - No
    - If No, go to question 15
    - If Yes, indicate the average number of drinks per week that you consumed during the past 5 years (consider a drink as being equivalent to a 12 oz. can of beer or to a 4 oz. glass of wine or to 1.5 ounces of hard liquor such as gin, vodka, whiskey, scotch, rum, tequila, etc.).
      a) Beer: ______ cans per week
      b) Wine: ______ glasses per week
      c) Liquor: ______ drinks per week
REPRODUCTIVE HISTORY

In this section of the questionnaire we would like to know about your reproductive health including all your pregnancies as well as miscarriages and abortions.

15. At what age did you have your first menstrual period? _______ years

16. To the best of your knowledge, are you currently pregnant?
   - Yes
   - No
   - Don’t know

17. Have you ever been pregnant before?
   - Yes
   - No
   If Yes, go to question 20
   how many times? _______ times

18. How many of your pregnancies resulted in:
   a) livebirths: ______ 
   b) stillbirth: ______ 
   c) miscarriage______ 
   d) abortion: ______ 

19. How many of your full-term pregnancies resulted in:
   a) vaginal deliveries: ______ 
   b) cesarean sections: ______

page 5
SEXUAL HISTORY

The next questions are about your sexual history. We realize this is a personal subject, but it is very important to the study. Please take the time to recall this information as accurately as possible. Note that some questions in this section refer to your entire life as an adult, whereas others refer only to your recent experience. We would like to remind you that all the information you give us will be kept entirely confidential.

20. Have you ever engaged in vaginal sexual intercourse?
   ☐ Yes  ☐ No
   → If No, go to question 27
   If Yes, how old were you when you first had vaginal sexual intercourse? ________ years

21. THROUGHOUT YOUR LIFE, what is the number of male partners with whom you have had vaginal sexual intercourse?
   Number (approximately) ________

22. With how many of these male partners did you have a sexual relationship involving intercourse on a regular basis for three months or longer?
   Number ________  ☐ None

23. For MOST OF YOUR SEXUALLY ACTIVE LIFE, how often on the average, did you have vaginal sexual intercourse? Please give your answer in number of times per week, month, or year, whichever is easiest:
   Number of times per week ________
   OR
   Number of times per month ________
   OR
   Number of times per year ________
   OR
   Less than once a year ☐

24. During THE LAST YEAR ONLY, what is the number of male partners with whom you have had vaginal sexual intercourse?
   ________ Number  ☐ None in the past year
   → How many of those partners were new? ________ Number
25. In THE LAST YEAR ONLY, how often on the average did you have vaginal sexual intercourse? Please give your answer in number of times per week, month, or year, whichever is easiest:

   Number of times per week_______
   OR
   Number of times per month_______
   OR
   Number of times per year ________
   OR
   Less than once a year □

26. When you are having your menstrual periods, do you have vaginal sexual intercourse?
   □ Yes  □ No

27. THROUGHOUT YOUR LIFE, has anyone ever performed oral sex on you?
   ___________ Number  □ None
   If No, go to question 31

28. How often on average, did you receive oral sex? Please give your answer in number of times per week, month, or total per year, whichever is easiest:

   Number of times per week_______
   OR
   Number of times per month_______
   OR
   Number of times per year ________
   OR
   Less than once a year □

29. During THE LAST YEAR ONLY, how many people performed oral sex on you?
   ___________ Number  □ None in the past year
30. During THE LAST YEAR ONLY, how often on average, did you receive oral sex?
Please give your answer in number of times per week, month, or year, whichever is easiest:

Number of times per week_______
OR
Number of times per month_______
OR
Number of times per year _______
OR
Less than once a year □

31. Do you ever practice anal intercourse?
□ Yes □ No

   If yes, would you say that you have had anal intercourse:
□ Frequently □ Occasionally □ Rarely

32. With whom do you usually have sex?
□ Men □ Women □ Both

33. Do you ever masturbate?
□ Yes □ No

   If yes, do you ever insert objects into the vagina for stimulation:
□ Yes □ No
CONTRACEPTIVE HISTORY

*Here we would like to know about methods of birth control or family planning that you and your husband/partner used. It would be important to indicate all the methods you’ve used since you became sexually active. If you answered “No” to question 20 you may skip this section entirely and go to question 38.*

34. The following is a list of common birth control methods. Read along the list and check if you and a sex partner have ever used any of them (check all that apply) either occasionally or regularly.

**BY REGULARLY WE MEAN AT LEAST 3 MONTHS CONSECUTIVELY**

<table>
<thead>
<tr>
<th>Method</th>
<th>Regularly</th>
<th>Sometimes</th>
<th>Never</th>
</tr>
</thead>
<tbody>
<tr>
<td>a) oral contraceptive (birth control pill)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>b) condom (rubber)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c) foam, jelly, cream, or suppository</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>d) loop, coil, or other intrauterine device</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>e) diaphragm</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>f) cervical cap</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>g) sponge</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>h) vaginal douche</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>i) rhythm, calendar, or natural method</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>j) withdrawal/pulling out</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

35. If you have used oral contraceptives or birth control pills, please indicate how old you were when you first took them?

<table>
<thead>
<tr>
<th>Age: _______ years</th>
<th>Never used oral contraceptives</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Go to question 37</td>
</tr>
</tbody>
</table>
36. Considering only the times when you were taking the pill, for how long have you been relying on this method of birth control (add together all periods during which you took any oral contraceptives)?

___ months

OR

___ years

OR

☒ all periods combined were less than 3 months

37. Now, considering ONLY THE LAST YEAR, on the average, which of the following birth control methods have you or your partner come to rely upon? (check all that apply)

**BY REGULARLY WE MEAN AT LEAST 3 MONTHS CONSECUTIVELY**

<table>
<thead>
<tr>
<th>Method</th>
<th>Regularly</th>
<th>Sometimes</th>
<th>Never</th>
</tr>
</thead>
<tbody>
<tr>
<td>a) oral contraceptive (birth control pill)</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>b) condom (rubber)</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>c) foam, jelly, cream, or suppository</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>d) loop, coil, or other intrauterine device</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>e) diaphragm</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>f) cervical cap</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>g) sponge</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>h) vaginal douche</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>i) rhythm, calendar, or natural method</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>j) withdrawal/pulling out</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
</tbody>
</table>
### PERSONAL HYGIENE HABITS

This section concerns your personal hygiene habits. As with other topics in this questionnaire, this is also a personal subject of great importance to the study. Again, rest assured that we will treat your answers confidentially. Please take the time to recall this information as accurately as possible. A good guess is always better than leaving the question blank.

#### 38. How many times per day or per week do you usually bathe or shower?

<table>
<thead>
<tr>
<th></th>
<th>per day</th>
<th>OR</th>
<th>per week</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

#### 39. Aside from those baths and showers, do you ever wash your genital area?

(Do not consider the times you may wash after sexual intercourse.)

- [ ] Yes
- [ ] No

If Yes,

in the last year, on the average, how many times per day, week, or month did you wash your genital area?

<table>
<thead>
<tr>
<th></th>
<th>per day</th>
<th>OR</th>
<th>per week</th>
<th>OR</th>
<th>per month</th>
<th>OR</th>
<th>less than once a month</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

#### 40. Have you ever used a vaginal douche?

- [ ] Yes
- [ ] No

If Yes,  

in the last year, on the average, how many times per day, week, or month did you use a vaginal douche?

<table>
<thead>
<tr>
<th></th>
<th>per day</th>
<th>OR</th>
<th>per week</th>
<th>OR</th>
<th>per month</th>
<th>OR</th>
<th>less than once a month</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

If No, go to question 42

#### 41. If you answered yes in question 40, what kind of douche did you use?

- [ ] Water only
- [ ] Water and vinegar
- [ ] Other (specify:_____________)
42. Have you ever used a feminine genital spray or vaginal deodorant?
   - Yes
   - No

   If Yes,
   - in the last year, on the average, how many times per day, week, or month did you use a vaginal deodorant?
     - Per day
     - Per week
     - Per month
     - Less than once a month

43. When you are menstruating, what do you use to collect the blood? (check all that apply)
   - Sanitary pads/napkins
   - Tampons
   - Other

44. Following sexual intercourse, do you usually wash your genital area within the hour?
   Choose the category that best reflects your behaviour during most of your adult life: (skip this question if you never had sexual intercourse).
   - Always
   - Sometimes
   - Rarely
   - Never

45. Following oral sex, do you usually wash your genital area within the hour? Choose the category that best reflects your behaviour during most of your adult life: (skip this question if you never practice oral sex).
   - Always
   - Sometimes
   - Rarely
   - Never

46. What is your preferred hygiene practice after each bowel movement? (Check all that apply)
   - Use toilette paper with a BACK-TO-FRONT hand motion for wiping
   - Use toilette paper with a FRONT-TO-BACK hand motion for wiping
   - Wash with water only
   - Wash with water and soap
MEDICAL HISTORY

The next questions are about the frequency with which you have taken PAP smears and about some medical problems including sexually transmitted diseases. We realize that this is a sensitive subject but, again, it is very important to the research. We appreciate your honesty and want to remind you that all information you give us is kept private and confidential.

47. Thinking back over your adult years, how often have you usually had a PAP smear? Choose one category below:
   - this is my first PAP smear
   - 2-3 times
   - 4-5 times
   - 6-10 times
   - more than 10 times

48. What is the month and year of the last PAP smear you had?  
   Month _______/_______ Year

49. Did a doctor ever tell you that you had one of the following conditions? Check all that apply, if you are in doubt check the "don't know" column.
   a) Vaginal yeast infections:  
      - Yes  
      - No  
      - Don't know
   b) Trichomonas vaginal infections:  
      - Yes  
      - No  
      - Don't know
   c) Venereal warts, condylomas, or papilloma virus infections:  
      - Yes  
      - No  
      - Don't know
   d) Chlamydia:  
      - Yes  
      - No  
      - Don't know
   e) Genital herpes:  
      - Yes  
      - No  
      - Don't know
   f) Syphilis:  
      - Yes  
      - No  
      - Don't know
   g) Gonorrhea:  
      - Yes  
      - No  
      - Don't know
   h) Ulcers or genital sores:  
      - Yes  
      - No  
      - Don't know
50. Thinking back over all your adult life, have you experienced other genital conditions such as vaginal discharge, itching or irritation?
   - Never
   - Less than once a year
   - More than once a year

51. Now, only during the last year, have you experienced other genital conditions such as vaginal discharge, itching or irritation?
   - Never
   - Once or twice
   - More than 3 times last year

52. Sometimes women are given female hormones by their doctors because of a variety of reasons (alleviate acne, regulate or eliminate painful periods, menopausal symptoms, reduce discomfort during intercourse due to vaginal dryness, prevent miscarriage, among others). To the best of your recollection, were you ever prescribed any female hormones by your doctor?
   - Yes
   - No
   If Yes, go to question 55
   If No, go to question 55
   in what month and year did you start taking them and also, in what month and year did you last take them?
   Start: _______ / _______     End: _______ / _______
   month     year      month     year

53. Between the above two dates, for how long (number of months) did you take the female hormone medication on a continual basis, altogether?
   _______ months

54. Was the female hormone medication in the form of (check all that apply):
   - pills
   - shots
   - creams or suppositories
55. Would you please indicate the date when you finished filling in the questionnaire?

\[ \underline{\text{DAY}} \quad \underline{\text{MONTH}} \quad \underline{\text{YEAR}} \]

USE THE SPACE BELOW IF YOU HAVE ANY ADDITIONAL INFORMATION YOU FEEL WOULD BE IMPORTANT FOR US TO KNOW:

________________________________________________________________________

________________________________________________________________________

________________________________________________________________________

________________________________________________________________________

________________________________________________________________________

This is the end of the questionnaire. We would like you to take a few seconds to review your answers in all sections of the questionnaire. If you answered the sections on different days, take a moment to reflect if you agree now with your answers from previous days. Again, try to answer all questions; a good guess will be more useful to the study than leaving the question blank.

THANK YOU VERY MUCH FOR YOUR COOPERATION
Appendix F: Follow-up Questionnaire

STUDY NO: ______

WOMEN'S HEALTH STUDY FOLLOW-UP QUESTIONNAIRE

McGill University Student Health Services
Departments of Oncology and Epidemiology & Biostatistics

INSTRUCTIONS FOR THE QUESTIONNAIRE

This questionnaire is composed of the following sections:

General information
Sexual history
Contraceptive history
Personal hygiene habits
Medical history

Most questions require that you simply check a box ✅ with an "X" to indicate your choice. Other questions require a specific answer, such as age, date, or another number. Depending on your answer for some questions, you will be told to skip the next question and go to a different part of the questionnaire. This is to save you time, so that you won't have to go over questions that do not apply to you.

Many questions also refer to the period since your last visit to the clinic, a few months ago, when you were given a similar questionnaire. In those instances the questions will start with "since your last visit... ". There are no right or wrong answers to any question. Many questions require that you think back over your adult years, particularly over the past year, to recall specific information. Please take the time to reflect. If you prefer, you can answer sections of the questionnaire on different days. If you choose to do so, check your answers from previous days to make sure you agree with them before mailing the questionnaire back to us. You will be surprised that by being "forced" to recall specific information of one type, some of the answers for other questions may come more naturally to you later on. If you can't possibly remember the information skip the question, but we would like to encourage you to try to answer all questions. A good guess is always better than no information at all. If you'd like to tell us more about any specific items please use the available space at the end of the questionnaire.

Once you have completed the questionnaire please return it to the study nurse at the clinic.

WE APPRECIATE YOUR COOPERATION WITH THE STUDY
STUDY NO: ______

GENERAL INFORMATION

This portion of the questionnaire concerns general information about you and where you live.

1. What is your date of birth? _____ / _____ / _____ (very important)
   D   M   Y

2. In what country were you born? ___________________________
   If born in Canada: indicate province: _______________________

3. What is your current marital status?
   ☐ Married
   ☐ Single
   ☐ Unmarried, but living with a partner
   ☐ Divorced/separated
   ☐ Widowed

4. On average, how many cigarettes have you smoked since your last visit?
   Number of cigarettes: ______ per day OR ______ per week OR ☐ Non-smoker

5. Indicate the average number of drinks per week that you consumed since your last visit (consider a drink as being equivalent to a 12 oz. can of beer or to a 4 oz. glass of wine or to 1.5 ounces of hard liquor such as gin, vodka, whiskey, scotch, rum, tequila, etc.).
   a) Beer: ______ cans per week
   b) Wine: ______ glasses per week
   c) Liquor: ______ drinks per week
   d) None since last visit ☐
SEXUAL HISTORY

The next questions are about your sexual history. We realize this is a personal subject, but it is very important to the study. Please take the time to recall this information as accurately as possible. Note that some questions in this section refer to your entire life as an adult, whereas others refer only to your recent experience. We would like to remind you that all the information you give us will be kept entirely confidential.

6. Have you ever engaged in vaginal sexual intercourse?
   - Yes
   - No
   _____ If No, go to question 11
   - Yes,
   how old were you when you first had vaginal sexual intercourse? _______ years

7. THROUGHOUT YOUR LIFE, what is the number of male partners with whom you have had vaginal sexual intercourse?
   Number (approximately) _______

8. SINCE YOUR LAST VISIT, what is the number of male partners with whom you have had vaginal sexual intercourse?
   _______ Number
   - None since last visit
   How many of those partners were new? _______ Number

9. SINCE YOUR LAST VISIT, how often on the average, did you have vaginal sexual intercourse? Please give your answer in number of times per week, or month, whichever is easiest:
   Number of times per week _______
   OR
   Number of times per month _______
   OR
   Never since last visit

10. When you are having your menstrual periods, do you have vaginal sexual intercourse?
    - Yes
    - No
11. SINCE YOUR LAST VISIT, how many people performed oral sex on you?
   ________ Number     ❑ None since last visit
   ➔ How many of those partners were new? ________ Number

12. SINCE YOUR LAST VISIT, how often on average, did you receive oral sex? Please give your answer in number of times per week, or month, whichever is easiest:
   Number of times per week ________
   OR
   Number of times per month ________
   OR
   Never since last visit ❑

13. SINCE YOUR LAST VISIT, have you practiced anal intercourse?
    ❑ Yes          ❑ No
    ➔ If yes, would you say that you have had anal intercourse:
    ❑ Frequently   ❑ Occasionally   ❑ Rarely

14. SINCE YOUR LAST VISIT, with whom do you usually have sex?
    ❑ Men               ❑ Women               ❑ Both

15. Do you ever masturbate?
    ❑ Yes          ❑ No
    ➔ If yes, do you ever insert objects into the vagina for stimulation:
    ❑ Yes          ❑ No

16. To the best of your knowledge, are you currently pregnant?
    ❑ Yes               ❑ No               ❑ I don't know
CONTRACEPTIVE HISTORY

Here we would like to know about methods of birth control or family planning that you and your husband/partner used. If you answered "No" to question 6 you may skip this section entirely and go to question 18.

17. SINCE YOUR LAST VISIT, on the average, which of the following birth control methods have you and your partner come to rely upon? (check all that apply)

BY REGULARLY WE MEAN AT LEAST 3 MONTHS CONSECUTIVELY

<table>
<thead>
<tr>
<th>I was not sexually active</th>
<th>Go to Question 18</th>
</tr>
</thead>
<tbody>
<tr>
<td>a) oral contraceptive (birth control pill)</td>
<td>□ Regularly □ Sometimes □ Never</td>
</tr>
<tr>
<td>b) condom (rubber)</td>
<td>□ Regularly □ Sometimes □ Never</td>
</tr>
<tr>
<td>c) foam, jelly, cream, or suppository</td>
<td>□ Regularly □ Sometimes □ Never</td>
</tr>
<tr>
<td>d) loop, coil, or other intrauterine device</td>
<td>□ Regularly □ Sometimes □ Never</td>
</tr>
<tr>
<td>e) diaphragm</td>
<td>□ Regularly □ Sometimes □ Never</td>
</tr>
<tr>
<td>f) cervical cap</td>
<td>□ Regularly □ Sometimes □ Never</td>
</tr>
<tr>
<td>g) sponge</td>
<td>□ Regularly □ Sometimes □ Never</td>
</tr>
<tr>
<td>h) vaginal douche</td>
<td>□ Regularly □ Sometimes □ Never</td>
</tr>
<tr>
<td>i) rhythm, calendar, or natural method</td>
<td>□ Regularly □ Sometimes □ Never</td>
</tr>
<tr>
<td>j) withdrawal/pulling out</td>
<td>□ Regularly □ Sometimes □ Never</td>
</tr>
</tbody>
</table>
PERSONAL HYGIENE HABITS

This section concerns your personal hygiene habits. As with other topics in this questionnaire, this is also a personal subject of great importance to the study. Again, rest assured that we will treat your answers confidentially. Please take the time to recall this information as accurately as possible. A good guess is always better than leaving the question blank.

18. Aside from baths and showers, do you ever wash your genital area? (Do not consider the times you may wash after sexual intercourse.)
   - Yes
   - No

   If Yes,
   since your last visit, on the average, how many times per day, week, or month did you wash your genital area?
   ———— per day
   OR
   ———— per week
   OR
   ———— per month
   OR
   less than once a month

19. SINCE YOUR LAST VISIT, following vaginal sexual intercourse, do you usually wash your genital area within the hour? Choose the category that best reflects your behaviour, since your last visit: (skip this question if you never had sexual intercourse).
   - always
   - sometimes
   - rarely
   - never

20. SINCE YOUR LAST VISIT, following oral sex, do you usually wash your genital area within the hour? Choose the category that best reflects your behaviour, since your last visit: (skip this question if you never practice oral sex).
   - always
   - sometimes
   - rarely
   - never

21. SINCE YOUR LAST VISIT, did you use a vaginal douche?
   - Yes
   - No

   If Yes, how many times? ——————

What kind of douche did you use?
   - Water only
   - Water and vinegar
   - Other (specify: ——————)
MEDICAL HISTORY

The next questions are about some medical problems including sexually transmitted diseases. We realize that this is a sensitive subject but, again, it is very important to the research. We appreciate your honesty and want to remind you that all information you give us is kept private and confidential.

22. SINCE YOUR LAST VISIT, did a doctor tell you that you had one of the following conditions? Check all that apply, if you are in doubt check the "don't know" column.
   a) Vaginal yeast infections:  □ Yes  □ No  □ Don't know
   b) Trichomonas vaginal infections:  □ Yes  □ No  □ Don't know
   c) Venereal warts, condylomas, or papilloma virus infections:  □ Yes  □ No  □ Don't know
   d) Chlamydia:  □ Yes  □ No  □ Don't know
   e) Genital herpes:  □ Yes  □ No  □ Don't know
   f) Syphilis:  □ Yes  □ No  □ Don't know
   g) Gonorrhea:  □ Yes  □ No  □ Don't know
   h) Ulcers or genital sores:  □ Yes  □ No  □ Don't know

23. SINCE YOUR LAST VISIT, have you experienced other genital conditions such as vaginal discharge, itching or irritation?
   □ Never since last visit
   □ Once or twice
   □ More than 3 since last visit
24. Would you please indicate the date when you finished filling in the questionnaire?

__/__/__
DAY MONTH YEAR

USE THE SPACE BELOW IF YOU HAVE ANY ADDITIONAL INFORMATION YOU FEEL WOULD BE IMPORTANT FOR US TO KNOW:

__________________________________________________________________________________________

__________________________________________________________________________________________

__________________________________________________________________________________________

__________________________________________________________________________________________

This is the end of the questionnaire. We would like you to take a few seconds to review your answers in all sections of the questionnaire. If you answered the sections on different days, take a moment to reflect if you agree now with your answers from previous days. Again, try to answer all questions; a good guess will be more useful to the study than leaving the question blank.

THANK YOU VERY MUCH FOR YOUR COOPERATION