In compliance with the
Canadian Privacy Legislation
some supporting forms
may have been removed from
this dissertation.

While these forms may be included
in the document page count,
their removal does not represent
any loss of content from the dissertation.

LONGITUDINAL RELATIONSHIP BETWEEN HUMAN PAPILLOMAVIRUS INFECTION AND THE INCIDENCE AND PROGRESSION OF PRECURSOR LESIONS OF CERVICAL NEOPLASIA

Nicolas Schlecht

Joint Departments of Epidemiology, Biostatistics and Occupational Health

McGill University, Montreal, Canada

November 2002

A thesis submitted to the Faculty of Graduate Studies and Research in partial fulfillment of the requirements of the degree of Doctor of Philosophy

Your file Votre référence
ISBN: 0-612-88575-5
Our file Notre référence
ISBN: 0-612-88575-5

SHORT TITLE OF THESIS

Longitudinal Relationship of HPV and CIN

TABLE OF CONTENTS

PREFACE

Description of Thesis

The format of this thesis follows that of a manuscript-based thesis. This dissertation consists of a collection of papers of which the student is an author or co-author. As per McGill University requirements, the papers must have a cohesive, unitary character making them a report of a single program of research. The structure for the manuscript-based thesis must conform to the following:

1. Candidates have the option of including, as part of the thesis, the text of one or more papers submitted, or to be submitted, for publication, or the clearly duplicated text (not the reprints) of one or more published papers. These texts must be bound together as an integral part of the thesis. (Reprints of published papers can be included in the appendices at the end of the thesis.)

2. The thesis must be more than a collection of manuscripts. All components must be integrated into a cohesive unit with a logical progression from one chapter to the next. In order to ensure that the thesis has continuity, connecting texts that provide logical bridges between the different papers are mandatory.

3. The thesis must conform to all other requirements of the "Guidelines for Thesis Preparation" in addition to the manuscripts. The thesis must include the following: (a) a table of contents; (b) an abstract in English and French; (c) an introduction which clearly states the rationale and objectives of the research; (d) a comprehensive review of the literature (in addition to that covered in the introduction to each paper); (e) a final conclusion and summary.

4. As manuscripts for publication are frequently very concise documents, where appropriate, additional material must be provided (e.g., in appendices) in sufficient detail to allow a clear and precise judgment to be made of the importance and originality of the research reported in the thesis.

5. When co-authored papers are included in a thesis the candidate must have made a substantial contribution to all papers included in the thesis. In addition, the candidate is required to make an explicit statement in the thesis as to who contributed to such work and to what extent. The supervisor must attest to the accuracy of this statement at the doctoral oral defense.

List of Original Papers

This thesis is based on the following original manuscripts:

I. Schlecht NF, Kulaga S, Robitaille J, Ferreira S, Santos M, Miyamura RA, Duarte-Franco E, Rohan TE, Ferenczy A, Villa LL, Franco EL. Persistent human papillomavirus infection as a predictor of cervical neoplasia. ***JAMA***. 2001 Dec 26; 286(24):3106-14.

II. Schlecht NF, Trevisan A, Duarte-Franco E, Rohan TE, Ferenczy A, Villa LL, Franco EL. Viral load as a predictor of lesion risk in the natural history of cervical intraepithelial neoplasia. ***Int J Cancer*** 2002; (*in press*)

III. Schlecht NF, Platt R, Negassa A, Duarte-Franco E, Rohan TE, Ferenczy A, Villa LL, Franco EL. Modeling the time dependence of the association between human papillomavirus infection and cervical cancer precursor lesions. ***Am J Epidemiol*** (*submitted*)

IV. Schlecht NF, Platt R, Duarte-Franco E, Rohan TE, Ferenczy A, Villa LL, Franco EL. Human papillomavirus infection as a predictor of the rate of progression and regression of cervical cancer precursor lesions.

Statement of Originality

This thesis project was based on an existing larger cohort study conducted in São Paulo, Brazil. The primary difference between my PhD project and others based from the Ludwig-McGill cohort study is that no one else has exploited the longitudinal nature of the cohort nor has anyone elsewhere evaluated the time dependent nature of the disease. This includes taking into account both the variability of infections by HPV their intermittency and the transient nature of precursor lesions of cervical cancer.

Other studies currently being carried out using this cohort's data are looking at HPV as an outcome and ancillary risk factors that are related to its acquisition. Therefore, the plan of this thesis is not to go into depth on co-factors – which include information on attitudes and beliefs, and diet (currently subjects of Master's thesis projects by other students in the Division) and reproductive health factors (subject of another PhD project on the natural history of the virus) – but to evaluate viral covariates (HPV persistence, viral burden, and time since infection) related to the incidence, progression, regression and sojourn time of precursor lesions of cervical cancer.

In the literature review, I will describe the difference between the approaches I am using in this project and those used by other research groups. I attempt to explain how differences in HPV status can influence the occurrence disease states for cervical neoplasia over time. Due to the transient nature of the disease, this requires specific statistical methods to handle the repeated measurements over time and sensitive testing methods of exposure and outcome.

Contributions of Authors

Members of the Ludwig-McGill Cohort Study and PhD Thesis supervisory committees who were listed as co-authors on the four articles provided epidemiological and methodological expertise (Drs. Franco and Rohan), expertise in laboratory methods and microbiology (Drs. Franco and Villa), clinical guidance (Dr. Ferenczy), and statistical expertise (Drs. Franco and Platt) in the design and conduct of the research and in the interpretation of the results.

The contributions of the other co-authors are gratefully acknowledged. Ms. Andrea Trevisan is a PhD student in microbiology from the University of São Paulo and Ludwig Institute for Cancer Research, under the supervision of Dr. Luisa Villa, who together contributed significantly to my understanding of biology of the human papillomavirus and of the testing methods for viral load. Ms. Monica Santos and Silvaneide Ferreira and Mr. Romulo Miyamura were responsible for testing of cervical specimens for HPV DNA. Maria L. Baggio and Lenice Galan are not listed as co-authors but were involved in the management of the study patients and specimen collection. Likewise João Pereira Sobrinho, Lara Termini, and José M. Prado were also responsible for HPV

testing. Dr. Eliane Duarte-Franco is a pathologist and family medicine physician who was central to the management of the Ludwig-McGill Cohort Study. Dr. Abdissa Negassa is a biostatistician who contributed to the initial framework for the statistical analyses on time-dependence developed for this project. Ms. Sophie Kulaga is a PhD student in epidemiology at McGill University who was involved in preliminary analyses of the Ludwig-McGill cohort study while working on her Master's thesis under the supervision of Dr. Eduardo Franco. Ms. Juliette Robitaille and Dr. Ferenczy performed all cytological analyses used in this project.

Statement of Financial Support

During the tenure of the project my work was funded by a pre-doctoral Scholarship from the Canadian Institutes of Health Research (CIHR) (DR56680). This PhD dissertation project was the subject of a CIHR grant (MOP49396) to fund the development of longitudinal analyses for the Ludwig-McGill cohort study of which I was responsible for writing. Preliminary results from this thesis were presented on different occasions at the International Papillomavirus Conferences and the European Research Organization on Genital Infection and Neoplasia (EUROGIN) International Multidisciplinary Conference for which I received travel scholarships from the Alma Mater Fund of McGill University, the Ludwig Institute for Cancer Research, my CIHR pre-doctoral fellowship, and from the International Papillomavirus Society (Charleston Award).

The Ludwig-McGill Cohort study is supported by intramural grants from the Ludwig Institute for Cancer Research and by grants from the U.S. National Cancer Institute (CA813100-01A1, CA70269) and the Canadian Institutes of Health Research (MOP49396, MA13647).

During the period of the doctoral thesis I was also employed as a research assistant in the Division of Cancer Epidemiology under the supervision of Eduardo Franco and was responsible for the coordination of the statistical analysis of the CIHR grant (MOP49396) study as well as for special projects derived from other studies of which Eduardo Franco is principal investigator or co-investigator.

Acknowledgements

This project could not have been accomplished without the contributions and assistance of many people. First and foremost, none of this would have been possible without the guidance and support of my supervisor Dr. Eduardo Franco, who was a source of encouragement and inspiration. His extensive knowledge in microbiology, epidemiology and public health made him invaluable resource for learning about cancer and epidemiology. Dr. Franco provides his students with the right balance between the freedom to be creative and develop our own research questions, and the guidance to keep us on the right track. He always has his door open to his

students and is always ready to tender a kind ear and lend a hand if we had trouble. I owe him a debt of gratitude for his constant support during my studies, as much financially and academically as emotionally, as well as for his help in my search for a career in academia. I am honored to have had him as a mentor and will always cherish his kindness and sincere friendship.

I am grateful to the team at the Ludwig Institute for Cancer Research for their expertise in the biology of the virus and their dedicated work "behind the scenes". I cannot begin to thank them for the countless hours of laboratory work they perform to generate the HPV data so instrumental to this research project. Dr. Luisa Villa's compassion and support, both for her students and for me, were always a source of encouragement. I wish to thank Dr. Paulo Maciag and Ms. Andrea Trevisan for their friendship, energy and enthusiasm for the area of research.

I woud also like to thank the members of my supervisory committee. As a physician and epidemiologist, Dr. Thomas Rohan provided much insight in the natural history of the disease and in cancer epidemiology in general. I knew I could always count on him for very prompt and in-depth reviews of my work. I am grateful for his confidence in me and for giving me the opportunity to work with him in the Department of Epidemiology and Social Medicine at Albert Einstein College of Medicine.

Dr. Robert Platt provided guidance and feedback on the thesis and helped in developing the statistical analyses used in this study as well in running some of the analyses in S-Plus. I am very grateful for his patience and availability.

I would also like to thank Dr. Abdissa Negassa for his assistance with the conception of some of the statistical models for correlated data used in this study and for his comments on the manuscript.

I feel privileged to have had the opportunity to work with the students and team in the Division of Cancer Epidemiology. Dr. Javier Pintos has been a long and cherished colleague. I have had the pleasure of working with him on several projects during my Master's and Doctoral training, including on head and neck cancer and on cancers of the cervix and vulva. Marie-Claude Rousseau, Harriet Richardson, Anita Koushik and Susan Brogly provided helpful feedback and suggestions in the preparation of this project. I have enjoyed sharing the PhD experience with them. Candida Pizzolongo, our Division secretary, was very helpful in relieving me of many administrative tasks. I am also grateful to Dr. Eliane Duarte-Franco for her help with my learning of the pathology of the disease and for making me feel welcome in the division.

I could not have completed this thesis without the encouragement of my family. I am grateful for their unconditional support and interest in my endeavors. Finally, special thanks to Concetta, for her encouragement and constant support during the moments spent analyzing and writing. She revived my self-confidence in times of doubt and was always ready to help from beginning to end.

Abbreviations used in thesis

AG, Anderson-Gill; ASCUS, atypical cell of undetermined significance; *bp*, base pair; CDC, Centers for Disease Control; CI, confidence interval; CIN, cervical intraepithelial neoplasia; CIS, carcinoma in situ; df, degrees of freedom; ELISA, Enzyme-Linked Immunosorbent Assay; FDA, Food and Drug Agency; GEE, generalized estimating equation; HC, Hybrid Capture™ ; HLA, human leukocyte antigen; HSIL, high-grade squamous intraepithelial lesion; HPV, human papillomavirus; HR, hazard ratio; IARC, International Agency for Research on Cancer; LCR, long control region; LSIL, low-grade squamous intraepithelial lesion; MHC, major histocompatibility complex; NBCCEDP, National Breast and Cervical Cancer Early Detection Program; NCI, National Cancer Institute; OR, odds ration; Pap, Papanicolaou; PCR, polymerase chain reaction; PWP, Prentice-William-Peterson; RFLP, restriction fragment length polymorphism; RR, rate ratio; SCC, squamous cell carcinoma; SD, standard deviation; SIL, squamous intraepithelial lesions; SE, standard error; WHO, World Health Organization; WLW, Wei-Lin-Weisfeld

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

Introduction: Human papillomavirus (HPV) infection is now believed to be the central cause of cervical cancer. However, most of the epidemiological evidence has come from retrospective, case-control studies, which do not provide information on the dynamics of a cumulative or persistent HPV infection.

Objectives: 1) To measure the risk of incident neoplastic cervical lesions over time related to prior cumulative and persistent HPV infections. 2) To evaluate the influence of HPV viral burden on lesion risk longitudinally. 3) To estimate the progression rates and sojourn time for precursor squamous intraepithelial lesions (SILs) and how they relate to HPV infection status.

Design and Methods: In 1993, the Ludwig-McGill study team began a large longitudinal study of the natural history of HPV infection and cervical neoplasia in the city of São Paulo, Brazil. Follow-up involved repeated measurements on individual subjects over time. 2462 women were enrolled into the study and were seen every 4 months in the first year (0, 4, 8 and 12 months), and twice yearly thereafter for a period of up to eight years. In addition to obtaining risk factor information via questionnaire, cervical specimens were taken for Pap cytology and HPV testing at every visit. Statistical analyses entailed: 1) using different modalities for defining HPV persistence by type and intensity; 2) using modeling approaches that take into account the repeated measurements of HPV and SIL over time within individuals; 3) analyzing changes in transition states between different cervical lesion grades and the rate of progression from one state to the next.

Rationale: A longitudinal, repeated measurement cohort investigation, such as this one, permits an accurate and unbiased assessment of the relationship between cumulative HPV exposure and lesion incidence. An elevated relationship between persistent HPV infections and SIL incidence supports the proposal for the application of type-specific molecular HPV DNA testing as a screening tool for the detection of cervical neoplasia. Better understanding of the natural history of disease can help in developing effective and efficient public health programs in prevention for cervical cancer.

RESUME

Introduction: Le virus du papillome humain (VPH) est considéré comme la cause principale du cancer du col utérin. Néanmoins l'évidence épidemiologique, montrant cette association provient des études rétrospectives de cas-contrôles, qui ne donnent pas d'informations sur le dynamisme d'une infection cumulative persistante.

Objectives: 1) Mesurer le risque d'une lésion néoplastique cervicale incidente dûe á une infection de VPH persistanteantérieure. 2) Evaluer l'influence du taux de virus sur le risque longitudinal d'une lésion néoplastique. 3) Estimer la vitesse de progression et la durée clinique des lésions préinvasives.

Designs et Méthodes: En 1993, l'équipe de l'étude de Ludwig-McGill a commencé á São Paolo, Brésil, une étude de cohorte de l'évolution historique du VPH et du cancer du col utérin. L'étude comprend des dépistages répétés. 2462 femmes ont participé á cette étude. Les participantes ont été examinées tous les quatre mois pendant la première année, et tous les six mois après cela pour une période maximale de huit ans. Un questionnaire, un test de VPH et un prélèvement de cellules du col utérin pour une évaluation cytologique ont été effectués à chaque visite. Les analyses statistiques comprenaient: 1) différentes modalités de définition d'une infection persistante de VPH; 2) une implication de différentes approches statistiques tenant compte de prélèvements répétés au cours de l'étude; 3) l'analyse des transitions entre les grades neoplastiques des lésions cervicales.

Rational: Une étude de cohorte longitudinale avec prélèvements répétés permet ainsi de mesurer précisement et correctement l'association entre le VPH et le risque du cancer du col utérin. Une association elevée entre une infection persistante et l'incidence des lésions neoplasticques confirment l'utilité des tests du VPH comme méthode de dépistage du cancer. Une meilleure compréhension de l'evolution naturelle des lésions cervicales peut aussi aider á établir des programmes de santé publique visant la prévention de cancer du col utérin.

1. INTRODUCTION

Cervical cancer and its precursors can follow one of two histological lineages depending upon whether they originate in squamous or in glandular cervical epithelium. The former type, squamous cell carcinomas (SCC) make up 80%-90% of all cervical cancer cases, glandular adenocarcinomas make up the rest. Age standardized global rates of adenocarcinoma and other carcinomas between 1973-1991 were 11.4 and 12.7, respectively [*253*]. Given the differences between these two diseases in terms of their etiology, and approaches to their detection and prevention, epidemiological investigations tend to evaluate their natural history separately [reviewed in *71*]. The focus of this thesis is on the natural history of preinvasive squamous lesions of the uterine cervix.

The World Health Organization's (WHO) International Agency for Research on Cancer (IARC) released in 1995 a review of epidemiologic and biologic evidence demonstrating human papillomavirus (HPV) infection as the central cause of cervical cancer [*110*]. Asymptomatic, latent genital HPV infections are found in 5%-40% of women of reproductive age with most women who engage in sexual activity acquiring an HPV infection sometime in their lifetime. Most of these infections will be transient with only a small proportion initiating a cancerous lesion on the cervix [*73;75;103;105;106;139;174;188;244*]. It is hypothesized that without sufficient immunity or with repeated exposure, infections can become persistent and more severe [*106;107;171*]. While the role of HPV in the natural history of cervical intraepithelial neoplasia (CIN) is established, HPV is considered as a necessary but not sufficient cause of cervical cancer. Establishment of productive and persistent infections by oncogenic types of HPV is thought to be a key early event in the natural history of cervical cancer [*15*].

A consistent relationship between HPV infection and the incidence of cervical cancer has been identified by case-control studies [*110*], and confirmed more recently by longitudinal studies [*99;106;107;131;138;171;174;185;224;255;262;268*]. Although an HPV infection is said to produce CIN grade 1, consensus on using detection of HPV as a predictor of risk of higher grade disease has been lacking since most HPV infections will not lead to cervical cancer [*66;202*]. Other factors, including variations in defining what constitutes a prognostically significant infection are starting to be investigated to improve prediction of risk for the disease. To understand the role of HPV and the pattern of the dynamic changes in the natural history of cervical neoplasia one must conduct studies that collect data repeatedly on both risk factors, like HPV, and on outcomes (i.e., cervical lesions) on multiple occasions.

When multiple measurements of the virus cannot be made, a single measure of viral burden may be sufficient to identify a subset of HPV infected women at higher risk of developing cervical

cancer. Given the positive relationship between viral load and the likelihood of persistent HPV infection, and the strong relationship between the latter and the risk of cervical neoplasia [*42*], a single measurement of viral load in cervical specimens may be a suitable bio-marker for the development of CIN. Measurement of viral burden using quantitative methods has been under recent investigation, though preliminary results have been controversial [*114;248*].

The time dependence of the association between HPV and the incidence of precursor lesions of cervical cancer is still a matter of debate [*16;64*]. The cytopathological precursor stages of cervical neoplasia, squamous intraepithelial lesions (SIL), can occur repeatedly over time. The likelihood of a lesion persisting or progressing onto more advanced grades or to cancer may depend on the characteristics of HPV infections [*142;224;248;252*] as well as on other co-factors [*16;63*]. The strength of the observed association is influenced by the accuracy with which the exposures are measured [*65;72*]. To date, cohort studies of HPV and CIN have differed in the frequency with which subjects have been re-evaluated over time for exposure and outcome status [*106;107;131;138;171;174;185;255;262*], as well as in the time between the repeated measurements [*99;268*].

In the natural history of cervical cancer, women can progress from a normal state where no neoplastic or pre-neoplastic changes are detected in the squamous epithelium, to varying states of cellular abnormalities in the cervical epithelium including carcinoma *in situ* (CIS) [*154*]. Women may develop SILs of low and high grade and progress on to CIS or regress back to a normal state [*109*]. While some research on the rates of progression and regression of cervical neoplasia has been done [*109;152;155;162;164;187*], to date no study has evaluated sojourn time (preclinical duration of preinvasive neoplasia), or regression and progression of precursor lesions according to HPV infection status. Identifying those lesions that are at high risk of progressing quickly to more severe stages can provide potential targets for detection by screening and subsequent aggressive follow-up, chemopreventive therapy or vaccination.

1.1. Objectives

The overall aim of this thesis was to investigate the time dependent relationship between HPV infection and cervical neoplasia incidence. This was done through an extensive analysis of longitudinal data from an ongoing prospective cohort study in São Paulo, Brazil. The specific objectives were:

1. To measure the risk and cumulative incidence of SILs for women with persistent HPV infections determined by repeated samplings over time from cervical specimens (Paper I);

2. To evaluate the predictive role of HPV viral burden on SIL incidence among HPV positive subjects using a true quantification method of viral load (Paper II);
3. To evaluate how changes in study design can affect the observed relationship between HPV infection and cervical lesion risk longitudinally. This involved modeling approaches that take into account the correlated nature of the repeated measurements within individuals (Paper III);
4. To estimate the progression, regression rates, and sojourn time of precursor lesions of cervical cancer according to HPV infection status. This involved actuarial analysis of transition rates between different lesion grades (Paper IV);

2. LITERATURE REVIEW

2.1. Burden of disease

Cervical cancer is the second most common neoplastic disease affecting women today. It comprises approximately 12% of all cancers prevalent in women worldwide [*159*]. In the year 2000, 468,000 new cases of cervical cancer were detected, representing 4.7% of all new cancers for both men and women combined. Cervical cancer is less frequent in industrialized countries, and almost 80% of all cases occur in developing countries, where it is second only to breast cancer [*168*]. In many developing regions, it is the most common cancer in women. Age-adjusted incidence rates (standardized according to the world population of 1960) for São Paulo County in Brazil were 27.4 per 100,000 women-years for 1993 [*160*]. This rate was similar to the overall rate for Brazil (23.8 per 100,000 women-years) [*193*].

In 2000, an estimated 233,000 deaths occurred from cervical cancer worldwide, representing 8.7% of all cancer deaths. In developing countries, where screening for cervical cancer is limited, the average 5-year survival rate following a diagnosis of invasive cervical cancer is 49% [*193*]. The North American perspective is brighter. Although incidence of cancer in Canada averages around 1500 cases per year, mortality is much lower compared to developing countries. [*159*].

The prognosis of cervical neoplasia is good when the disease is detected and treated at an early preinvasive stage. However, preinvasive lesions are asymptomatic and are generally discovered through cytological examination using the Papanicolaou technique (or Pap test), or through visual inspection methods such as cervicography performed from a photographic image of the cervix taken during a gynecologic examination. Their presence is confirmed via magnification by colposcopic examination and biopsy. These follow-up examinations themselves can be rather invasive [*71*].

For every 50-100 cases of abnormal smears identified by Pap cytology screening consistent with a precursor lesion grade, one new case of invasive cancer is found [*71*]. In the clinical management of abnormal cytological smears, cases of high-grade squamous intraepithelial lesions (HSIL) are immediately referred for colposcopy and biopsy because of their high likelihood of progressing to cancer. Clinical follow-up and treatment can involve invasive colposcopic examination, biopsy or surgery. Altogether, atypical and SIL findings account for over 10% of all Pap smears that are processed in screening programs [*32*].

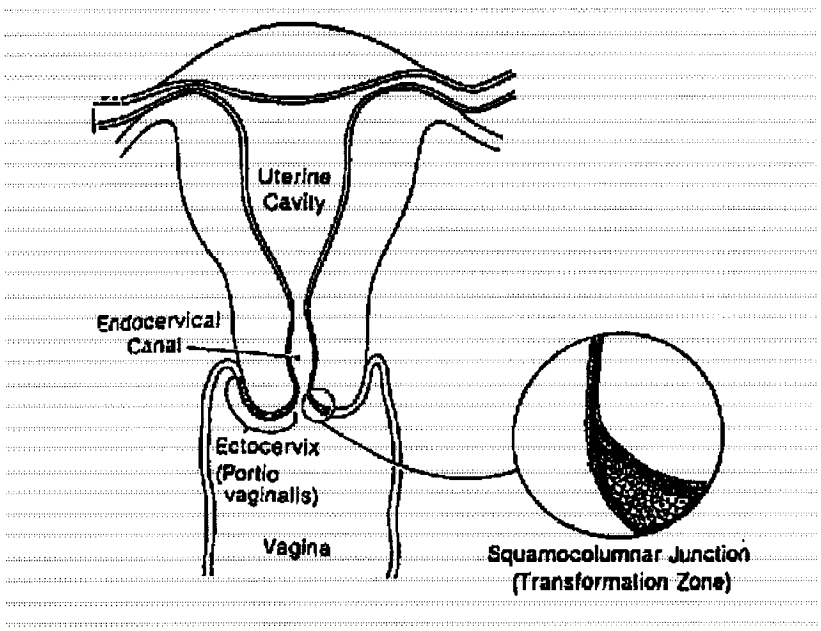
2.2. Natural history and etiology of cervical neoplasia

Over the past 30 years, epidemiological research has been consistent in providing evidence of a progressive disease model in cervical neoplasia, with similar risk factors identified for both

cervical cancer and its precursor lesions. The similarity of risk factor profiles observed in different studies [*18*] has lent credibility to the natural history model specifying that the abnormal changes seen in the cervical epithelium follow a continuum leading to invasive cervical carcinoma. At each stage of the disease, there is a likelihood of spontaneous regression or remission without medical intervention.

At the entrance to the endocervical canal, adjacent to the squamo-columnar junction, epithelial cells in the transformation zone of the uterine cervix are in a constant state of maturation and replacement. The transformation zone of the squamo-columnar junction and uterine cervix are illustrated in Figure L-1. In the normal maturation process, these cells go through several transitions from basal to columnar to squamous form and eventually migrate to the ectocervix before shedding off [*154*]. The natural history of cervical cancer begins as a slow process of disruption of the normal maturation of the transformation zone. Initially, this process of abnormal changes is limited to the cervical epithelium and does not involve the adjacent connective tissue. This preinvasive phase is known as dysplasia or an intraepithelial neoplasia.

Figure L-1: The squamo-columnar junction and transformation zone of the uterine cervix



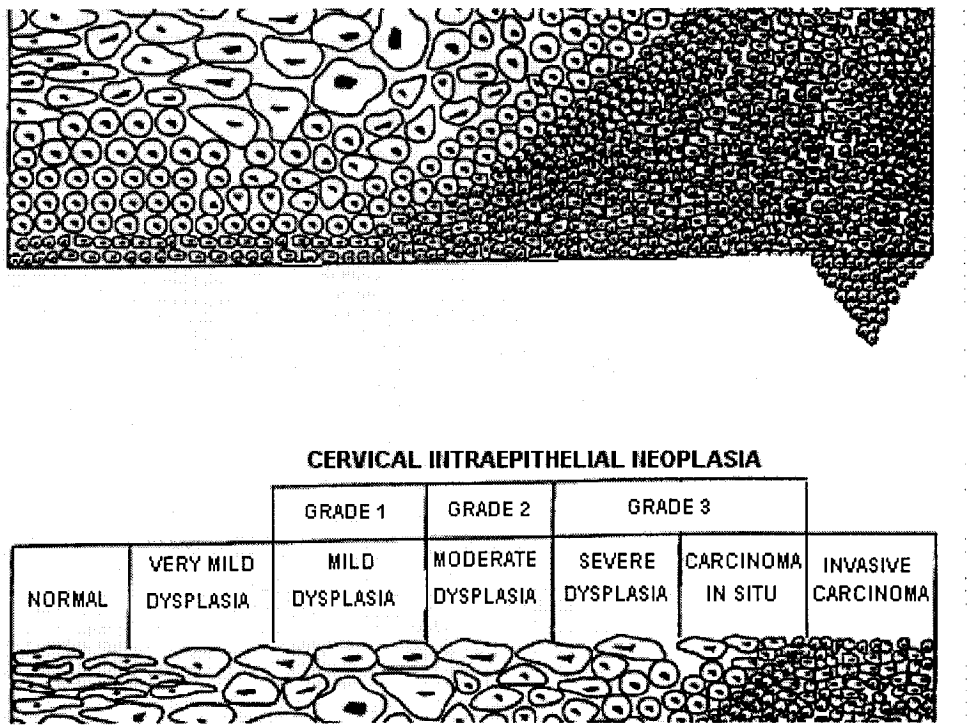
* Reprinted from Ylitalo [*265*].

Cells exhibiting morphological changes characteristic of dysplasia usually arise in the basal layer and spread up towards the squamous layer. Mild dysplasias are commonly identified by the presence of enlarged, irregular nuclei with perinuclear halos in the superficial or intermediate cells of the squamous epithelium. Otherwise known as koilocytotic atypia, such atypical changes are

characteristic of an HPV infection and are classified as CIN 1 under the WHO classification system. In North America this is classified as a low-grade squamous intraepithelial lesion (LSIL) under the Bethesda system [*231*]. Figure L-2 illustrates the principle of cytological detection based on the clinical progression of the disease.

Moderate and severe dysplasias contain malignant basal/parabasal cells of different numbers. In these cells the cytoplasm is scarce and the nuclei are large, hyperchromatic, and irregular producing a nuclear/cytoplasmic ratio in favor of the nuclei. If more than a third of the squamous epithelium shows inclusion by these dysplastic cells, the lesion is said to be moderately dysplastic and is classified as CIN 2. When more than two thirds are involved, the lesion becomes CIN 3 or severe dysplasia [*59*]. If left untreated, moderate and severe dysplasia can spread the width of the squamous epithelium and become CIS. Subsequently, the disease may break through the basal lamina and infiltrate the underlying connective tissue becoming invasive cancer. Once the cancer has penetrated the stroma, the lesion can grow and spread through the circulatory or lymphatic system to other sites in the body and become metastatic.

Figure L-2: Clinical progression of cervical intraepithelial neoplasia



* Modified from Meijer et al. [*154*].

2.2.1. Cytology classification

Designed to detect morphological changes in epithelial cells of the transformation zone, cytology currently represents the most accepted method of lesion detection in clinical management of

cervical cancer. Since the Pap smear was implemented widely for cervical cytologic screening, incidence and mortality from cervical cancer have markedly declined in many countries with screening [*159*]. Cytologic screening has been used mainly to detect preinvasive, SILs before the development of cervical cancer. A substantial proportion (5%-20%) of the Pap smears read in cervical cancer screening programs are classified as minor grade lesions or atypias of undetermined significance [*71*].

Table L-1 shows the correspondence between reporting terminologies for cervical cytology and pathology reports according to different international classification systems for cervical cytopathology. A number of classification systems have arisen over the years expanding on the original scheme by Papanicolaou [*190;203;207;208*]. In 1988 a US National Cancer Institute (NCI) led workshop proposed the Bethesda system [*231*] and later revision [*184*] to try and resolve the disparity between the different classification schemes.

The Papanicolaou system was developed as a grading scheme for cytological evaluations. Unlike histological analysis, it is not considered a diagnostic tool. Cytology, however, does not require excision of cervical tissue foir evaluation as occurs with biopsies [*184;190*]. According to the Bethesda system, combined categories were created for mild dysplasia or LSIL and for moderate to severe dysplasia (represented as a combined HSIL category), based on the understanding that morphological changes in the cervical epithelium should reflect the etiological effect of HPV infection [*71*]. HPV-associated changes in the absence of other squamous abnormalities are classified as LSIL and more advanced degrees of dysplasia and CIS (corresponding to histological CIN grades 2 and 3) are combined into HSIL as a single lesion grade. This classification system, however, combines koilocytotic atypia and flat condylomas, pathological signs of a productive HPV infection, with mild dysplasias and dyskaryosis (histological CIN grade 1) which represent cytopathic results of a present and productive HPV infection. A more detailed form of the Bethesda system is sometimes used in epidemiological studies that separates LSILs into those that represent koilocytotic atypia or show effects of a productive HPV infection (LSIL/HPV) and those with mild dysplasia or squamous abnormalities (LSIL/SQ) equivalent to CIN 1 [*168*]. HSILs are also subdivided into CIN 2 and CIN 3.

The Bethesda system also created a category for equivocal atypias: atypical squamous cells of undetermined significance (ASCUS). This classification has no etiological basis in the disease continuum [*231*]. The aim was to resolve the ambiguity of the class II category of the Papanicolaou scheme. ASCUS readings can represent from 3 to 7% of all Pap smears tested [*184*]. In the latest revision of the Bethesda classification system, a separate sub-category, ASC-H was created for ASCUS smears where the possibility of an HSIL cannot be excluded [*232*].

While this latter sub-category is believed to include 5 to 10% of all ASCUS cases, it is not highly reliable. In this cohort study, a similar approach was adopted introducing two ASCUS sub-categories: ASCUS favor benign atypia and ASCUS rule out SIL to distinguish the putative severity between these two subcategories.

Table L-1: Disease classification according to different international guidelines

Papanicolaou class system	Dysplasia terminology	Original CIN terminology	Bethesda system (SIL terminology)
I	Normal	Normal	Within normal limits
II	Atypia (multiple qualifiers)		Benign cellular changes (infection or repair)
II	Atypia (epithelial cell abnormalities)		ASCUS/AGCUS with qualifier *
II or III		Koilocytotic atypia, flat condyloma, without epithelial changes	LSIL
III	Mild dysplasia or dyskaryosis	CIN grade 1	LSIL
III or IV	Moderate dysplasia or dyskaryosis	CIN grade 2	HSIL
IV	Severe dysplasia or dyskaryosis	CIN grade 3	HSIL
IV or V	Carcinoma in situ	CIN grade 3	HSIL
V	Invasive carcinoma	Invasive	Invasive carcinoma

* Whether a reactive or premalignant/malignant process is favored. Reproduced from Franco and Ferenczy [*71*].

2.2.2. Performance of cytology in research

Pap cytology was introduced as a screening tool in the 1960's and was designed to identify cervix cancer precursors [*191*]. The rationale for the use of cytology in a research setting includes the arguments in favor of using it as a screening tool (i.e. its ease of use, reproducibility, cost and relative non-invasiveness). It also has the advantage of not interfering in the normal maturation process of the epithelium and natural history of early lesions by removing only the cells at the surface of the cervical epithelium and entrance to the endocervical canal.

Pap cytology involves the sampling of epithelial cells from around the endocervical canal of the cervix, transfer of these sampled cells on to a glass slide, the fixing and staining of these cells to the slide and subsequent visualization of the cervical cells under 10x magnification by a trained cytopathologist. The methods for sampling cervical cells for Pap smears that exist in clinical practice include the cytobrush, Accelon brush (no longer commercially available), the exocervical spatula and endocervical (Dacron) swab, and the cervical broom. While all of these methods are designed to sample both ecto-cervical and endo-cervical cells, there may be differences in the quality of the slides that can interfere with diagnostic performance [*79;249*]. More recent tools like

the cytobrush and Accelon brush were designed to respond to these concerns by improving the method of transfer of collected cells onto the cervical slide to facilitate visualization.

Pap cytology is used almost exclusively in medical practice worldwide. As a result cytology is the outcome event *per se* when referring to screening studies and is an important clinical outcome for patient management and treatment. Screening performance of Pap cytology observed in different studies ranges from 60 to 80% sensitivity with a specificity of 70 to 100% [*150*]. Despite its demonstrated efficacy as a screening test, false-negative results do occur for a number of reasons. Reported false-negative results in the literature vary widely, from 1% to 90%, with the most comprehensive studies showing results between 20% and 30% [*77*]. Variability in interpretation (depending of the severity of the lesion) of abnormal Pap smears may contribute to the wide variation in false-negative results [*54*]. In this regard, the Pap test works best when lesions have progressed further than desirable. However, much of the evidence has come from passive studies of cervical cancer screening without correction for verification bias. A recent meta-analysis of studies with bias-free measures of sensitivity indicated an overall performance of 51% sensitivity and 98% specificity [*182*]. While variability exists between studies as a result of population characteristics, sampling and evaluation methods are also suspected to influence test performance [*71;159*].

While biopsies represent the gold standard for diagnosis of cervical neoplasia, it cannot be used as a screening tool for ethical reasons or to study the natural history of SCC. The procedure is invasive and requires the removal of one or more samples of tissue spanning the depth of the epithelium. Targeted biopsy performed at colposcopy may therefore remove the lesion entirely and terminate the progression of the disease thus interfering in the natural history of cervical neoplasia. Cone biopsy procedures can remove the transformation zone entirely [*168*]. Cytology therefore, serves both as an ethical alternative for screening and follow-up of women, and a viable method of outcome ascertainment to study the natural history of cervical cancer.

2.2.2.1. Alternative cytology methods

Recently, a method of preparation and transfer of cells to the Pap smear slide has been introduced. Thin-layer liquid-based cytology involves the suspension of sampled cells in a stabilizing solution before transfer of the cells on to slides through a filter designed to remove any extra-cellular material [*58*]. This technology is thought to improve Pap cytology screening by producing uniformly cleaner slides [*6*]. This method has undergone limited testing in epidemiologic studies, and large-scale prospective studies using this method are needed to estimate the test's diagnostic performance [*69*].

Methods involving computer-assisted image analysis of Pap smear slides have also been proposed to reduce the level of variability (intra- and inter-observer agreement) observed between cytopathology laboratories [*162*]. Automated screening technologies may also help reduce the subjectivity inherent in Pap smear screening [reviewed in *61*]. These technologies, however, have been used to a limited extent due to the elevated costs required to set up and run them and are feasible only in laboratories processing large volumes of slides.

2.2.2.2. Visualization methods

While cytological analysis remains by far the predominant method used for screening for cervical carcinoma and its precursor stages, alternate technologies based on methods of visualizing the cervix have also been developed. The most familiar in this category are: direct visual inspection, cervicography, speculoscopy and colposcopy. A potential problem common to all of these methods however, is that with increasing age, the squamo-columnar junction migrates inward towards the endocervical canal. As a result lesions forming adjacent to the junction become progressively more difficult to detect. The scoring scheme used by many of these methods must accommodate this obstacle to visualization [an example of medical reports sheets can been found at http://www.nationaltestinglaboratories.com].

Direct visual inspection or cervicoscopy relies on detecting morphological changes conducive of neoplasia revealed by the application of 5% acetic acid. Variations of this test involve treatment of the cervix with Lugol's iodine, also known as Schiller's test, and the use of a hand-held magnification device. Direct visual inspection can be performed by non-medically qualified screeners. Some studies have found direct visual inspection to have a lower specificity than cytology due to the use of aceto-whitening as the only criteria of positivity in almost all of the studies [*51;227;246*], as well as the presence of cervico-vaginal co-infections [*51*].

Adaptations on direct visual inspection include cervicography and speculoscopy. Cervicography consists of taking a static photographic image (Cervigram® Slide, National Testing Laboratories, MI) of the ectocervix after application of 5% acetic acid solution [*235*]. Speculoscopy uses a chemiluminescent blue-white light called a Speculite® (Trylon Corporation, CA) attached to the upper dilator blade of the speculum [reviewed in *163*]. Such methods have been suggested to play a role in screening in remote locations where well-trained colposcopists cannot be recruited easily [*98*]. In a large population-based study of 8,460 women, cervicography performed poorer than Pap cytology with a lower sensitivity for detecting HSIL and only slightly higher for cancers [*227*].

2.2.2.3. *Colposcopy*

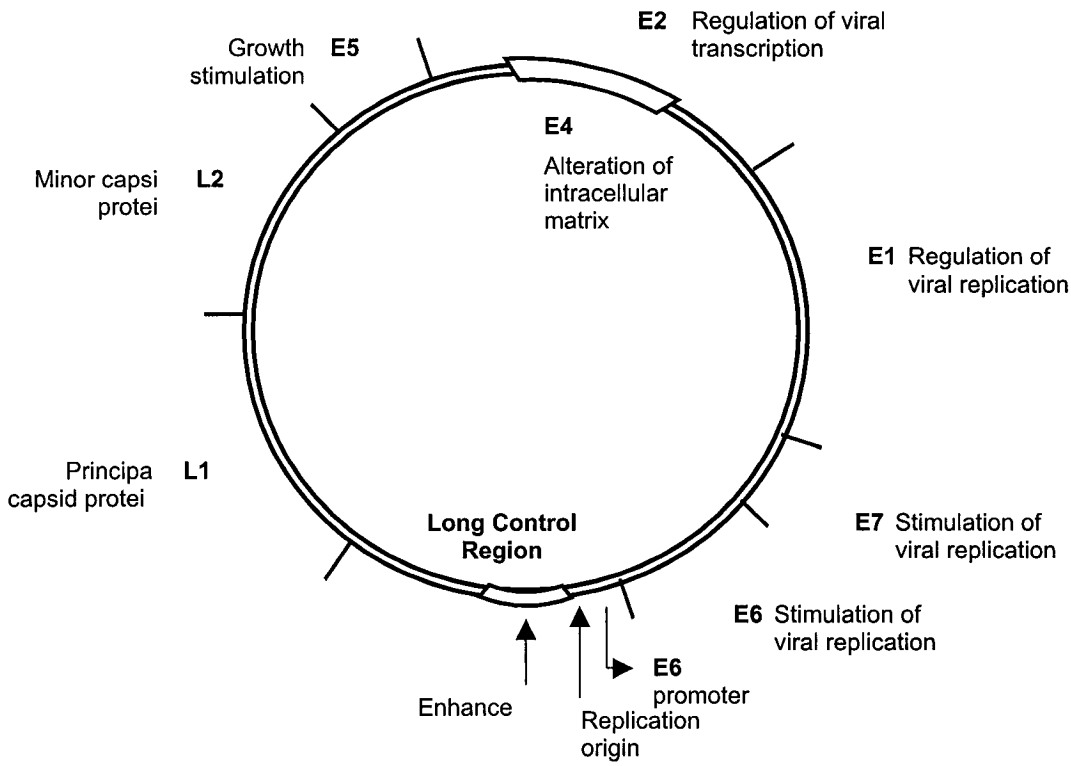
Colposcopy is normally employed to verify abnormal findings discovered by Pap cytology or cervicography and as a method to help direct biopsies for confirmation of screening results. Similar to direct visual inspection, colposcopy is performed under magnification by a medically trained health provider. A more invasive procedure providing the patient with considerably more discomfort, colposcopy is used to identify lesion plaques on the cervix for directed biopsies. The method also requires extensive training and more experience than other visualization technologies or cytopathology. Evaluation of colposcopy as a triage tool for cervical cancer is currently under way [*218*]. In a meta-analysis of colposcopy performance for the diagnosis of SIL, Mitchell et al., [*163*] found an average weighted sensitivity and specificity of 96% and 48%, though the discrimination of LSIL from HSIL and cancer was weaker, 85% and 69%, respectively.

2.3. Role of human papillomavirus

Pisani et al. [*199*] estimated that 15.6% (1,450,000 cases) of the worldwide incidence of cancer in 1990 can be attributed to infection with either the hepatitis B and C viruses, the human papillomaviruses, EBV, human T- cell lymphotrophic virus I, HIV, the bacterium Helicobacter pylori, schistosomes, or liver flukes. The presence of a venereal agent in the natural history of cervical neoplasia was identified early on in epidemiological studies seeking a causal agent of cervical cancer [*21;263*]. Confirmation of the role of HPV in the natural history of cervical neoplasia arose with the advent of polymerase chain reaction (PCR) based techniques to detect the presence of HPV DNA sequences in cytological specimens. Over 95% of cervical tumor specimens have been found to harbor HPV DNA, of which HPV types 16 and 18 account for 50% and 14% of the infections, respectively [*17;254*]. Over the past decade, there has been strong evidence supporting a causal relationship between HPV and cervical cancer [*16;110;178;220*].

HPVs are small, non-enveloped double-stranded DNA virus particles, of approximately 55 nm in diameter. The HPV genome is circular, 8,000 base pairs in length and is encapsulated in an icosahedral protein capsid containing 72 capsomers [*198*]. Taxonomically, papillomaviruses are part of the Papovaviridae family, with HPV being highly specific to their human hosts. The HPV genome consists is divided into three sections: the early (E) region involved primarily in growth stimulation and replication control, the late (L) region the codes for capsid formation, and the noncoding long control region (LCR) (Figure L-3) [*110*]. Different HPVs are classified as types on the basis of DNA sequence homology in the E6, E7, and L1 genes.

Figure L-3: The HPV 16 genome map with functional characteristics



Although the majority of women of reproductive age acquire an HPV infection sometime in their lifetime [*110*], most of these infections will be transient, with only a small proportion becoming persistent [*72;73;105;106;139;174;244*]. Without a sufficiently strong immune response or with repeated exposure, infections can become persistent and more severe. Studies have observed a substantial increase in risk of cervical neoplasia for women who develop persistent, long-term infections with oncogenic HPV types [*131;171*].

2.3.1. Detection of HPV

Over 120 different HPV genotypes have been isolated of which 40 are identified as infecting the genital mucosa [*49;50;204;273*]. Two basic methods for testing the presence of HPV DNA in cervical specimens exist today. These are: Hybrid capture™ HPV DNA assays and PCR-based methods.

2.3.1.1. Non-amplified DNA assays

Initial detection and typing of HPVs employed non-amplified DNA samples and RNA hybridizing techniques [*143*]. The HybridCapture™ tube assay (HC, Digene Diagnostics, Gaithersberg, MD) was the first US FDA approved test that detected 9 types of oncogenic HPV (16, 18, 31, 33, 35, 45, 51, and 52) and 2 non-oncogenic types (6, 11) using separate probe mixtures. HC uses RNA

probes complimentary to the genomic sequence to form hybrids that are captured by antibodies and subsequently detected. The second-generation test, HC2, includes a further 4 oncogenic types (39, 58, 59, 68) and 3 non-oncogenic types (42, 43, 44). Other detection methods, including ViraPap™ [*28;88;173*] and HPVProfile™ [*124*], have also been used involving a similar method of RNA-DNA dot blot hybridization. ViraPap has since been replaced by HC. Typing, however, is usually not possible with these methods as the detection method employs a cocktail of probes that prevent differentiation among types. An additional limitation, the threshold level for detection is high with HC to permit clear expression of a positive test, resulting in a high false negative rate for low viral load (in number of viral copies) specimens.

2.3.1.2. PCR-based methods

PCR-based methods rely on target amplification using generic and specific primers to achieve higher sensitivity. The sensitivity and specificity of such methods vary depending on size of the PCR product used, the primer set, reaction conditions and the performance of the DNA polymerase reaction. Transport and storage of sample can also affect the sensitivity. While PCR is a more sensitive method and can replicate one million copies of a single-stranded DNA molecule after 30 cycles, care must be taken when processing to avoid cross-contamination of samples. PCR testing for HPV began with the introduction of the MY09/11 degenerate primer set [*149*]. This degenerate primer set yields a 450 base-pair amplicon that can detect over 40 types of HPV. The system incorporates one primer set designed to amplify a highly conserved L1 domain and a second primer set designed to amplify a domain within the E6 gene. An alternative primer set, the GP5/6 consensus primer set, was developed soon after [*247*], and was then expanded with the introduction of the GP5+/6+ primer set [*45*] which can detect 37 mucosotropic HPV types. The PCR-based methods using HPV specific degenerate primers have been shown to be more reliable and sensitive than consensus primers [*5;84;105;201*]. A more recent set developed by Gravitt and Manos [*85*] (PGMY09/11) has been demonstrated to be more sensitive, specific and efficient than MY09/11 [*130*]. Based on a highly conservative section of the L1 region of the HPV genome, these PCR-based methods can amplify many HPV types, which can then be typed for research purposes.

Resulting amplification products are then either typed with type-specific oligonucleotide probes or analyzed with restriction fragment length polymorphism (RFLP). Visualization of presence of types is generally performed by gel electrophoresis, but can be adapted to different formats, including dot or line blot hybridization. Amenable for use with non-amplified samples, Southern blot hybridization was one of the first methods used to detect and type for HPV DNA. Hybridization is performed with radioactive probes. The size of fragments and stringency of hybridization are used to identify HPV types [*110*]. In a study comparing several HPV

hybridization methods with Southern blot, Schiffman [*219*] found PCR-based methods to be more sensitive than Southern blot alone.

When considering concurrent multiple infections by HPV types, HC2 does not discriminate between different types of HPV while PCR-based methods can. Due to the fact that different primers detect different HPV types in PCR-based methods, their sensitivity varies when detecting multiple infections [*111;129*]. Though research on the relationship between HPV co-infections and cervical neoplasia is limited, results from a case-control study have shown an increased risk association for multiple infections [*213*] Among oncogenic HPV types, HPV 16 and 18 have been identified in the majority of cervical tumors and precursor lesions [*41;122;153;156*] and are targeted in clinical vaccine trials [*176*].

2.3.1.3. Serology

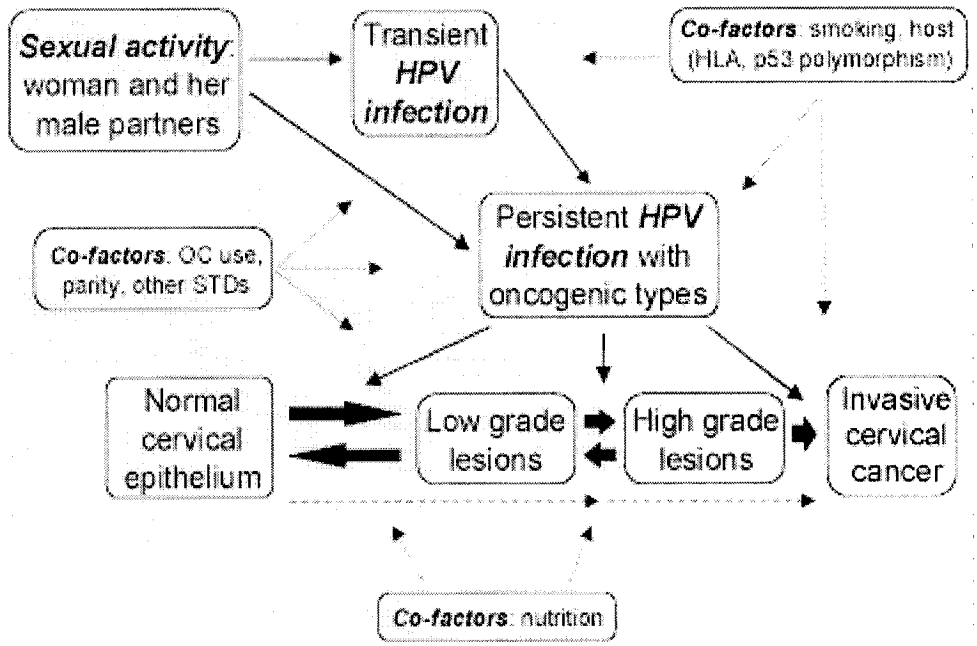
Other testing methods have looked at assessing immune response in women by using serology to measure humoral antibody response to HPV. Enzyme-Linked Immunosorbent Assays (ELISA) are used with virus-like particles as antigens to detect the level of response. Studies have shown that there is a positive association between the detection of HPV antibodies and risk of cervical disease [*46;137;239*]. This approach, however, is limited in that the presence of antibodies may reflect either current or past infections. Some studies have associated IgG response to HPV 16 capsids with a lifetime exposure to HPV, while IgA response is thought be a marker of recent or ongoing infection [*31;256*]. Nonetheless, measurement of changes in HPV infection status by repeated sampling is required to assess temporal associations for risk of cervical neoplasia with HPV infection status, which cannot be evaluated through presence of these antibodies.

2.4. Other risk factors

Other risk factors have been associated with the natural history of cervical cancer, and act as mediators or remote variables in the causal pathway from HPV infection to cervical cancer. That is, most of the identified co-factors for cervical cancer do not confound the association between HPV and cervical neoplasia per se, but can instead act further downstream or upstream to the infection. For example, once one is detected with the virus, the role of behavioral co-factors like sexual activity disappears [*67*]. These factors include both exogenous and endogenous factors that, acting in conjunction with HPV, influence the acquisition and persistence of the virus and progression of precursor lesions of cervical neoplasia to cancer. Others are believed to act directly on later carcinogenic steps and may behave as confounders only through indirect correlation with acquisition of HPV infections (e.g. diet and reproductive health). Figure L-4 shows the current sequence of disease progression believed to exist in cervical cancer and the established co-factors involved. Co-factors can be classified as behavioral, viral or host-related.

Though the focus of this dissertation lies with the primary cause of cervical neoplasia, HPV infections, I review below the literature on these co-factors to illustrate how they may fall in the natural history of the disease.

Figure L-4: Current understanding of the natural history of cervical intraepithelial neoplasia



*Adapted from framework presented in Franco et al. [*69*].

2.4.1. Behavioral co-factors

While acquisition of the virus may occur primarily through sexual contact and intercourse, other reproductive health factors such as parity [*22;23;179*], use of oral contraceptives (OC) [*167*] and other sexually transmitted diseases (STD) – HIV, HSV and chlamydia – can also affect HPV incidence and persistence [*11;175;250*]. Other co-factors including: i) tobacco smoking [*134*]; ii) diet (including vitamin A, E, C, beta-carotene and consumption of fruits and vegetables) [*132*] have been found to affect the progression of the disease.

2.4.1.1. Sexual activity, parity, OC use and STDs

Early evidence from numerous studies has pointed to markers of sexual history as the most prominent behavioral risk factors in the epidemiology of cervical cancer. These markers of sexual behavior are considered surrogate measures of the sexually transmitted pathogen HPV [*168*]. In particular, two measures of sexual activity, number of sexual partners and age at first intercourse have shown strong associations [*20;74*], as well as the sexual behavior of the woman's male partners [*24*]. The consistency of the sexually-transmitted disease model for cervical neoplasia

led much of the laboratory and epidemiologic research designed to identify the putative microbial agent or agents acting as the intermediate cause of cervical cancer [*71*]. Mediation analysis has demonstrated that apparent associations between incidence of cervical lesions and markers of sexual activity are explained by the presence of HPV in the epithelial cells of the cervical epithelium [*65;67;119*].

The association between parity and CIN is equally inconsistent and confounded by sexual activity. Estimates of risk from case-control and cohort studies have provided evidence of an association among women with multiple pregnancies compared to nulliparous women independent of HPV [*34;179*] while others failed to show a relationship [*91;113;186;192*]. Suggested biological mechanisms of risk effect have been proposed including cumulative traumatic or immunosuppressive effects from multiple pregnancies facilitating acquisition of HPV, and from the interaction between progesterone levels during pregnancy and HPV proteins [*195;225*]. Morphological changes that occur to the cervix after giving birth could also influence visual detection methods such as cervicography or colposcopy.

Use of oral contraceptives has also been investigated in several case-control studies [*34;167*] though the association was seen to disappear after accounting for HPV infection, sexual history, and frequency of cytological screening [*133;171*]. The suggestion is that OC use may play a role in the progression of pre-malignant lesions to malignant disease through hormonal interaction with the E6 and E7 genes of the HPV genome [*165*].

Initial studies of sexually transmitted vectors for cervical cancer identified HSV as the primary agent in cervical tumors [*151*]. These studies, however, found that this association was not consistent when HPV infection was controlled for because of a strong correlation between infections for the two viruses [*47;112*]. In addition to HSV, studies have also shown a positive association between Chlamydia infection and risk of cervical cancer [*47;131*].

2.4.1.2. Tobacco smoking and diet

Tobacco metabolites have been found in the cervical mucosa [*221*] while other studies indicate a link between smoking and immune response to HPV [*134;188*]. Smoking, however, is strongly correlated with sexual behavior making adjustment for confounding difficult. The apparent association between smoking and acquisition or persistence of HPV may therefore be a result of residual confounding due to incomplete exposure assessment [*71;223*].

Studies looking at vitamin A, C, and E in association with risk of CIN have produced contradictory findings. Some studies show a negative association or no association between beta-carotene,

vitamin A, C, and E with CIN risk [*100;132;230;272*], while others show a positive association between beta-carotene and risk of CIN [*48*]. Variations between these studies existed in the method collection of information on dietary factors including subjective assessment by questionnaire and measurement of plasma nutrient levels.

2.4.2. Host co-factors

Host factors include ancillary genetic factors such as human leukocyte antigens (HLA), pRb and p53 polymorphisms [*13;146;147*] as well as level of immunosuppression [*189*].

The E6 and E7 proteins of the HPV genome interact with human tumor-suppressor genes (p53 and pRb), leading to their inactivation and increased susceptibility to cancer. The likelihood of this inactivation may increase for women harboring persistent HPV infections with oncogenic types [*57*]. Results from various case-control studies that have investigated the relationship between p53 and CIN have been inconsistent [*95;102;115*]. Many of these studies, however, have been plagued by misclassification and variability in testing methods as shown by Makni et al. [*147*].

Certain polymorphisms for major histocompatibility complex (MHC) genes have been associated with genetic susceptibility to cervical cancer [*142;257;258*]. Many of the attributed polymorphisms are located in human leukocyte antigen (HLA) genes found in the MHC region that encodes for cell surface class I (HLA-A, B and C) and class II (HLA-DR, DQ and DP) molecules. Immunologic responses are mediated by HLA genes that can influence the acquisition and persistence of certain HPV infections [*145*]. Previous studies conducted in different populations have obtained disparate results regarding MHC polymorphisms and susceptibility for cervical cancer, although most are restricted to HLA class II alleles. [*4;94;96;104;183;212;214;257;258*]. Assessment of risk is complicated by interactions between alleles making up the MHC, as well as the genotypic make-up. Evaluation of interactions with specific HPV types may also help to uncover why some subjects are more likely to harbor persistent infections and develop carcinoma [*145*].

Studies that have looked at immunosuppression have involved such patients as HIV positive individuals and transplant recipients [*56;189;238*]. Both populations are more prone to develop cervical cancer [*127;170*]. Studies have shown that there are higher quantities of HPV the lower the CD4 level is in HIV positive patients [*92*].

2.4.3. Viral co-factors

2.4.3.1. Viral burden

Cross-sectional studies of HPV have observed an association between viral load and cervical carcinoma. Longitudinal studies of viral load and incidence or progression of HPV infections to

high-grade lesions have relied on different strategies to measure viral burden [*35;106;107;116;268*]. It is unknown whether the viral quantity is a result of one cell with a large number of virons or a large number with few virons. Differences in sampling methods of cervical specimens from which viral load is tested; varying methods used to quantify viral load; and differences in classifications of disease among studies have made interpretation of results and comparison across studies difficult. Of those studies that employed PCR methods of quantification, some have shown an increased risk of lesion incidence with higher viral load [*116;248;268*].

2.4.3.2. Viral integration

A key event in the natural history of cervical neoplasia is the integration of the HPV viron into the host cell genome. While evidence of integration has been found in cervix tumors [*60;126;144;181;196*], little is known about this area of research. Changes in viral load may reflect this integration process and concomitant loss of episomal viral forms, and serve as a marker of the morphogenesis of the squamous epithelium into neoplasia and carcinoma. Integration results in the inactivation of tumor suppressor genes, which lead to cell immortalization [*57*]. Integration is believed to be an irreversible process. During integration E1 and E2 viral genes are frequently disrupted, E6 and E7 viral oncogenes are retained. Recent methods have been adapted to measure the ratio between E2 and E6 protein genes of HPV 16 as a marker of the degree of integration [*181;196*].

2.4.4. Age

While age can also be included as a host co-factor, it is often evaluated separately. After sexual activity, age is the second most influential factor in the risk of HPV infection. The prevalence of HPV tends to fall after the age of 25-30 years and peaks again at 55 years. The first peak is likely related to onset of sexual activity while the second may be related to changes in hormonal status and immune response due to menopause [*99*]. Incidence rates for invasive cervical carcinoma have been found to be higher for older women (>50 years) [*168*]. However, an independent risk effect of CIN cannot be ruled out. In this study, the RRs of incident LSIL and HSIL differed somewhat with age even after considering the effect of HPV infection status [*224*] (Paper I).

2.5. Evidence of the role of HPV in the development of cervical cancer

Most epidemiologic studies of pre-invasive cervical lesions are based on prevalence data collected from cytopathology case series or case-control studies [reviewed in *71*], few of which are population-based [*125*]. RR estimates for the association between HPV infection and risk of CIN/SIL vary considerably among these studies [*71*], a reflection of the viral DNA detection methods used, HPV types tested for, method of lesion ascertainment, lesion grade evaluated,

case-control sampling methods, and source populations used. Nonetheless, these studies serve as the foundation on which use of HPV infection and preinvasive lesions as surrogate markers of risk for cervical cancer is based.

Of those studies of incidence of cervical cancer that are population-based, most derive data from tumor registries or screening programs, where little information is collected on the incidence of pre-invasive lesions or CIS [*71;125;255*]. This only allows for a limited assessment of the incidence of such lesions.

Natural history studies that have been longitudinal in nature are less prevalent, of which a few are based in North America [*53;106;107;131;138;171*], Europe [*262*], and a couple in developing countries [*62;101*]. These investigations obtain baseline cytologic information on specific population groups and then measure the incidence of lesions among those who were free of abnormalities at enrollment. While various exposure and outcome measurement methods have been employed in the above studies, only a few contain multiple measurements of either exposure or outcome at repeated intervals. Among these, two studies have reported on HPV persistence and subsequent SIL incidence over an extended period of time [*171;262*].

2.5.1. *Studies of cervical neoplasia incidence*

Estimates of the incidence rate of SIL vary across studies. Table L-2 shows the crude and age adjusted incidence rates estimated from case-control and cohort studies conducted around the world. Those that were able to remove prevalent lesions detected at enrollment by Pap cytology observed crude rates for dysplasia of 0.134 to 3.577 per 1000 woman-months, depending on the method of detection and source population [*71*]. Rates of CIS and those based on histology tended to be lower, ranging from 0.003 to 0.75 per 1000 woman-months. While previous studies have not calculated incidence rates of dysplasia according to HPV status, long-term studies of lesion incidence have shown higher cumulative risks for LSIL and HSIL among individuals with oncogenic HPV infections [*171;262*].

Table L-2: Summary of published crude and age-adjusted incidence rates of preinvasive cervical malignant lesions and associated abnormalities*

Study reference	Population, study period	Ascertainment of outcome	Outcome	Incidence rate (/1000 woman-months)	Study description
Stern and Neely [*236*]	US, Los Angeles	Histology	Dysplasia, 1955-64 CIS	0.092 0.003	Screening population, all ages
Parkin et al. [*194*]	England, Leeds and Wakefield	Histology or persistent abnormality on follow-up smear	Dysplasia, 1976-77 CIS	0.199 0.058	Screening population, all ages, N=81890, estimates corrected for unconfirmed diagnoses

Miller et al. [*158*]	Canada, Toronto	Cytology	Dysplasia, 1962-81 CIS	3.577 0.020	Retrospective cohort within a cytopathology database, all ages, 16053 woman-years
Gram et al. [*83*]	Norway, Troms and Finnmark	Histology	CIN 3, 1980-89	0.198	Screening population, N=43016, ages 23-72
Kainz et al. [*120*]	Austria, Vienna	Cytology	CIN, 1980-84 CIN, 1985-89	0.361 0.72	Screening population, N=12604, all ages
The New Zealand Health Study Group [*242*]	New Zealand	Cytology followed by histology or DNA ploidy	Dysplasia or worse lesions, 1980-86	0.742	Cohort using contraception, N=7200, ages 20-39
Morrison et al. [*169*]	Canada, British Columbia	Histology	CIS + invasive, 1949-92	0.05, cohort 1 0.75, cohort 2	Screening program cohorts, N=119,000
Bos et al. [*14*]	Denmark, Maribo	Histology	CIN, 1966-82	0.158	Screening population, all ages, 106,000 woman-years
Sawaya et al. [*216*]	US, NBCCEDP	Cytology	LSIL, 1991-98 HSIL ASCUS	0.564 0.134 2.192	Screening population, N=128805, 15.7 months follow-up, all ages
Sawaya et al. [*215*]	US, multi-center	Cytology, colposcopy, and histology	SIL (Pap) SIL (colpo/histology) ASCUS/AGCUS (Pap)	0.222 0.136 1.652	Hormone replacement trial of postmenopausal women, N=2561, 4895 woman-years

* Modified from Franco and Ferenczy [*71*].

2.5.2. *Longitudinal studies of HPV and cervical neoplasia*

RRs for the association between HPV infection via viral DNA detection and risk of CIN or cancer as estimated in several epidemiologic studies conducted during the past 10 years [reviewed in *71*] are all of high magnitude – in some studies RRs greater than 100 have been observed [*99;140;186*]. The magnitude of the association with HSIL (or equivalently CIN grades 2/3) is greater than that for LSIL (equivalently CIN 1). In addition, associations tend to be stronger when viral exposure definition is restricted to HPV 16 [*186;217*], the main viral genotype found in cervical cancers worldwide [*17*]. Some studies ascertained exposure to a broad spectrum of HPVs whereas others restricted detection to only a few of the so-called oncogenic types. In all, the associations are very strong; no other risk factor for cervical neoplasia is of comparable magnitude.

There are some underlying differences between the above-mentioned cohort studies with respect to the characterization of HPV infection status, and in particular, with respect to the definition of HPV persistence. The first approach involves evaluation of HPV status at two points in time: the first at enrollment and the second being a prevalence measure collected at the same moment as the outcome is diagnosed [*138;255*]. These studies are similar to cross-sectional studies where the time sequence between exposure and outcome is unclear.

A second approach involves repeated measurement of HPV before the onset of disease [*53;106;107;131;174;224*]. Of those studies that do follow a longitudinal approach with scheduled visits at repeated intervals [*53;107;174*], investigations are limited by small sample sizes, variable or less sensitive methods of HPV detection, or come from highly selected populations [*174*] or from other studies such as randomized control trials [*107*].

Some prominent studies investigating the natural history of HPV and CIN are based on high risk populations including women who had already received diagnoses of ASCUS or LSIL [*106;131;142*] or who were targeted for their high STD infection histories (either for HPV or HIV) [*53;174*]. As a general observation, these studies have shown higher risk associations for SIL outcomes for both persistent and productive HPV infections.

2.5.3. *Studies of HPV persistence*

There are a few ongoing cohort studies of the natural history of HPV infection and cervical lesions that are collecting data on HPV on multiple follow-up opportunities. By necessity, these studies are smaller. The types of study design vary considerably, as well as the criteria to define persistence and transience (also called intermittence by some) of infections.

As part of a NIH study in Portland, Oregon, Hildesheim et al. [*105*] followed up 393 cytologically normal women for a repeat cervical sample over a period of 30 months. Persistence of HPV infection determined by consensus PCR was found in 64 (16.3%) of these women, which represented less than half of all women who tested positive at least once (141 subjects: cumulative positivity of 35.9%). Older age (>30 years), interval time between cervical samplings and presence of an oncogenic type in the first specimen were independent predictors of persistence. However, the authors defined persistence in broad terms, identifying specimens positive for one or more HPV types during both visits as persistent.

Following young women with at least one positive HPV result by consensus PCR Xi et al. [*264*] found short-term persistence by HPVs 16 variants to be a frequent event, whereas long-term persistence was practically nonexistent after one year. Persistence is defined in the Seattle cohort as same-type infections. Subjects who were repeatedly HPV-16 DNA-positive over 2 to 8 4-monthly visits showed identical single-stranded conformational polymorphism patterns at every visit.

In a study of adolescent women attending an STD clinic at the University of California at San Francisco [*171-174*] persistence of HPV infection was defined as overall positivity by dot blot hybridization using separate probe mixtures for non-oncogenic and oncogenic types. Among the

288 women aged 13-22 who were evaluated, persistence declined steadily with time since enrollment. Almost half of those initially positive became negative within 30 months and 40% had new or different types than the ones detected at enrollment. With such a high variability in patterns of HPV positivity defined on the basis of broad spectrum probe mixtures it is likely that variability would have been even more pronounced had persistence been defined by same-type infections only.

Ho et al. [*106*] used a repeated-measurements, longitudinal cohort study design to evaluate the relationship between persistent HPV infections and risk of lesions in young women. The median duration of HPV infection was 8 months and by 12 months of a positive result 70% of the women had cleared their infections.

Ylitalo et al. [*266*] used a novel single-tube nested PCR system for assessment of HPV 16 persistence in a retrospective study of registry data from a cytological screening program for CIS in Sweden. Infection with HPV 16 in the two most recent Pap smears before diagnosis was associated with an increased risk of CIS. Viral load estimates were also performed by 5'exonuclease (Taqman) PCR method on 478 cases and 608 matched controls that revealed a 60-fold increase in risk of CIS among women with the highest amounts of HPV DNA [*116*].

2.5.4. *Estimation of rates of progression*

A few review and meta-analytical studies have attempted to summarize rates of progression and regression along the continuum of cervical dysplastic changes using different criteria for selecting investigations and for combining natural history data. For preinvasive stages of disease it has been shown that less than 2% of CIN lesions (including grades 1-3) progress to invasive cancer and 25% of CIS cases progress to invasive carcinoma if left untreated, with the transition from CIN to CIS to invasive carcinoma occurring over a period of 10 to 20 years [*164*].

Several retrospective studies have examined the ongoing effects of preinvasive lesion stages (decribed in chronological order below). While these studies report rates of progression and regression, they are beset with several limitations. First, most of the studies had small sample sizes, came from highly selected study populations, or had insufficient follow-up time to evaluate the progression from mild dysplasia to severe dysplasia or cancer. Second, previous studies report only crude rates of progression and regression without regard for proper actuarial analysis of cumulative risk over time. Third, several of the studies relied on pooled estimates that were based on variable methods for detecting lesion development during follow-up. In particular, those using histology to determine outcome may have altered the course of the natural history of the disease because frequent cervical biopsies may remove the entire lesion. These problems tend

to affect the comparability of results across studies [*109*]. Fewer studies have tried to estimate the mean preclinical duration of cancer or sojourn time [reviewed in *71*]. None of the previous studies of lesion progression looked at risk factors like HPV, and few investigated the effect of age. Of these studies, most were based on screening program results and extrapolated data from different sources such as screening studies, national surveys and fee schedules, and published literature in order to make predictions using Markov models [*8;80;180*].

McIndoe et al. [*152*] followed patients diagnosed histologically with CIS for a period of five to 28 years. The authors observed differences in cumulative incidence of invasive carcinoma after 28 years among subjects who displayed normal cytology following biopsy (1.5%), and those who continued to produce abnormal cytology (22.1%). The majority of those in the latter category (95%) persisted with CIS or progressed to an invasive stage. While special care was taken to reduce the disturbance of the lesion at the initial diagnosis of CIS, 60% of the subjects had normal cytology after biopsy.

In a review of studies on the natural history of cervical neoplasia conducted between 1950 and 1993, Östor [*187*] looked at the progression, regression and persistence of CIN determined by biopsy and cytology. The results show that the likelihood of regression of a CIN 1 lesion is 57%, persistence 32%, progression to a CIN 3 lesion was 11%, and progression to invasive carcinoma was 1%. For CIN 2 lesions, the likelihood of regression was 43%, 35% persisted, 22% progressed to CIN 3 and 5% to invasive carcinoma. Of the CIN 3 lesions, 33% regressed and greater than 12% progressed to invasive carcinoma. Less than 56% persisted as CIN 3. The results demonstrated that even severe lesions regress in the majority of cases. In this review evidence was obtained from case studies, subjects were not followed over time at regular intervals, and time to regression or progression was not measured actuarially over time using methods such as Kaplan-Meier curves or life table analysis.

In a review of 31 studies by Mitchell et al., [*162*] the authors derived pooled probabilities of regression, persistence, and progression of all CIN grades detected by cytology of 34%, 41% and 25%, respectively. Regarding the latter progression figure, 10% of the lesions progressed to CIS and 1% to invasive cancer. The studies included in the review followed untreated patients sequentially with Pap smears or biopsies, and were selected on the basis of their similarities in observed rates of progression and regression of cervical neoplasia. These studies, however, were affected by short follow-up, incomplete outcome assessment, loss to follow-up, and selection bias due to non-random subject selection and differential outcome assessment. Mitchell et al. observed a higher probability of regression, and less persistence or progression using

biopsy follow-up: 45%, 31% and 23%. These latter values are not exclusive for biopsy, however, as most of studies included employed both cytology and biopsies for follow-up.

In a systematic review of 15 studies between 1970 and 1996, Melnikow et al. [*155*] evaluated the pooled rates of progression and regression at 24 months after an initial abnormal cytology or biopsy result. Studies with at least 6 months of follow-up were included. The rate of regression to normal from LSIL was 47% and 35% from HSIL. The rates of progression to HSIL at 24 months were 7% for ASCUS, 21% for LSIL and 23% persisted. Rates of progression to invasive carcinoma at 24 months were 0.25% for ASCUS, 0.15% for LSIL and 1.4% for HSIL. The authors therefore concluded that borderline dysplasia and LSILs detected by cytology show a low risk of invasive cervical cancer up to 24 months. This review was limited due to the fact that the different studies had varying periods of follow-up from 6 months to 24 months, thereby preventing the calculation of time to regression or progression. This heterogeneity between source studies was not explained by regression analyses of study-specific variables including: performance of biopsy, duration of follow-up, percentage of subjects lost to follow-up, mean or median age, and a quality score determined from study design. Furthermore, assessment of lesion status included both biopsy and cytology without differentiation.

Using cytology results from a large provincial database from Ontario, Holowaty et al. [*109*] estimated the rates of progression and regression for different severities of dysplasia before treatment. Overall, progression at 2 years to severe dysplasia and worse was 2.1% for mild dysplasia and 16.3% for moderate dysplasia. A higher proportion of cases progressed to the next level within 24 months while a smaller proportion progressed after this period relative to the previous 2 years. With mild dysplasia taken as the referent, RRs of CIS were 8.1 within a 2-year period for moderate dysplasia and 22.7 for severe dysplasia. The equivalent RRs of invasive cancer were 4.5 and 20.7 for the latter lesion grades, respectively. However, these rates did not consider persistence of lesions. That is, without active evaluation of lesion status at repeated intervals, the study could not determine whether observed instances of progression were not in fact new occurrences of transient lesions. The average number of smears taken per subject was low (2.85) over a mean period of follow-up of 159 months. The observed occurrences of progression within 24 months may have been due to an under call of the initial abnormal specimen. The authors could not account for censoring due to treatment and could have underestimated the rates of progression. Nonetheless, the results for this study [*109*] remain important for the clinical management of women with mild or moderate lesions.

Previous studies of disease progression and regression were not able to differentiate lesions according to biomarkers such as HPV. In a recent study of disease regression of 136 cases of

CIN 1 or 2 determined histologically, Kadish et al. [*118*] observed increased odds of regression within 12 months for subjects showing an immune response to specific HPV 16 E6 and E7 peptides. Previous evaluation of the method used to test for cell-mediated immune response to the above peptides showed a higher likelihood of loss of HPV infection [*117*]. While the sample size for the former evaluation was too small to produce significant results for all of the peptides, this study indicates an effect of immune response to HPV infection on the rate of disease regression.

2.5.5. *Considerations for studies on disease progression*

Many of the above studies were performed in high-risk populations, such as STD clinics [*174*] or in immunosuppressed patients [*53*]. Other studies relied on subjects with dysplastic lesions at enrollment for comparison [*2;171*], or determined HPV status at the moment of detection of a neoplastic event [*138;255*]. Estimates derived from these studies may have limited generalizability. Without interviews and long-term follow-up, previous prospective studies were also not in a position to evaluate the whole natural history process from acquisition of an HPV infection and persistence to SIL incidence and progression.

Earlier studies were also limited in their use of HPV detection methods that are less sensitive than the current state-of-the-art methods [*107;131*] and are more prone to misclassification reducing the reliability of risk estimates [*65;67*]. Reliance on single point assessments of infection status cannot determine persistence of HPV infections. Early studies of viral burden have likewise been limited to semi-quantitative methods such as strength of hybridization signal [*89;237;240*].

While misclassification of outcome status due to the use of cytology is also a concern in this study, it remains one of the more sensitive methods of SIL detection available and the most widely established in medical practice. Reliance on alternative detection methods in previous studies such as histological analysis of biopsies for confirmation, which remove tissue from the transformation zone, may not have been appropriate to investigate the natural history of precursor lesions [*206*]. Results from such studies therefore should be considered with discretion and may not provide useful recommendations for screening or general clinical management, which rely almost exclusively on cytology-based follow-up.

3. STUDY DESIGN AND SUBJECTS

This thesis project and the respective manuscripts included are based on a larger cohort study being carried out in Brazil as product of the collaboration between two research centers: McGill University and the Ludwig Institute for Cancer Research in São Paulo, Brazil. The study design and methods of data collection are summarized below and repeated to some extent in each of the manuscripts.

3.1. Description of cohort

This ongoing cohort investigation of the natural history of HPV infection and cervical neoplasia in a population at high risk for cervical cancer in Latin America began in November 1993. The study population is a systematic sample of women attending a comprehensive maternal and child health program catering to low income families in the city of São Paulo, Brazil, as part of a network of primary, secondary, and tertiary health care institutions maintained by the municipal health department. The study design has been presented elsewhere [*62*]. Women were selected at random from the daily lists of outpatients in the family medicine, gynecology, and family planning clinics. Study nurses approached each patient to determine eligibility and to explain the purpose of the study. Those willing to comply with all the study requirements were then asked to sign an informed consent form.

The investigation has received ethical approval from McGill University, University of Toronto, University of Arizona, Ludwig Institute for Cancer Research (institutions with which the laboratory co-investigators are affiliated), and from the local clinic setting in São Paulo (Appendix E). Continued approval has been granted on annual reviews by all ethical boards concerned since 1995. The criteria for eligibility were as follows: (i) age between 18 and 60 years, (ii) permanent city residence, (iii) not currently pregnant nor intend to become pregnant during the next 12 months, (iv) having an intact uterus and not being a current referral for hysterectomy, (v) not having used vaginal medication in the last 2 days, (vi) not having been treated for cervical disease in the last 6 months.

3.2. Cohort management

Table D-1 summarizes the various study procedures and instruments used at the initial visit and during each of the pre-scheduled returns for all subjects. Subjects enrolled into the study are being followed up over a 10-year period in pre-scheduled returns every 4 months, in the first year, and twice yearly thereafter. In the first 4 visits and on the annual returns, subjects are submitted to an interview using a structured questionnaire specific for the current visit and have cervical specimens taken for Pap cytology and HPV testing. During the semester (6 month) returns between annual follow-up visits only the cervical specimen is collected for cytology and HPV

testing (no questionnaire). A cervicography is performed once in the first year during any one of the visits, when it is mutually convenient for the patient and for the nurses. Additional cervicographies are performed every 2 years thereafter.

Table D-1: Design of the Ludwig-McGill cohort study

Procedures / Instruments	**Entry**	**4 mos**	**8 mos**	**1 yr**	**1.5 yrs**	**2 yrs**	**2.5 yrs**	**3 yrs**	**3.5 yrs**	**4 yrs**	**4.5 yrs**	**5+ yrs**
Viral markers:												
HPV testing and typing	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Viral load	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Molecular variants	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
HPV serology	✓	✓	✓	✓		✓		✓		✓		✓
Host susceptibility markers:												
HLA typing	✓											
p53 polymorphism	✓											
Cervical pathology:												
Local Pap cytology	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Cytology review	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Cervicography			✓			✓				✓		
Colposcopy + biopsy	<---------- ✓ (whenever needed if HSIL) ---------->											
Questionnaire information:												
Sociodemographics	✓											
Diet		✓								✓		
Reproductive health	✓		✓									
Sexual behavior, smoking	✓	✓	✓	✓		✓		✓		✓		✓
Health attitudes and beliefs										✓		
Compliance incentive:												
Meal tickets (in US$)	5	10	15	20	20	20	20	20	20	20	20	20

The study nurses telephone the patients a few days before the scheduled returns to remind them of the pending visits. Missed appointments are followed by phone and/or letter. These attempts at contacting subjects are repeated once a month until an appointment can be scheduled or the woman explicitly states that she will drop out of the study. The process of follow-up of subjects recruited in the study is illustrated in Appendix A.

Several questionnaires are scheduled during the course of the study follow-up corresponding to each of the first year visits and for the annual follow-up returns (Appendix B). Questionnaire 1 is administered at enrollment, being the most detailed with 107 questions. The information that is collected in these questionnaires cover all classes of proven and suspected risk factors for HPV infection and cervical neoplasia, i.e., sociodemographics, reproductive health, sexual practices,

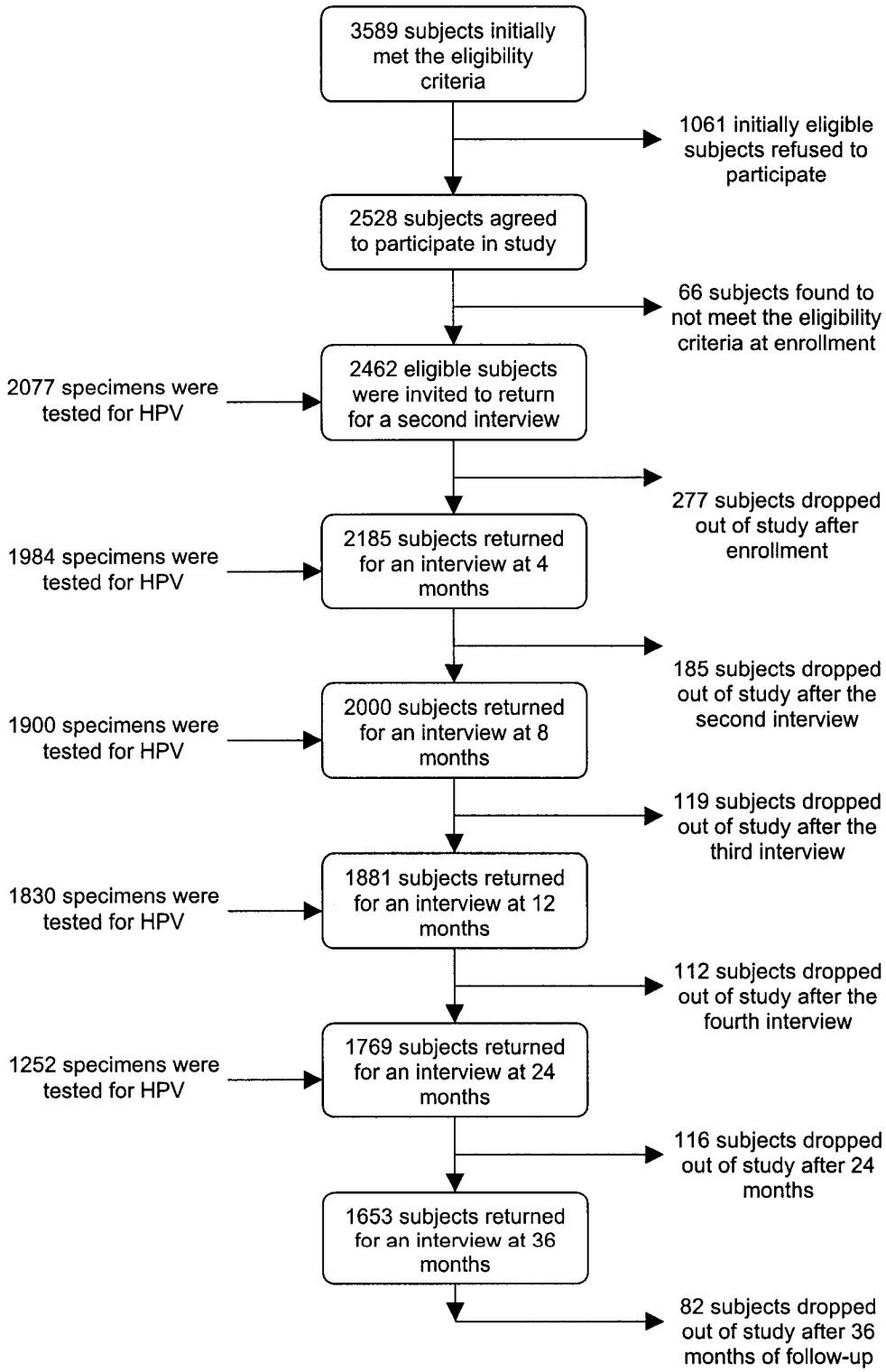
smoking, and diet. The emphasis on a given class of factors changes depending on the visit number. While co-factor information on demographics, sexual behavior, reproductive health and smoking were evaluated for confounding in this thesis, factors related to diet and host co-factors were the subject of separate projects.

Because of the importance of retaining subjects for the entire duration of the study participants were told at enrollment that they would receive cash-equivalent incentives. Participants were offered meal tickets, which are widely used in Brazil as part of employee benefits for salaried workers and are honored in supermarkets in exchange for meals and groceries. The cash-equivalent incentives begin at $5 at the enrollment visit and increase $5 per subsequent visit to a maximum of $20, which then remains in all subsequent appointments that are kept by the participant. This strategy has ensured excellent rates of follow-up compliance (see below) despite the complexity of the procedures used in the study.

3.3. Study status

As of March 2002 the following statistics applied to the study. A total of 2528 women had been recruited through March 1997 when the target sample size for the main protocol was attained (N=2300). Figure D-1 illustrates the recruitment and follow-up status of the Ludwig-McGill cohort study. The scheduling of semester visits, however, was not introduced into the study until 1997. As a result, some women enrolled at the beginning of the study had missing data for some of the initial semester returns. Follow-up since this change in study design had kept to the predetermined 6-month schedules as described above. After elimination of ineligible subjects found after enrolment, 2462 subjects were followed over time. Actuarially calculated proportions of women who have been compliant with all scheduled follow-up visits were 76%, 72%, 67%, 64% and 55% at 12, 24, 36, 48 and 60 months, respectively. A total of 17,614 clinic visits have been logged since enrollment, corresponding to 130,748 women-months of follow-up (means of 9.2 visits and 53.1 months per woman). The median follow-up time is 65.5 months. In terms of crude proportions of visits completed including baseline, 25% had 12 or more, 50% had 10 or more, and 75% had 4 or more. A total of 20,999 cytology reports and 9437 HPV test results (thus far, the latter covering specimens collected during the first 2 years) have been entered on separate databases after checking for accuracy. Newly documented infections have occurred in the cohort at the rate of 1.32 per 100 women-months (95% confidence interval: 1.2-1.5). The cumulative rates of incident lesions detected by cytology documented in the cohort are as follows: ASCUS, 6.3%; LSIL, 4.5%; HSIL, 1.0%. These proportions are based on a hierarchical coding reflecting the worst cytological diagnosis found on all cervical smears during the study for all women.

Figure D-1: Flow chart of subject recruitment and compliance in the Ludwig-McGill Cohort for the first three years of follow-up



* Based on March 2002 update of Ludwig-McGill Cohort

3.4. Cervical cell specimens

An Accelon biosampler (Medscand, Inc.) is used to collect a sample of ectocervical and endocervical cells at each of the visits. After the smear is prepared onto a glass slide and fixed in 95% ethanol, the sampler containing the exfoliated cells is immersed in a tube containing Tris-EDTA buffer pH 7.4 and kept at 4°C at the clinic for at most 5 days. Once brought to the laboratory at the Ludwig Institute the tubes are vortex-mixed to release the cells from the sampler, the latter is discarded, and the tubes containing cell suspensions are frozen until testing.

The Pap smears are fixed in absolute ethanol, stained, and read at Ludwig Institute's cytopathology laboratory for an initial diagnosis and are then shipped to the principal investigator, in Montreal, Canada, where they are coded and then sent to the laboratory of Professor Alex Ferenczy, at the Jewish General Hospital, one of McGill University's teaching hospitals, for classification. Statistical analyses for this thesis are based on the Canadian cytology due to the initial lack of correspondence between the Brazilian and Canadian scoring. The Montreal cytopathology reports are based on the 1992 Bethesda system for cytological diagnoses and are blinded to all other screening results for the same sample [*97*]. Results from the Brazilian cytology were initially based on the original Papanicolaou scheme. These are currently being reread and classified using the Bethesda system. Wherever HSIL is detected by either cytology system or by cervicography the woman is referred for colposcopy and biopsy by a hospital gynecologist. In two instances women with HSIL could not be reached until several years later, and two women were referred for biopsy after detection of LSIL at two consecutive visits.

3.5. HPV DNA detection

All viral testing methods were performed at the Ludwig Institute for Cancer Research (LICR) in São Paulo. Cervical specimen DNA is extracted and purified following standard techniques. In brief, cells are digested with 100-ug/ml proteinase K for 3 hours at 55°C, followed by organic extraction and ethanol precipitation. Cervical specimens were tested for the presence of HPV DNA by a previously defined PCR protocol amplifying a highly conserved 450 bp segment in the L1 viral gene (flanked by primers PGMY09/11) [*10;105*]. Typing of the amplified products was performed by hybridization with individual oligonucleotide probes specific for all 27 HPV genital types whose nucleotide sequences for probes within the MY09/11 fragment have been published in the literature [6/11, 16, 18, 26, 31, 33, 35, 39, 40, 42, 45, 51, 52, 53, 54, 55, 56, 57, 58, 59, 66, 68, 73, 82, 83, and 84] [*105*]. Amplified products that hybridize with the generic probe but with none of the type-specific probes were tested further by RFLP analysis of the L1 fragment [*12*] to distinguish among unknown HPVs. Use of the RFLP analysis extends the range of identifiable HPV types to include additional HPVs 32, 34, 44, 62, 64, 67, 69, 70, 71, 72, 81, CP6108 and IS39 plus other unknown types. To verify the specificity of the hybridizations, more than 30 type-

specific positive controls were included in all membranes. In order to check the integrity of the host DNA material extracted from the specimens, assays also included an additional set of primers (GH20 and PC04) to amplify a 268 bp region of the ß-globin gene [*10*]. All HPV assays were done blindly on coded specimens with no identification linking specimens from the same woman. Appropriate precautions were taken to reduce the possibility of specimen contamination.

3.6. Measuring viral burden

All cervical specimens found to be positive with the main PCR protocol (MY09/11) were reprocessed by a quantitative, low-stringency PCR to measure viral burden in exfoliated cervical cells [*29*] Using this method I evaluated the association between viral burden and incidence of cervical neoplasia in Paper II. The method uses the general primers GP5/6. This PCR protocol is well known and detects a broad spectrum of HPVs [*247*]. The quantification protocol employs low stringency conditions to co-amplify the specific HPV DNA fragment along with DNA sequences from the human genome present in the starting PCR mixture. The amplified DNA is then run on polyacrilamide gels to allow visualization of bands of HPV and human DNA and subsequent quantification using silver staining. Standards consisting of mixtures containing varying amounts of reference HPV 16 plasmid are included in duplicate in every assay added to a constant background of normal human DNA (corresponding to 4, 20, 100, 500, and 2500 viral copies per cell). In addition, control samples consisting of DNA from two cervical carcinoma cell lines with known quantities of HPV copies (HeLa, 20-40 copies of HPV-18; Caski, 400-600 copies of HPV-16) are included in duplicate in every assay. The silver-stained gel bands corresponding to the HPV and to the constant human genome fragments are quantified by densitometry [*29*]. The logarithm of the ratio between these two bands is directly proportional to the logarithm of the amount of HPV DNA in the individual samples. Proper quantification is obtained by linear interpolation in a standard curve constructed with the results from the control mixtures.

The viral load quantification method used in this study involves a PCR-based protocol designed to detect the L1 region of the viral genome. Integrated forms of the virus should have been amplified along with episomal (or extra chromosomal) forms. Testing under a wide range of conditions has shown that the PCR-based technique used to measure viral load in this study is reproducible and adaptable to large scale testing in epidemiologic studies [*29*]. The method has also been demonstrated to be reliable over repeated samplings since women harbor similar levels of viral load from one visit to the next [*245*]. Dysplastic cells found in CIN 2 or 3 lesions express fewer intercellular adhesion molecules than normal cells [*127*] and could possibly be sampled more readily than normal cells. The method used in this study corrects for this over sampling by normalizing the numbers of copies of viral DNA against the quantity of host DNA permitting the calculation of true viral load in terms of the number of copies per cell, thus

eliminating the fluctuation due to variation in cell content among specimens from different subjects and from the same subject over time. Furthermore, the results from the inclusion of cervical carcinoma cell lines as secondary controls (HeLa and Caski cells) in every testing batch indicate that the method is sufficiently accurate and precise in allowing a quantitative assessment of the number of HPV copies per host cell [*29*]. The dose-response relationship is linear at concentrations as high as 5000 copies per cell and is independent of the amount of DNA present in the reaction mixture, of the number of PCR amplification cycles, of staining intensity, and of the choice of human genome bands used as reference [*29*]. In all, unlike previous methods based on semi-quantitative assessment of PCR signals using external standards [*241*] this technique appropriately satisfies the criteria for quantitative measurement of viral load in cervical specimens. That is, it is reproducible, has an adequate linear range for dose-response, and provides results standardized for cell content.

4. STATISTICAL ANALYSES

A general introduction and description of the statistical methods used for the analyses in this thesis and the connecting manuscripts are presented below. While each manuscript includes a description of the individual statistical analyses performed, the sections below further discuss the assumptions and limitations of the respective methods and serve to compliment the manuscript texts. Each method is itemized according to the manuscript in which it was applied.

4.1. Exposure definition

The primary exposure variable evaluated in this dissertation was HPV infection status. While other factors have been identified in the natural history of cervical cancer and its precursors, these factors only served as measures of confounding or as mediators for the main etiological relationship with HPV that was investigated in this study.

For the investigation reported here, HPV types were grouped by oncogenic potential. Non-oncogenic HPVs included types 6/11, 26, 32, 34, 40, 42, 44, 53, 54, 55, 57, 62, 64, 66, 67, 69, 70, 71, 72, 73, 81, 82, 83, 84, CP6108, IS39, and other unknown types. Oncogenic HPV types included 16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, 59, and 68, based on an expanded classification of Bauer et al. [9]. Stratification by HPV oncogenicity was hierarchical and based on mutually exclusive categories.

HPV infection status was determined both at enrollment, and at the enrollment and first follow-up visits combined to ascertain initial persistence of HPV infection (Paper I). Algorithms to define persistence considered the following *a priori* definitions of persistence: (1) Subjects with at least two consecutive HPV positive cervical specimens for any types during the first year. This was the most inclusive definition but does not provide useful information for risk association given neither oncogenicity nor persistence can be evaluated. (2) Subjects with at least two consecutive cervical specimens positive for HPV DNA of non-oncogenic or of oncogenic types during the first year, regardless of whether they contain the same type(s) or not. (3) Subjects with at least two consecutive cervical specimens positive for the same oncogenic HPV type during the first year. Statistical evaluations based on the latter two of the three algorithms were presented in this thesis. HPV types 16 and 18 were also treated separately in some analyses if subjects presented positive for either HPV type at the index visits regardless of the presence of other types. Following a conservative approach to exposure classification, subjects with invalid HPV test results at any of the visits were excluded from the analyses of persistence (Paper I and III) or substituted with the last valid result (Paper IV). Average values calculated from valid results at repeated measurements were used in some analyses (Paper II). HPV persistence extended over a period of observation of three

visits, was also assessed in another series of analyses (Paper I) using the third algorithm defined above.

4.2. Co-factor information

I adjusted *a priori* for age and ethnicity using the following categories for age: 18-24, 25-34, 35-44, and 45-60 years, and for ethnicity: white and non-white. The original categorization schemes used on the interview questionnaires were used for most of the other covariate factors. Categorization of continuous values such as number of sexual partners, age at first intercourse, number of pregnancies and years of previous OC use, were based on percentiles or on *a priori* schemes that were in line with biological implications.

4.3. Outcome definition

Most studies on CIN to date have considered cytology results for cervical cancer precursors using the Bethesda classification system with ASCUS, LSIL, and HSIL. This system was adopted for comparison with current North American clinical standards.

In the current cohort, Pap smears were also reread by the Canadian cytopathologist using an expanded version of this classification scheme. Each instance of Normal, ASCUS, LSIL, and HSIL was subcategorized into two groups to denote the severity of the diagnosis (Paper II): Normal smears were classified into smears showing epithelial cells within normal limits or with benign cellular changes; ASCUS diagnoses were redefined as either ASCUS 'favor benign' or ASCUS 'favor SIL'; LSILs were separated into lesions showing changes suggestive of koilocytotic atypia (LSIL/HPV) or squamous abnormalities equivalent to mild dysplasia (LSIL/SQ) or CIN 1; and HSILs were redefined as moderate (HSIL/CIN2) or severe dysplasia (HSIL/CIN3). This classification system represents a mix of cytological and histological nomenclature that has been adopted in other institutions [*168*] and allows for the creation of four actual precursor lesion grades (2 for LSIL and 2 for HSIL) within a spectrum of severity (from LSIL/koilocytosis to severe dysplasia) and two categories for ASCUS, for a total of six states of cytological abnormality for which risks of transition were calculated as a function of covariates. While allowing for more detailed evaluation of outcome stages, this classification system was used at a cost in power and precision and was performed only in certain circumstances.

Another outcome looked at was the incidence of a persistent lesion. This was defined as the occurrence of two or more visits positive for any SIL event detected by cytology during follow-up allowing for one negative or non-SIL interval visit. As such, subjects with SIL detected at two separate instances within three subsequent 4-month or 6-month interval visits were considered to have a persistent SIL event.

4.4. Cohort analyses (Papers I and II)

The following analytical methods were applied to the statistical analyses of Papers I and II. While all methods are summarized in the included manuscript texts, they are explained in more detail below for reference. All statistical analyses were performed using the statistical program PEPI version 2.07 (Abramson, Jerusalem, Israel), SPSS® version 10.0 and 11.0 (SPSS, Chicago, IL), STATA® versions 6.0 and 7.0 (StataCorp, College Station, TX), and S-PLUS® for Windows (Insightful, Seattle, WA).

4.4.1. Incidence of lesions

Incidence of SIL was estimated both by lesion severity (first confirmed SIL event of any grade and first occurrence of HSIL), and by persistence of SIL for two or more consecutive visits allowing for at most one negative intermediate visit representing a period of 6 to 12 months (Paper I). Prevalent cases of lesions detected at enrollment were excluded from all longitudinal analyses.

The incidence density rates for the entire follow-up period at time of analysis was calculated for first incidences of either any SIL, only HSIL or persistent SIL. Calculation of incidence rates was based on the following formula for closed populations [*128;209*]:

$$incidence\ rate = \frac{No.\ disease\ events}{\sum_{persons} \Delta t}$$

where the No. of disease events refers to occurrences of a first SIL outcome event, and
t = the length of time spent at risk for each participant.

Lesion incidence rates were calculated by summing the number of first occurrences of event divided by the total accrued women-months of follow-up. Person-time (Δt) was defined as the sum of all individual units of time spent in the study of subjects at risk until incidence of a disease event, censoring or loss to follow-up. Due to the ongoing nature of the cohort study, subsequent manuscript analyses (Papers II-IV) included longer periods of person-time as the time of censorship was extended as a result of longer follow-up time for uncensored subjects.

The calculation of incidence rates under this method assumes the source population is closed. That is, although recruitment may have occurred over a period of four years, no new subjects were added after recruitment was terminated and losses can only arise due to the occurrence of an event, death or loss to follow-up [*209*].

4.4.2. Cox proportional hazards regression

I analyzed the risk of post-enrollment occurrence of SIL as an incident finding in relation to HPV infection status at enrollment and during the first two visits (Paper I and II). The longitudinal rate ratios (RRs) were estimated using proportional hazards ratios (HRs). Respective 95% confidence intervals (CIs) of incident cervical lesions over time were modeled by Cox regression [*39;87*].

Time to event was measured from enrollment to the first instance of a lesion event or to the last recorded return visit date for censored subjects. This created a moving baseline time that followed the recruitment process from study initiation in 1993 to the last scheduled enrollment visit in March 1997. While follow-up is ongoing, right censoring could occur due to loss of follow-up before incidence of a cervical lesion. Therefore, disease-free time following these visits is not known. Censorship could have occurred either from loss of follow-up or due to unavailability of cytology results at the time of analysis. All time-to-event analyses included complete follow-up time for all subjects regardless of censoring.

To deal with such time-to-event data as a function of explanatory variables, Cox [*38*] introduced the concept of the hazard rate thought of as proportional to the instantaneous probability of an event at a particular time (see Appendix C for formulas).

Two important assumptions must be made when using Cox proportional hazards models: 1) event observations must be independent of one another, and 2) while the hazard function can vary over time, the ratio of the hazards for two contrasted units of observation should be proportional over the period of estimation. Evaluating time to occurrence of a first SIL event of interest as defined above, the assumption of independence between individual outcome events is maintained. I later allowed for the occurrence of repeated outcomes within subjects using variants of the Cox proportional hazards model, which are described later. I verified the assumption for proportional hazards over time using a graphical method. This involved plotting graphs for log(-log(survival at time t)), ln(-ln)S(t), over time from onset of study according to the stratification variables for HPV infection described below. Parallelism between graphical lines was taken as evidence of proportionality.

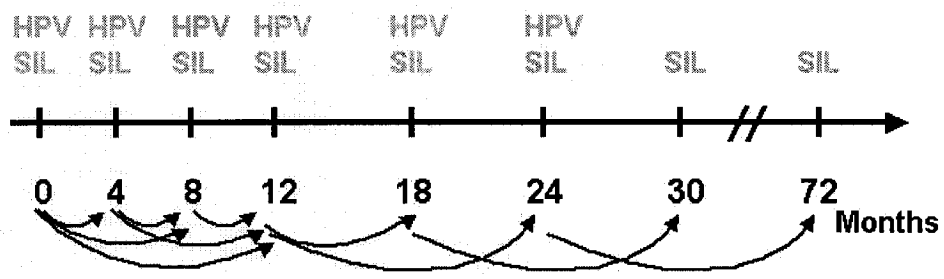
The handling of ties was based on the Breslow method [*19*] when outcome events were detected at the same point in time. This method assumes that the risk pool for the tied failure events corresponds to the largest risk set of each of the failure events including both event observations. This method has the advantage of being easier to handle and approximates the exact marginal likelihood when there are few tied events. Given the rarity of SIL events in this cohort study, this assumption was believed to be safe.

4.5. Analysis of repeated measurements (Papers II and III)

Due to the transient nature of both outcome and exposure, a woman can transit through a series of precursor stages of neoplasia form LSIL to HSIL before developing cervical CIS or cancer, or transit directly to a HSIL (Figure S-1). At any given lesion grade, a woman can also regress back towards a normal state. Likewise, a woman can test positive for HPV at any given visit and can then either eliminate the infection by the next visit or remain with a persistent infection. The situation resulting can be illustrated by the following example for the follow-up of a woman who presents without a lesion at enrollment (t_o) but then develops a lesion at the first return visit (t_i). She may then persist with the lesion until the next visit (t_{i+1}) or eliminate it at the next. Therefore, though the data layout treats observations independently, the correlation between observations within each subject must be taken into account when comparisons are made between populations (i.e. based on covariate patterns defined by their exposure history).

By considering such data as a general second level structure we can apply a standard set of marginal modeling techniques that allow any pattern of measurements while providing statistically efficient parameter estimation (Paper III). An illustration of the application of such methods is the repeated measures design as described above where measurements are 'nested' within individual subjects. These methods contrast with traditional regression models where each summary variable is analyzed simply as a function of covariates. Marginal regression models for multiple outcome events were employed since in clustered observations within individuals variation tends to be greater between individuals than between events within individuals [*82*].

Figure S-1: Repeated measurement of exposure and outcome over study follow-up



4.5.1. Cox regression for time-dependent variables

The partial likelihood function for the Cox model also allows for the inclusion of time-dependent variables [*93*]. Time-dependent Cox proportional hazards models can incorporate the transient nature of HPV infections by readjusting the HPV results over the course of follow-up to reflect the latest infection status observed at the previous visit. For the evaluation carried out on this study,

HPV status for each individual was allowed to change for the first six scheduled follow-up visits, corresponding to the end of the second year of follow-up (Paper III). HPV DNA test results from subsequent visits were not available for evaluation. As a result of the inclusion of time-dependent covariates, an expanded formula for hazard rates was used [*128*]. The formula used to derive the time-dependent hazard estimates is presented in Appendix C.

I evaluated the effect of HPV as a time-varying exposure based on a function of time since measurement. Time-varying hazards for non-oncogenic and oncogenic types declined with increasing time to event (data not shown). A test for proportionality of hazards was then done using Schoenfeld residuals (Figure A-1, Appendix D), and fit models for non-proportional hazards using the extended Cox model (Paper III) [*243*]. This statistic represents the sum of score process arrays across individuals plotted over time where a non-linear pattern of residuals is indicative of non-proportionality. With increasing duration of follow-up, a decrease in the magnitude of the HRs for the associations described above was observed with each new cohort update. While the log(-log) survival curves showed parallel relationships over time between HPV groups, the p-value for the test of proportionality of hazards using Schoenfeld residuals was 0.08, suggesting possible non-proportionality over time. This statistic was borderline, however, and more likely a result of a small number of events than effect modification.

Two additional assumptions are made when dealing with time-dependent covariates [*86*]. First, the models implicitly assume the exposure (HPV in this case) does not effect any covariate used to create strata or evaluated as a confounder. Second, there should be no confounding within strata or within levels of other covariates in the model. This becomes a problem if the covariate plays an intermediate role in the causal relationship between the exposure under evaluation and the disease. As demonstrated by prior studies of HPV and cervical neoplasia, most of the covariate risk factors analyzed in this thesis would not fall into this category (i.e. p53, HLA, smoking, OC use, parity, and age). Furthermore, given the blinding to HPV results of subjects at interview, it is also unlikely that changes in reporting or sexual behavior, sociodemographic status, smoking or reproductive health practices occurred as a result of changes in HPV status. While time-dependent Cox models do not generally allow individual predictive time-to-event curves, as do fixed models, the exponentiated coefficients give the instantaneous RRs of SIL over time as a function of the current HPV status.

Implicitly a temporal relationship is maintained in a time-varying model. That is, while the measurement of exposure may be allowed to change as a step function at repeated evaluations, it should not coincide in time with the outcome assessment. Consequently I linked measures of HPV status with ascertainment of cervical lesions at a subsequent visit in the longitudinal

analyses. All prevalent cases of SIL detected at enrollment from all these analyses were excluded.

4.5.2. Generalized Estimating Equations for marginal models

Generalized Estimating Equations (GEEs) have been adapted for use with longitudinal type data [*271*]. These models, based on the generalized linear model method are used to make inferences at the population level by modeling marginal means of repeated outcomes while taking into account the clustering within individuals (see Appendix C for formula). Although the correlation structure between outcomes can be parametric, it is treated as a nuisance parameter. Inference is then based on the coefficients for the covariates in the model that can be time-dependent, such as HPV infections, or time-independent (age and race) as defined above [*78*].

In a recurrent outcome event situation, the data can impose a time series structure to this correlation that is translated into an autoregressive function. In the context of this study, women with lesions may be more likely to persist until the next visit, and the strength of this correlation will decrease with time or as the number of interval visits between reference visits increases. Additional correlation structures were also evaluated including an exchangeable matrix structure assuming equal correlations between interval comparisons and an unstructured matrix that allows for separate estimations of correlation between intervals. Choice of correlation structure was initially based on biology. The resulting regression model coefficients were compared for consistency.

GEE models require that the data conform to a particular balanced structure. Broadly speaking, these procedures require that the measurement intervals and number of intervals be consistent across individuals. Visit returns in this study were always scheduled every 4 months, for a first year cohort visit (a total of four visits), or every 6 months for visits five and beyond. If a subject returned for her next visit after the scheduled date, the following interview was delayed to allow a complete programmed inter-visit period (4 or 6 months) to elapse. This management approach helped decrease the irregularity in visit interval times. Under these situations, however, GEE models can lose efficiency and the correlation structure between outcomes becomes less precise, as follow-up continues, if more individuals are measured irregularly. As a result, estimation of elaborate correlation structures is difficult, but can be partially circumvented by specifying an exchangeable or autoregressive correlation structure.

A second assumption for the GEE model is that missing data must be missing at random [*52*]. As described above, the occurrence of missing data was minimized in this study by allowing for delays in returning for a given appointment. While observations of missing HPV and cytology data

at different return visits in this dataset were excluded, their occurrence was rare. Furthermore, the testing procedures used in this study were highly sensitive and underwent standardized procedures for quality control. Missing values were generally considered as a separate dummy variable category in multivariate regression analyses if the variable was not measured.

In addition to the above, a third assumption for GEEs is that the outcome events do not influence subsequent exposure status. Two methods were used to reduce this affect. First, exposure status (HPV testing) was ascertained blindly without knowledge of either the previous outcome result or previous HPV result. Second, highly sensitive and objective PCR techniques were used to ascertain HPV positivity and viral burden. Therefore, the potential for false negative results was minimized. The situation may arise, however, where a prior event may serve as a predictor of a next event or viral infection through its association with an underlying pathological process initiated by a previous infection. As a result, although it may appear as if the above assumption is violated, the change in baseline risk is not considered to be a consequence of the outcome event or the biology of the disease. Furthermore, this correlation between outcome and exposure may be offset by the adjustment for relevant covariates. Therefore, in the case of an association between HPV at any given visit and repeated occurrences of SIL, the assumption should hold given that the changing profile of HPV infection is known and adjusted for in the models used on this data. Removing subjects with lesions at enrollment also reduces the potential influence of unmeasured covariates such as HPV infections before the study started.

4.5.3. Marginal models for multivariate incomplete failure time data

Since not all women have yet completed the maximum number of scheduled visits, due to insufficient follow-up time or loss to follow-up, the number of repeated observations for each individual may vary. Also, although every attempt is made to ensure quality of specimens, some cervical smears and HPV tests are inadequate and constitute missing or censored data. Therefore, time to development of multiple ordered neoplastic events was evaluated using survival analysis techniques for marginal data (Paper III) [*33;243*]. Marginal regression models adjust for correlation between observations in a way analogous to GEE. That is, the estimates are determined first by fitting the model ignoring the correlation and then correcting the variance by taking into account the clustering by individual. To look at the relationship between HPV status at repeated visits and recurrent instances of SIL, I considered several marginal models for Cox regression [*3;200;259*]. The main differences between these models exist in how they follow repeated observations and condition on repeated events.

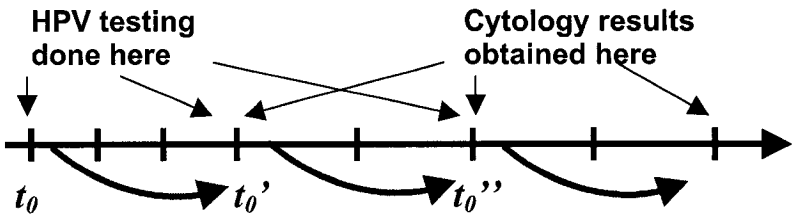
A first marginal model approach, proposed by Andersen-Gill (AG) [*3*] assumes that the repeat outcome events are independent but conditional on the covariate pattern. That is, the occurrence

of one event does not predict a second. This assumption holds provided that all relevant covariates are included. Omission of important risk factors can lead to an underestimation of the standard error (SE) of the coefficients. Preliminary HRs for occurrence of SIL (both incident and repeated events) generated by the AG model did not change much compared to the traditional time-dependent Cox proportional hazard estimates for a first SIL event published in Paper III, although the gain in precision for the AG estimates was negligible (data not shown). The lack of viable results for HSIL events with this method suggested a lack of power to estimate the hazards for different repeat HSIL events due to the fact there are few subjects with multiple such events.

Additional methods for multiple ordered data that have been proposed include the Wei-Lin-Weisfeld (WLW) model [*259*]. Based on the Cox proportional marginal hazards assumption, information from the k^{th} recurrent event time is used to estimate β_k. An alternative model proposed by Prentice, William and Peterson (PWP) [*200*] also conditions on the entire event history of subjects with recurrent SIL outcome events [*243*]. This model assumes that a subject cannot be at risk for a (k^{th}) recurrent event until the *k-1*st event has already occurred. The Cox models are stratified to allow the baseline hazard functions to differ for the covariate groups identified. Strata are assigned for each repeated time-to-event observation. A consequence of the stratification, however, is that these models require a larger number of individuals with repeated outcome events than are present in the latest version of this cohort study. Regression model fitting procedure for both SIL and HSIL outcomes failed to converge.

Based on the approaches described for the above models by Therneau and Grambsch, I created a repeated events data structure to conform to the requirements of a marginal hazards model. The objective of the analysis was to evaluate the relationship between HPV infection status at any given visit with lesion incidence at different follow-up visits while allowing for repeated observations within individuals. Observation periods consisted of two points of assessment, one for HPV exposure and the other for outcome lesions that were measured at different follow-up visits. As such, subjects could have multiple periods of observation. The duration of each observation period was dependent on the interval times between visits as defined by cohort study design, i.e., for the evaluation of HPV infection at any given visit and the occurrence of SIL 4-months later, subjects could have contributed up to three separate periods of interval observation (between enrollment and visits 4, 8 and 12 months). A layout of the data structure used in this analysis is presented in Figure S-3 and in Paper III. The primary difference between this model approach and the previously described marginal models is that instead of stratifying by event, stratification was performed on time (periods of observation). In addition, HPV status was also allowed to change for each repeated observation period using the HPV test results obtained at each index visit (t_0).

Figure S-2. Follow-up design and layout of marginal data model for repeated observations of HPV infection status at any given visit and occurrence of SIL at different interval returns



Using increasing stringency in exclusion of prevalent lesions, I refit the marginal Cox model described above on (i) all subjects regardless of baseline status by cytology, (ii) on a subset of subjects after removing prevalent cases of SIL at enrollment, and (iii) restricting for periods of observation with no lesions at the baseline visit. The latter approach removes observation periods that have a detected lesion by cytology at the beginning of the period, i.e., at each baseline (t_0) visit where HPV exposure is assessed, forcing repeated events to occur after an interval period without SIL. For the second level of restriction, women with lesions at enrollment are removed in order to exclude those women who may have entered the study manifesting a more extensive or untreated lesion. Thus, while women with incident lesions were allowed to contribute to analyses of next visit associations, these are more likely to be at the beginning of the disease process. Subjects could have had repeat events if there was an event at t_0' and at t_0''. Point estimates of RR were consistent across the three groups of subjects (see Figure A-2, Appendix D). Inclusion of prevalent lesions at each baseline visit tended to produce slightly higher point estimates away from the null compared to the other model estimates. However, differences in HR between the three models were small.

The above sensitivity analyses were also performed using GEEs. A similar stratification on observation period order and clustering on individual procedure was used according to the requirements for GEE modeling (Appendix C). Repeated events at each time period were also permitted. Correlation between repeated events was accounted for using an exchangeable correlation pattern. The consistency in point estimates for the different sensitivity analyses using varying levels of restriction also indicated that the occurrence of prior events do not affect the prediction by HPV indicating that the third assumption for GEEs holds for this data.

Point estimates derived by both methods, including the latter marginal hazards model presented above and GEE, were very similar. This was anticipated since, by design, the follow up time is stipulated for everyone (4 months, 6 months, etc) and the grouping of the risk sets over time created by the survival analyses is tight. Point estimates derived by GEE using the above selection

criteria were also very similar across sub-groups showing little difference whether subjects with prevalent lesions were included or if all baseline lesions were excluded at the beginning of each observation period (Figure A-3, Appendix D).

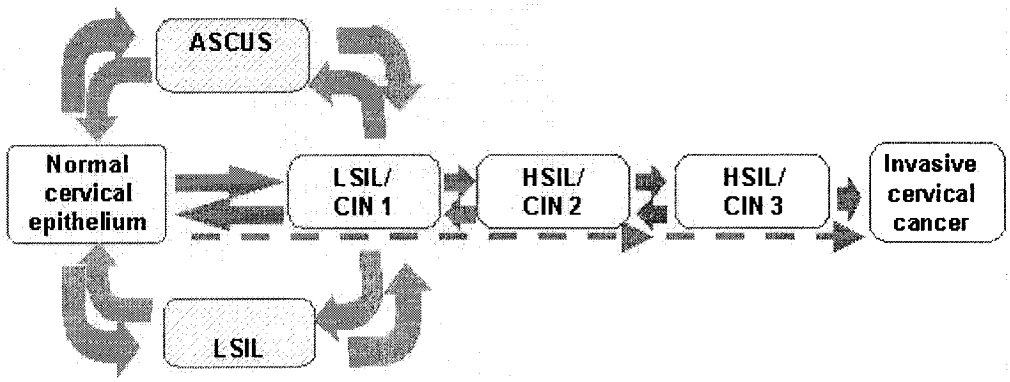
A robust variance was applied in all models with correlated data [*259*]. This approach estimates the variability of the coefficients directly from the observed correlation structure of the residuals in the sub-populations using a sandwich estimator rather than relying on the binomial or Poisson regression likelihood estimate of variation, which is coupled with equal correlations. This method should provide valid SEs even if the correlation between groups is not as hypothesized by the specified correlation structure in GEE [*271*].

4.6. Estimation of disease progression and sojourn time (Paper IV)

Sojourn time, the time spent in the preclinical detectable phase for chronic diseases, plays an important role in the design and assessment of screening programs. For the analysis of sojourn time, actuarial methods were applied to measure the average time to regression or progression within the spectrum of preinvasive lesions of cervical neoplasia. Based on the proposed natural history of cervical cancer, the possible transitions that can exist between the different lesion states will likely follow the course shown in Figure S-4. From a modeling point of view this involved modeling the length of time spent in each precancerous transition state, relating this to both constant factors such as an individual's age at entry to the study, and to dynamic time dependent factors such as HPV infection status. Given the delay in collection and recording of HPV test results and questionnaire data, the last recorded value for each was used.

The recorded dates at interview and cervical sampling for cytology were used to mark the start and end points of transitions between outcome grades. This was deemed equivalent to using mid-point values of interval periods. While duration will be somewhat overestimated due to the use of interval-censored data, this was not thought to be significant because of the short interval times (4 to 6 months) used in this study. Estimates of mean and median duration were derived by actuarial methods for cumulative survival probability (Appendix C).

Figure S-3: Potential transition states of cervical lesions



The cumulative probabilities of remaining with a lesion or progressing to a next stage were estimated by actuarial analysis [*121*] as a function of the length of follow-up for HPV type or grouped infection, among women who had no detected lesion at enrollment (Paper IV). The proportion of subjects remaining with a lesion or abnormal cytology result without regressing or progressing to a next stage at 6, 12, 18 and 24 months were calculated by life table analysis [*38*]. Cumulative risks of lesion incidence over time were also illustrated by Kaplan-Meier survival curves depicting the actuarially estimated cumulative incidence over the entire follow-up period from four months to six years post-enrollment stratified according to HPV infection status at enrollment and first follow-up return (Paper I).

In addition to the latter actuarial estimates of mean duration of preinvasive lesions an alternate formula was also used based on the epidemiologic tenet that, within a stationary population and in the absence of migration (Appendix C).

5. METHODOLOGIC ISSUES

Unlike most other cancers, in which multiple environmental, biologic, and lifestyle determinants contribute independently or jointly to carcinogenesis, cervical cancer has been shown to have a central causal agent, HPV infection. The following sections describe the general limitations and strengths of the cohort study. Methological issues related to each manuscript analysis are also discussed in the respective manuscript texts.

5.1. Selection bias

The subjects in the Ludwig-McGill cohort were recruited from a well-defined source population of residents living permanently in São Paulo city and attending a comprehensive maternal and child health program catering to low income families from neighborhoods located in the northern sector of the city (population: 12,000,000). Recruitment was done randomly without preference for age, ethnicity or sexual behavior from daily lists of outpatients in the family medicine, gynecology, and family planning clinics at the "Vila Nova Cachoeirinha" municipal hospital.

Selection criteria restricted enrollment to women at risk of HPV infection and incident cervical neoplasia, that is, they were sexually active, had an intact uterus and cervix, had not or were not scheduled for a hysterectomy or received treatment for cervical disease, and had not recently used vaginal medication. All women in the study (save one) had initiated sexual activity by the time of their first follow-up visit interview, and therefore had potentially been exposed to HPV through sexual transmission. Pregnant women and those planning to become pregnant within a year of enrollment were also excluded due to the concern that hormonal changes during pregnancy would influence the natural history of the disease. Selection of subjects and testing of samples for both HPV and cervical neoplasia was performed irrespective of age or risk of developing cancer.

Given the prospective nature of this cohort study, the opportunity for differential selection of subjects based on exposure and outcome status is reduced. Nonetheless the potential for selection bias due to loss to follow-up remains a possibility. Reports on the rate of loss to follow-up is lacking in many of the previous studies on the natural history of precursor lesions [*109;162*]. Comparison of the distribution of demographic factors for participants who dropped out of the study after the enrollment interview compared to those who remained in the study and returned for repeat interviews during the first year of follow-up showed no marked differences in distribution by age, ethnicity, sexual behavior or reproductive health characteristics (Table A-1.1 to A-1.3, Appendix D). Likewise, those who returned for second and third interviews only, returned for all four first year return visits, or remained in the study past the first year series of interviews, did not show substantial differences in distribution of demographic factors or other risk

factors for cervical cancer. To some extent, this may reflect the procedures that were implemented to encourage adherence to the follow-up schedule (Appendix A). The nurses called the patients a few days before the scheduled returns to remind them of the pending visits. Missed appointments were followed by phone calls and/or letters. These attempts at contacting subjects were repeated once a month until an appointment could be scheduled or the woman explicitly stated that she would drop out of the study. Therefore, the accrual of long-term test results presented in this study come mostly from those entering the cohort during the first two years of follow-up and who had long-term follow-up. It is unlikely, however, that subjects with delayed visits are more likely to manifest HPV or SIL. Return rates for scheduled follow-up returns were 75%, 70%, 65%, 62%, 61%, and 59% at 12, 24, 36, 48, 60 and 72 months, respectively; 11.2% of the subjects dropped out after the enrollment visit.

The point prevalence of HPV infection measured at each visit has ranged from 14.6% to 15.7%. Table M-1 shows the HPV prevalence rates among participants without lesions at enrollment for the first two years of follow-up. Excluding prevalent cases of SIL at enrollment, the prevalence of HPV among participants who had returned for the indicated visits were similar across visit returns. Testing for HPV past the first two years is currently under way.

Table M-1: HPV infection status at different scheduled returns for the first two years of follow-up

HPV status	Enrollment (N=2026)		4 months (N=1937)		8 months (N=1858)		12 months (N=1793)		18 months (N=568)*		24 months (N=1232)	
	No.	(%)	No.	(%)	No.	(%)	No.	(%)	No.	(%)	No.	(%)
Negative	1669	(82.4)	1592	(82.2)	1525	(82.1)	1500	(83.7)	475	(83.6)	1053	(85.5)
Only non-oncogenic types	129	(6.4)	142	(7.3)	142	(7.6)	121	(6.7)	36	(6.3)	76	(6.2)
Any oncogenic types	167	(8.2)	163	(8.4)	144	(7.8)	144	(8.0)	52	(9.2)	98	(8.0)
Beta-globin negative	61	(3.0)	40	(2.1)	47	(2.5)	28	(1.6)	5	(0.9)	5	(0.4)

* N=568 represents only the portion of participants who had not reached the 18 months follow-up point before the introduction of semester visits in 1997.

5.2. Information bias

Due to the prospective nature of the study, recall bias with respect to ancillary factors such as sexual activity, reproductive health and diet is unlikely. Subjects were unaware of their HPV infection status or cytological results at any given visit reducing the potential for subject's disease status affecting recall of co-factor information. Subjects presenting with HSIL were informed by phone after their interview when the cytopathologists' reports were received. In addition, while data on sexual activity are based on recalled data, reports of current sexual partners referred to recent (within 4 or 6 months) behavior. Likewise, history of OC use and smoking refers to both

lifetime exposure and recent behavior. As a result, misclassification of recent exposures due to recall error should be small, and non-differential. Even if residual effects were to remain after accounting for HPV status, such markers of sexual activity should not be considered as confounders in the natural history of cervical cancer.

In none of the interviews was ethnicity directly asked; the nurse checked the category most closely reflecting the ethnic group of the respondent among the main racial groups in Brazil (white, mulatto, black, oriental, or Indian descent) that are widely used in official census statistics. Early in the study nurses were trained to differentiate among these categories based on physiognomic criteria used by social workers. The reasons for not collecting information on self-reported race or ethnicity are as follows: First, many women in Brazil deem questions about racial background offensive. Asking the woman directly or even showing a list of possible choices would have disrupted the climate of non-judgmental rapport established between the nurse and the respondent. This could have made a respondent less cooperative and frank in completing the interview. Secondly, previous experience in conducting epidemiologic studies in Brazil indicate that a respondent's self-appraisal of race or ethnicity is frequently wrong [*251*] because many black or mulatto women identified themselves as being white.

Testing of HPV DNA in cervical specimens and cytopathology evaluations were done blindly with respect to each other and with respect to results from prior assessments for the same women. Differential information bias with respect to the main risk factors (i.e. HPV status and viral load) is improbable as the collection of the markers of HPV infection was done using objective laboratory methods.

5.2.1. Misclassification of HPV status

Non-differential misclassification can be more or less of a concern depending on what laboratory method of testing for HPV is used. While this was a more serious problem with first generation detection methods [*65;67*], the method used in the Ludwig-McGill study is highly sensitive and specific. The PGMY09/11 consensus primer set targets conserved sequences in the L1 gene and can amplify a wide spectrum of HPV types [*36*]. This method is capable of detecting a very small quantity of DNA. Stringent standardized quality control protocols were employed in the study to avoid cross-contamination. PGMY primers are the most recent PCR primer set currently used in epidemiological studies. The increased molecular sensitivity and specificity of these primers over existing methods has been demonstrated [*85*] reducing the likelihood of exposure misclassification.

The occurrence of residual misclassification was minimized using a consensus primer PCR detection method. The method can discriminate between 40 different HPV types including recent types 81, 82, 83, 84, CP6108 and IS39. A generic probe was also included in all HPV assays. Amplified products that hybridized with the generic probe but with none of the type-specific probes were tested further by RFLP analysis of the L1 fragment [*12*] to distinguish among unknown HPVs. Positive specimens for remaining unknown types were also included in all analyses as non-oncogenic types.

Despite the considerable improvement in laboratory techniques for the detection of HPV DNA, there is one important source of misclassification of HPV exposure status that cannot be readily corrected for by methodological advances: that caused by the fluctuation in viral load over time. Infections with low viral load may be labeled erroneously as HPV negative, and those with mildly productive transient infections at the time of testing will be classified as HPV positive. Fluctuations in viral load and specimen cellularity may also affect the comparison of risk factor profiles between HPV-positive and HPV-negative CIN case subjects by influencing the false-negative rate in specific groups. HPV infections may be cleared and exist only transiently in cervix epithelium or may become latent and exist in the basal squamous epithelium below the threshold of detection of even PCR [*71*]. Making multiple measurements of HPV over time therefore reduced the potential for such misclassification. However, the definition of a persistent infection was based on detecting viral DNA of the same taxonomic type using a consensus PCR protocol. Fluctuations in viral load below the detection threshold of PCR may have caused some cases of persistent infection to be misclassified as transient due to false negative results. In addition, it is impossible to ascertain via HPV DNA detection alone if test positivity is equated with true (active or latent) viral infection. On the other hand, it is reassuring to note that PCR typing alone may be sufficient for defining persistent infections. Longitudinal testing for molecular variants of HPVs 16 and 18 (which was performed in this study), while providing an enhanced level of taxonomic detail for ascertaining true persistence [*73*], indicated that persistently detected HPVs 16 or 18 were of the same molecular variant in each case [*252*]. This observation suggests that persistent detection of the same viral type may truly represent a persistent infection.

5.2.2. Misclassification of HPV viral load

The possibility of misclassification is more of an issue for the detection of viral load, which is based on an older method (GP5/6) of HPV DNA detection. While confirmation of HPV presence was established by PGMY09/MY11 in this study, measurement of viral burden may be more prone to error. The higher threshold value of the GP5/6 method may therefore have resulted in a higher rate of false negatives or underestimated the quantity of viral copies for samples with low viral loads. In the present study, these values are grouped into categories to increase precision in

risk estimates though the misclassification likely contributed to attenuation of the RR association. Other methods of viral load measurement exist including both non-quantified methods (HC), discrete definitions based on dichotomous test results for HPV detection using two tests of varied sensitivity [*107*], and true quantification methods such as real-time PCR. Earlier studies have used estimates of level of chemoluminescence released by bound antibodies in the HC assays to estimate viral burden per sample volume. Other methods of true quantification are more recent, including a real-time PCR method with Taqman® [*135*]. This latter method uses HPV 16 specific hybridization probes for L1, E2 or E6 with quantitative PCR that employs the 5'-exonuclease assay and real-time detection of the accumulation of fluorescence. The viral load quantification (copies/µl) is achieved by comparison of the threshold cycle number (C_t) for the sample (which represents the PCR cycle number at which the fluorescent signal exceeds the baseline signal) with the C_t for a range of samples with known starting copy numbers. Quantification in the Ludwig-McGill study has the advantage of being based on the L1 gene, which is more likely to give a reliable measure of viral load than methods based on E2 or E1, since this gene (L1) is preserved throughout the natural history of cervical neoplasia, with few exceptions.

Viral load measurements are also dependent on the cellular composition of the cervical specimens the testing is based on. As a result, even though the viral load quantification methods can produce a collective estimate of the number of virons per cell, there is no way to determine whether this measure represents few cells with high viral titer or many cells with lower viral numbers. Furthermore, the specimens were obtained from cervical smears, which may contain very few cells representative of a high-grade lesion amidst a large number of normal or even productively infected cells, in which episomes (extra chromosomal non-integrated DNA) are expected to predominate over integrated viral genomes, and replicate actively. Moreover, at least in the case of ASCUS and CIN2, misclassification can occur that additionally complicates the interpretation. Although the biasing effect of HPV misclassification may not be completely ruled out with new detection methods, the ultimate effect under conditions of perfect measurement of outcome (CIN versus non-CIN) would be to bias the estimates of RR to unity. The incremental effect of exposure misclassification in epidemiological studies of HPV and CIN has been demonstrated [*68*].

5.2.3. Misclassification of cytology

While the influence of the false positive rate on incidence and prevalence estimates is important for public health, for clinical management the occurrence of false negative results by cytology is a concern. Due to the low prevalence of the cervical neoplasia in most populations, slight increases in false positive rate by cytology compared to histology can result in elevated prevalence and incidence estimates for the disease. While incidence rates by cytology can range from 400 to

4000 per 100,000 woman-years, rates determined by histology are a hundred-fold lower [*71*]. However, follow-up diagnostic methods employed after a positive cytological assessment can also entail their own concerns for the patient with respect to medical, financial and emotional implications. Colposcopy and biopsy procedures subject the patient to discomfort, carry the risk of complications including bleeding and inflammation, and can place a substantial burden on the local health care system [*55;234;270*].

The merit of using tissue samples for outcome ascertainment instead of Pap smears is debatable. A primary concern of using biopsies to study disease progression is the potential of removing completely the lesion under evaluation and thus changing the natural history of the disease [*206;269*]. Among those with HSIL, the effect of biopsy on rates or regression was evaluated separately (Paper IV). While biopsy results represent the gold standard for diagnosis in cervical cancer, it would have been ethically untenable to follow-up women with histology from biopsies in such an epidemiological study.

It has been postulated that unless tissue extraction is done with the intent of sampling from the whole area of the endocervix and is performed using Schiller staining and colposcopy to direct biopsies, the gain in sensitivity or specificity will be negligible [*162*]. However, increases in rate of progression and reduction in sojourn time may be observed due to the delayed detection of lesions by cytology.

5.3. Confounding

In other epidemiological models of cancer, most identified risk factors are neither necessary nor sufficient causal factors of disease and therefore are usually investigated in conjunction with other covariates to isolate their independent role. This is not the case for HPV, which has been identified as a "necessary cause" of cervical cancer [*16*]. As studies are able to more precisely measure HPV, many established risk factors for the cervical neoplasia, including sexual behavior and smoking have been explained by HPV infection status [*168*]. Nonetheless, I investigated the potential effect of confounding by education, smoking history, ethnicity, feminine hygiene practices, parity, oral contraceptive use, condom use, menopausal status, age at first intercourse, number of sexual partners, history of sexually transmitted diseases, and occurrence of other genital infections that could influence the development of SIL using a change in point estimate cutoff of 10% [*148*]. Collection of information on these co-factors was done by comprehensive questionnaires at enrollment. I adjusted *a priori* for age using four groups contrasted as dummy variables (18-24, 25-34, 35-44, and 45-60 years), and two groups for ethnicity (white and non-white); no other variables were identified as confounders. To satisfy requirements for the first journal submission (Paper I), adjustment for ethnicity was included using the dichotomous

variable. Participants were predominantly white Hispanics. Of those in the non-white category: 12.4% were black, 20.9% mestiza, and 0.8% other. However, this covariate is not a likely candidate for confounding of the relationship with persistent HPV or viral load as no independent association with risk of cervical neoplasia exists. Differences in incidence rates for cervical cancer observed between geographic regions and ethnic groups taken from registry data are likely explained by differences in health care systems, intensity of screening programs, and exposure to HPV [*168*]. Nonetheless, RR estimates across studies from populations of varied ethnicity or geographic location have shown consistent risk associations for HPV with CIN lesions [*16;71;110*].

It is conceivable that increased sexual activity with a plurality of partners and higher levels of cigarette smoking both may have facilitated the establishment of more productive lesions, which are less likely to be missed in a single testing opportunity. Burger et al. [*26*] found that women with HPV-positive CIN had more sexual partners and tended to smoke cigarettes more than HPV-negative patients with CIN. Under the sexually transmitted disease model, markers of sexual activity serve as intermediate variables on the causal pathway of cervical neoplasia. The statistical association between these factors and risk of cervical neoplasia should only be present in analyses that do not adequately control for HPV status [*67*]. This situation is less likely in a study involving repeated assessments of exposure such as this. Long-term smoking has been found to increase the risk of CIS in younger women [*267*] indicating that an interactive association between metabolites of smoking and HPV DNA on the risk of SCC could exist. This association, however, was not the focus of the project and was not investigated.

5.4. Effect modification

To assess effect modification by age, incidence rates and RRs of SIL were evaluated for younger women and older women separately using different cut-offs of 30 and 35 years to form two age groups. Incidence rates of first episode SIL (any-grade) were highest in women between 18 and 24 years (2.44 per 1000 women-months, 95%CI: 1.7-3.2) and lowest for women 35-44 years of age (1.01 per 1000 women-months, 95%CI: 0.6-1.4) and then increased marginally to 1.08 per 1000 women-months (95%CI: 0.4-1.7) in women 45-60 years. There was a downward trend in incidence rates of persistent SIL that continued to the oldest age group, decreasing gradually from 0.58 per 1000 women-months (95%CI: 0.2-0.9) for 18-24 year-olds to 0.19 per 1000 women-months (95%CI: 0.1-0.5) for 45-60 years olds (p-value for trend = 0.037).

We also investigated whether RRs of first episode SIL differed for younger and older women by stratifying on age at enrollment. Separate analyses were performed for women ≤30 and >30 years of age. Higher RRs of incident HSIL were observed for older women with persistent

oncogenic infections at the first two visits (RR=29.35; 95%CI: 7.3-118.0) compared to younger women (RR=5.26; 95%CI: 1.0-27.1). Women 30 years and younger who harbored persistent infections for oncogenic HPV types during the first two visits were more likely, albeit non-significantly, to develop persistent SILs (RR=19.87; 95%CI: 5.0-79.5) than older women (RR=13.85; 95%CI: 3.7-52.4). There was no effect modification between age and HPV infection, regardless of outcome definition. Using a different cut point (25 years) for defining the two age strata yielded similar conclusions.

5.5. External validity

The Ludwig-McGill study was started in a population of low income women in São Paulo, Brazil, one of the highest risk areas worldwide for cervical cancer. However, no selection was performed based on risk level (i.e. restricting to women with high rates of STD infection, prevalent HPV infections or LSILs). Women with prevalent lesions at enrollment were excluded from the analyses. The age range for women in the study was also wide and well distributed from 18 to 60 years of age. The mean age at enrollment was 33.1 years (median age: 33) with most women (64.7%) being white and of European ancestry. Over 81% of the women had not attended high school and 82% were married or living in a common law relationship. These proportions are similar to those in other studies performed in Brazil [*251*].

Among HPV types, HPV 16 is the most prevalent oncogenic type in cervical cancers [*17*]. While the magnitude of the associations may vary between studies, the estimates of rates and RR from this study should be generalizable to other populations. However, given the unique algorithms used to define HPV over time and the time-dependent relationship with lesion incidence, these estimates need to be confirmed by other international studies for such inference to be substantiated.

5.6. Chance

The size of the cohort and prevalence of HPV exposure yielded good statistical power in most of the analyses. Nonetheless, the finding of statistically significant results in this study, based on two-sided tests at the 0.05 level, do not mean chance can be ruled out. The level of precision in estimates was represented by 95% confidence intervals wherever possible while p-values for two-sided tests were used to indicate statistical trend. Tests for trend were primarily performed for RR estimates of ordinal categories of viral burden to evaluate dose-dependency. The results for RR of incidence of SIL were significant and consistent across studies for the most part, although precision decreased for estimates of incidence of rarer severe outcomes such as HSILs.

5.7. Sample size and power

The initial objectives of the Ludwig-McGill cohort study were to investigate the natural history of cervical neoplasia and the factors, such as HPV persistence, involved in the risk of lesion incidence in a high-risk population. A sample size of 2400 was calculated to provide ample (80% or more) statistical power to address these objectives. The development of this longitudinal study for my doctoral thesis was introduced after recruitment was stopped for the primary study. Therefore, while the framework of the cohort could permit evaluation of these additional objectives for this thesis, statistical power was limited to the available sample size for the original cohort.

5.7.1. HPV persistence and SIL incidence

Expected rates of persistent HPV infection using the molecular variant definition were available only from preliminary analyses of the Brazilian cohort. Estimates of the cumulative rate of HPV persistence were 17%, for algorithm 1, and 5%, for algorithm 3 above. All cases classified as persistent based on algorithm 3 in the study had only one and the same molecular variant in all specimens over time [*73*]. Using these two figures as a range (5%-20%) and considering the first of the specific aims (assessment of HPV persistence), the cohort size required included 1821 subjects to measure rates of HPV persistence with adequate precision (an absolute error of at most 2%) at a confidence level of 95% (5% alpha error). The statistical analyses in this thesis are based on 2462 eligible subjects who have enrolled in the study since initiation the cohort of whom 1883 have already returned for 4 or more follow-up visits.

Of primary interest in this study in accordance with objective 1 was the attainable study power with respect to the analysis of cytologic endpoints. Assuming a 2-year risk for developing HSIL among low risk women (HPV negative) at 2% (from preliminary analyses), and also assuming that ratios of ASCUS:HSIL and LSIL:HSIL are 5:1 and 4:1, respectively [from the National Breast and Cervical Cancer Early Detection Program (NBCCEDP) *32*], the cumulative frequency of subjects with incident cytological abnormalities in this study (assuming 1600 subjects with complete follow-up at 5 years) would be about 44%, or 700 women, considering only the women who are cytologically negative at enrollment. Of these, approximately half, or close to 350 women should be SILs, as per the expected ratios of the NBCCEDP [*32*]. Based on these assumptions, for analyses of HPV persistence as a determinant of incident SIL, the statistical power was greater than 95% and 99% to detect a doubling in risk of incident SIL, for persistence assumed at 5% and 20%, respectively. Power was greater than 95% to detect a mere 50% increase in risk of SILs (RR=1.5) for a determinant present in 25% of the subjects' histories. Expectedly, however, power was somewhat less for the analysis of HSIL as an outcome (5% cumulative rate), i.e., 85% for persistence estimated at 20% and a RR=2, and 70% for persistence at 5% and a RR=2.5.

5.7.2. Progression rates and sojourn time

Previous studies looking at risk of cervical cancer and progression to invasive carcinoma have required large numbers of patients and long-term follow-up to allow for the time needed for development of cancer. As a result these studies have relied on registry data and retrospective follow-up [*109;162;266*]. Using preinvasive lesions as surrogate measures of cancer risk, a smaller sample size and less follow-up would be needed to measure differences in time to occurrence of an outcome event with a similar degree of statistical power.

Accrual rates as of 2002 indicated a cumulative incidence of 115 SIL events of which 19 are of HSIL among women with no prevalent lesions at enrollment. It is hypothesized that sojourn time should be shorter for women harboring HPV infections than for those with negative or transient infections. Preliminary results indicate that among women who develop SIL, those with persistent HPV infections (N=18) developed SIL two months sooner than those with negative HPV results. This difference in sojourn time (using log transformed data) can be found to be significant at p=0.05 (two-sided) with over 88.5% power. Currently comparison of sojourn time between HSIL cases with persistent infections and those with only transient or no HPV infections can be performed with 47% power at the 5% (two-sided) level. We would require 8 instances of HSIL among women with HPV persistent infections to detect the same difference in sojourn time at 80% power.

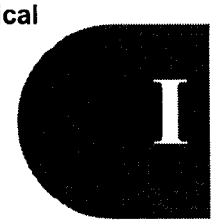
6. MANUSCRIPT RESULTS AND DISCUSSION

6.1. Preface to Manuscript I

Positive predictive values of HPV testing are low in most asymptomatic women because of the relatively high prevalence of subclinical HPV infection in the general population [*71*], particularly among sexually active women [*27*] and those less than 30 years of age [*9;205*]. Most of these infections are transient and are probably of little significance. The concern resides, however, in the small proportion of women who harbor persistent HPV infections. While the central role of certain HPV types in the causal pathway of cervical cancer is generally accepted, it is seen as a necessary but not sufficient factor in the natural history of the disease [*16;177*]. This is the first study to appropriately place HPV persistence in a time perspective.

Previous studies evaluating the association between HPV persistence and incidence of CIN have been limited in study design with respect the intervals between repeated sampling, length of follow-up and precision with which they could determine viral persistence. Collection of a single cervical specimen at the time of enrollment in a cohort study or at the time of diagnosis of CIN or of invasive cervical cancer in a case-control study provides little assurance that the laboratory determination of HPV positivity of that specimen accurately reflects the relevant past exposure to HPV infection that the subject may have had. Furthermore, a biological model for HPV carcinogenesis from an enduring HPV infection implicates the necessity for evaluation of a type specific infection. This requires testing by type-discriminate HPV assays and frequent repeated assessments of infection status. I therefore investigated the risk of developing cervical lesions in the presence of persistent cervical human papillomavirus (HPV) infection using a highly sensitive HPV detection method for data from the Ludwig-McGill cohort study. The analysis in the following manuscript represents the first of a series of investigations of the longitudinal relationship of HPV and CIN and is based on data collected until June 2000.

6.2. Manuscript I: Persistent human papillomavirus infection as a predictor of cervical neoplasia



(Running title: HPV persistence and CIN)

Nicolas F. Schlecht MSc[1,2]; Sophie Kulaga MSc[2]; Juliette Robitaille BSc[3]; Silvaneide Ferreira BSc[4]; Monica Santos BSc[4]; Romulo A. Miyamura BSc[4]; Eliane Duarte-Franco MD, MPH[2]; Thomas E. Rohan MD, PhD[2,5]; Alex Ferenczy MD[3]; Luisa L. Villa PhD[4]; Eduardo L. Franco PhD[1,2]

Author Affiliations:

[1] Departments of Epidemiology & Biostatistics, [2] Oncology, and [3] Pathology, McGill University, Montreal, Quebec;

[4] Ludwig Institute for Cancer Research, São Paulo, Brazil; and

[5] Department of Epidemiology and Social Medicine, Albert Einstein College of Medicine, Bronx, NY

ABSTRACT

Context Human papillomavirus (HPV) infection is believed to be the central cause of cervical cancer, although most of the epidemiological evidence has come from retrospective, case-control studies that do not provide information on the dynamics of cumulative or persistent exposure to HPV infection.

Objective To assess the risks of cervical neoplasia related to prior persistent HPV infections.

Design and Setting Longitudinal study of the natural history of HPV infection and cervical neoplasia in women residing in the city of São Paulo, Brazil, which was conducted between November 1993 and March 1997 and involved repeated measurements of HPV and lesions with follow-up until June 2000.

Participants A total of 1611 women with no cytological lesions at enrollment and HPV test results available from the first 2 visits.

Main Outcome Measure Cervical specimens taken for Papanicolaou cytology and HPV testing every 4 months in the first year and twice yearly thereafter. Incident cervical cancer precursor lesions ascertained by expert review of all cytology smears.

Results The incidence rate of squamous intraepithelial lesions (SILs) was 0.73 per 1000 women-months (95% confidence interval [CI], 0.5-0.9) among women free of HPV at the 2 initial visits and 8.68 (95% CI, 2.3-15.1) among women with HPV type 16 or 18 infections persisting over both visits. Relative to those negative for HPV oncogenic types at both initial visits, the relative risk (RR) of incident SIL was 10.19 (95% CI, 5.9-17.6) for persistent infections with any known oncogenic HPV types. The equivalent RR of incident high-grade SIL was 11.67 (95% CI, 4.1-33.3). The RRs of lesions were considerably higher for persistent infections with HPV type 16 or 18.

Conclusion A strong relationship exists between persistent HPV infections and SIL incidence, particularly for HPV types 16 and 18.

** This manuscript was published in JAMA. 2001;286:3106-14*

INTRODUCTION

Human papillomavirus (HPV) infection is believed to be the central cause of cervical cancer, although most of the epidemiological evidence has come from retrospective, case-control studies, [*110*] which do not provide information on the dynamics of cumulative or persistent exposure to cervical HPV infection. The social and economic implications of cervical cancer for public health programs worldwide and the recent interest in the development of HPV vaccines have compelled the initiation of prospective, long-term multidisciplinary studies of the natural history of cervical cancer to investigate the role of HPV infection in the development of preinvasive cervical lesions. To date, several studies have reported results on the prospective relationship between HPV infection and the incidence of squamous intraepithelial lesions (SILs) [*106;131;138;171;174;255;262;268*]; however, only 2 of these studies have reported on HPV persistence and subsequent SIL incidence over an extended period of time. [*171;262*]

In 1993, we began a longitudinal investigation (the Ludwig-McGill Cohort Study) of the natural history of HPV infection and cervical neoplasia in the city of São Paulo, Brazil, a high-risk area for cervical cancer. Women enrolled in this study were followed up over a period of several years at scheduled return visits during which they were screened concurrently for cervical lesions and HPV infection. This allowed us to assess the risks of SIL related to prior cumulative and persistent HPV positivity. We were also able to focus on persistence of HPV types 16 and 18, which have been linked with increased incidence of high-grade lesions [*131;138;262;268*] and higher probability of lesion persistence. [*107*]

METHODS

Subject Recruitment and Follow-up

Since 1993, we have carried out a cohort study involving repeated measurements on women attending a comprehensive maternal and child health program catering to low-income families living in neighborhoods located in the northern sector of the city of São Paulo, Brazil (population, 12 million). Using rosters of outpatients in the family medicine, gynecology, and family planning clinics at the Vila Nova Cachoeirinha municipal hospital, 2 nurses specifically trained for the study selected a systematic sample of 4990 women to be approached for interview. Of these, 3589 initially met the eligibility criteria, were given a detailed overview of the study, and were invited to participate. Between November 1993 and March 1997, a total of 2528 women were enrolled into the study, representing a response rate of 70.4%. Another 66 women who did not fit the eligibility criteria were excluded after enrollment. Subjects entered the study only after giving signed informed consent. The study protocol was approved by institutional ethical and research review boards of the participating institutions in Canada and Brazil.

Women were eligible to participate if they (1) were between 18 and 60 years of age; (2) were permanent residents of São Paulo (city); (3) were not currently pregnant and had no intention of becoming pregnant during the next 12 months; (4) had an intact uterus and no current referral for hysterectomy; (5) reported no use of vaginal medication in the previous 2 days; and (6) had not had treatment for cervical disease by electrocoagulation, cryotherapy, or conization in the previous 6 months. In addition to these criteria, women were considered ineligible if they were not interested in complying with all scheduled returns, at least for the subsequent 2 years.

All participants were seen every 4 months in the first year (0, 4, 8, and 12 months) and twice yearly thereafter. Delays in returning for a given appointment were allowed, with information and specimens collected during any post due visits being assigned to the delayed follow-up return, which precluded the occurrence of missing interval visits. Cervical specimens were taken for Papanicolaou cytology and HPV testing at every visit. An in-person interview was also performed at enrollment to collect information on risk factors for HPV infection and cervical neoplasia, including sociodemographics, reproductive health, sexual practices, smoking, and diet. For the analyses reported here, follow-up continued until June 2000, the development of SIL, death, or loss to follow-up, whichever occurred first. A detailed description of the design and methods of the study has been published. [*62*]

Cervical Cell Specimens

A cervical Papanicolaou smear was performed using an Accelon biosampler (Medscand, Inc, Hollywood, FL) to collect a standardized sample of ectocervical and endocervical cells. After the smear was prepared on a glass slide and fixed in 95% ethanol, the sampler containing the exfoliated cells was immersed in a tube containing Tris-EDTA buffer (pH 7.4) and agitated to release the cells. Samples were then sent to the laboratory at the Ludwig Institute for Cancer Research (São Paulo) for storage and HPV testing. The cervical smear slides were read locally and then shipped to Montreal, where they were coded and read specifically for the study by an expert cytopathologist (A. F.) who was unaware of any other results from the subjects. Cytopathology reports were based on the Bethesda system for cytological diagnoses. [*231*] For the purpose of this analysis, the following categories were used: within normal limits or benign cellular changes (normal); atypical squamous cells of undetermined significance (ASCUS) or atypical glandular cells of undetermined significance (AGUS); low-grade SIL (LSIL); and high-grade SIL (HSIL); or cancer. All detected events of HSIL were referred for colposcopic follow-up and biopsy if required according to National Institutes of Health approved guidelines.

HPV DNA Detection

DNA was extracted from all cervical specimens using digestion with 100-µg/mL of proteinase K for 3 hours at 55°C, and the DNA samples were then purified by spin column chromatography. Cervical specimens were tested for the presence of HPV DNA by a previously described polymerase chain reaction (PCR) protocol amplifying a highly conserved 450–base pair (bp) segment in the L1 viral gene (flanked by primers MY09/11). [*10;105*] Typing of the amplified products was performed by hybridization with individual oligonucleotide probes specific for 27 HPV genital types. [*105*] The PCR amplification products that hybridized with the generic probe but with none of the type-specific probes were tested further by restriction fragment length polymorphism analysis of the L1 fragment [*12*] to distinguish among unknown HPVs. To verify the specificity of the hybridizations, we included more than 30 type-specific positive controls in all membranes. To check the integrity of the host DNA material extracted from the specimens, assays also included an additional set of primers (GH20 and PC04) to amplify a 268-bp region of the $^{\beta}$-globin gene. [*10*] All HPV assays were done blindly on coded specimens with no identification linking specimens from the same woman.

Statistical Analysis

Lesion incidence rates were calculated over the accrued women-months of follow-up according to the HPV infection status determined at the enrollment and first follow-up visits combined to ascertain initial persistence of HPV infection. Human papillomavirus types were grouped by oncogenic potential: nononcogenic HPVs included types 6/11, 26, 32, 34, 40, 42, 44, 53, 54, 55, 57, 62, 64, 66, 67, 69, 70, 71, 72, 73, 81, 82, 83, 84, CP6108, IS39, and other unknown types; oncogenic HPV types included 16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, 59, and 68, based on an expanded classification of Bauer et al. [*9*] Stratification by HPV oncogenicity was hierarchical and based on mutually exclusive categories. We also treated HPV types 16 and 18 separately in some analyses. Following a conservative approach to exposure classification, subjects with invalid HPV test results at any of the visits were excluded from the analyses.

We analyzed the risk of postenrollment occurrence of SIL as an incident finding in relation to HPV infection status at enrollment and during the first 2 visits. The longitudinal relative risks (RRs) and respective 95% confidence intervals (CIs) of incident cervical lesions over time were modeled by Cox proportional hazards regression and graphically represented by Kaplan-Meier curves depicting the actuarially estimated cumulative incidence over the entire follow-up period from 4 months to 6 years postenrollment. Time to event was measured from enrollment to the first instance of a lesion event or to the last recorded return visit date for censored subjects. We investigated the potential effect of confounding by marital status, education, smoking history, feminine hygiene practices, parity, oral contraceptive use, condom use, menopausal status, age at first intercourse, number of sexual partners, history of sexually transmitted diseases, and

occurrence of other genital infections that could influence the development of SIL using a change in point estimate cutoff of 10%. We adjusted a priori for age using 4 groups contrasted as dummy variables (18-24, 25-34, 35-44, and 45-60 years of age) and for ethnicity (white and nonwhite); no other variables were identified as confounders and so were not included. Incidence of SIL was estimated both by lesion severity (first confirmed SIL event of any grade and first occurrence of HSIL) and by persistence of SIL for 2 or more consecutive visits allowing for at most 1 negative intermediate visit. All incident cases of SIL were compared with subjects who had no detected lesions or ASCUS during the entire period of follow-up. Prevalent cases of lesions detected at enrollment were excluded from all longitudinal analyses.

Extending the period of observation of HPV status to 3 visits, we also assessed the effect of loss of a persistent HPV infection. Subjects with persistent HPV positivity over 3 visits were compared with subjects persistent for the same HPV types at the first 2 visits only. Adjusted odds ratios (ORs) of incident SIL events occurring within 2 years of assessment of HPV status at the third visit were evaluated by unconditional logistic regression using subjects with no infections at all 3 visits as the reference group.

We performed tests of trend by including categorized risk factors as ordinal variables in the multivariate models. Relative risks of incident SIL events by HPV infection status during the first 2 visits were compared by age group, stratifying for subjects 30 years of age and younger vs older than 30 years at enrollment. Interaction between age and HPV was assessed by comparing multivariate models assuming independence of effects to the same models further incorporating a cross-product term for interaction using log likelihood ratio tests based on the χ^2 distribution with *df* equal to the number of parameters of interest. All statistical analyses were performed using the statistical program SPSS, version 10.0 (SPSS, Chicago, Ill).

RESULTS

Subject Characteristics

At the time of analysis, valid HPV typing results (excluding β-globin–negative samples) were available for the enrollment and up to 3 follow-up samples of 1862 women. Among the 1791 women with typing information at enrollment, 286 (16.0%) were found to be positive for at least 1 HPV type. Table I-1 shows the descriptive statistics on the primary factors that could influence the relationship between HPV persistence and cervical lesion incidence for the 1789 women with HPV test results and a complete questionnaire at enrollment. All women in the study (except 1) had initiated sexual activity by the time of their first follow-up visit interview and therefore had potentially been exposed to HPV through sexual transmission. The majority of women had only 1 to 2 partners in their lifetime. The mean age at enrollment was 33.1 years (median age, 33

years). The actuarial proportions of women who have been compliant with all scheduled follow-up visits were 89% at 12 months, 84% at 24 months, 79% at 36 months, 74% at 48 months, and 69% at 60 months.

Forty-one women were found to have cervical lesions at enrollment, and smears from 4 women were deemed inconclusive or were lost. Among women free of lesions at enrollment, the mean age at diagnosis of a first SIL was 32.7 years (SD, 8.9), whereas the mean age at first diagnosis for HSIL was 32.1 years (SD, 9.3).

Incidence Rates of SIL Events

Table I-2 shows the frequencies and incidence rates of SIL by HPV infection status during the first 2 visits (including enrollment visit) for 1611 women with valid HPV test results at both initial visits and no cytologically detected lesions at enrollment. Subjects with invalid HPV test results at either of the 2 visits were excluded from the analyses. Rates of any-grade SIL were high among women testing positive for oncogenic HPV types at enrollment and higher for those who tested positive for either HPV types 16 or 18 (data not shown). These rates further increased when infections persisted to the second visit for the same HPV types. For the most part, the patterns were similar for HSIL and persistent SIL. No HSIL or persistent SIL cases were observed for women testing positive only once for HPV type 16 or 18.

Rates of any-grade SIL were highest in women between 18 and 24 years of age (2.44 per 1000 women-months; 95% CI, 1.7-3.2) and lowest for women 35 to 44 years of age (1.01 per 1000 women-months; 95% CI, 0.6-1.4), then increasing marginally to 1.08 per 1000 women-months (95% CI, 0.4-1.7) in women 45 to 60 years of age. The downward trend in incidence rates of persistent SIL continued to the oldest age group, decreasing gradually from 0.58 per 1000 women-months (95% CI, 0.2-0.9) for 18- to 24-year-olds to 0.19 per 1000 women-months (95% CI, 0.1-0.5) for women 45 to 60 years old (*P* value for trend = .04).

Cumulative Risks of SIL Events

We evaluated the cumulative incidence of cytologically detected SIL over time in women free of lesions at enrollment. Figure I-1 illustrates the cumulative risk of any-grade SIL as a function of HPV infection status at enrollment alone (Figure I-1A) or during the first 2 visits including the enrollment visit (Figure I-1B). Subjects with oncogenic HPV infections at enrollment were more likely to develop SIL compared with those with nononcogenic infections or those who were HPV negative (Figure I-1A). The cumulative risk was somewhat more pronounced for women with HPV types 16 and 18 at entry compared with those who had other oncogenic types, but there was substantial overlap between the 2 curves. Persistent detection of HPV types 16 or 18 at the

enrollment and first follow-up visits was associated with a greater cumulative incidence of SIL compared with persistent infections with other types or transient infections. The cumulative detection of SILs among women with both initial visits positive for HPV types 16 or 18 approached 40% after 4 years (Figure I-1B).

Relative Risks of First Instance of SIL Events

Using Cox regression we estimated age- and ethnicity-adjusted RRs of a first instance of any SIL, HSIL, or persistent SIL over the entire period of follow-up among the 1611 women free of lesions at enrollment (Table I-3). The highest RRs for any SIL and HSIL were observed for women testing positive for oncogenic HPV types at enrollment and first follow-up visit compared with women testing negative for any HPV type at both visits. These RRs increased slightly when restricted to women with persistent infection with HPV types 16 or 18. Relative risks tended to be higher for persistent SIL compared with any SIL events, when considering persistence for oncogenic types except 16 or 18. Subjects with missing HPV test results at either visit, grouped as a separate category, showed slightly elevated RRs compared with the referent group only for any SIL (data not shown).

Human papillomavirus persistence was also defined using a less stringent method that grouped types by level of oncogenicity rather than by taxonomic classification. We distinguished subjects positive for oncogenic types at both enrollment and first follow-up visit from those displaying persistent infections with the same type at both visits. Although the RR of SIL was high for women with 2 positive visits with different oncogenic types, no cases of HSIL or persistent SIL were observed. Comparatively, elevated RRs for any SIL were also observed for women with transient infections involving an oncogenic HPV type at one of the visits and a nononcogenic one at the other. Most subjects (64/68) with oncogenic types detected at both visits, however, consisted of persistent infections with the same HPV types. We were also able to distinguish subjects with transient infections but positive at both visits for different HPV types. Twenty-five subjects were classified as having repeatedly positive visits, with one of them revealing an oncogenic type.

We also investigated whether RRs differed for younger and older women by stratifying on age at enrollment. Separate analyses were performed for women 30 years of age and younger and older than 30 years. Higher RRs of HSIL were observed for older women with persistent oncogenic infections at the first 2 visits (RR, 29.35; 95% CI, 7.3-118.0) compared with younger women (RR, 5.26; 95% CI, 1.0-27.1), but this difference was not significant. Women 30 years old and younger who harbored persistent infections for oncogenic HPV types during the first 2 visits were more likely, albeit nonsignificantly, to develop persistent SILs (RR, 19.87; 95% CI, 5.0-79.5) than older

women (RR, 13.85; 95% CI, 3.7-52.4). There was no interaction between age and HPV infection, regardless of outcome definition. Using a different cut point (25 years) for defining the 2 age strata yielded similar conclusions (data not shown).

Risk of SIL Events with Long-term HPV Infection

To measure the effects of longer-term persistence and of loss of the initial 2-visit persistence of HPV infection, we distinguished the 1611 women with persistent infections for the first 2 scheduled visits according to their HPV status at the third visit if available (Table I-4). Odds ratios estimated by logistic regression for women with persistent infections and testing positive at the third visit were compared with those derived for women with the same persistent infections but showing no HPV DNA at the third (referent, no HPV at visits 1-3). Only events occurring within 48 months following the ascertainment of long-term HPV persistence (visits 1-3) were included. The OR of incident SIL of any grades for women remaining HPV positive following a persistent infection for oncogenic types was 6.69 times (22.02/3.29) that for those eliminating their infections at the third visit. This effect was also observed among women with infections for HPV types 16 and 18 at both initial visits, though to a lesser degree, for whom the OR was 1.15 times (12.27/10.71) that of nonpersistors. Conversely, there was no incremental risk associated with continued positivity after a persistent nononcogenic infection (3.25/3.55). Corresponding ORs of developing a persistent lesion for women maintaining HPV infections after being initially persistent were also high. No HSIL or persistent lesion events were observed for women eliminating their infections by the third visit regardless of HPV classification.

COMMENT

The traditional epidemiological study designs of single-opportunity assessment of exposure and outcome do not allow questions of viral persistence or regression of cervical lesions to be addressed. [72] To understand the role of HPV and the pattern of the dynamic changes in the natural history of cervical dysplasia, studies that collect data repeatedly on risk factors (HPV) and screen for cervical lesions on multiple occasions during follow-up must be conducted. In our study, we have provided evidence that persistent HPV infection, particularly with oncogenic types, is associated with a much greater risk of incident cervical cancer precursor lesions than when HPV positivity is defined on the basis of a single-assessment measurement at enrollment.

Our study has a number of strengths and weaknesses. Among the former we include more elaborate algorithms to define type-specific viral persistence (and loss thereof) on the basis of multiple initial visits in the cohort and the assessment of long-term incidence of initial and recurrent SIL as a function of HPV persistence. However, the definition of a persistent infection was based on detecting viral DNA of the same taxonomic type using a consensus PCR protocol.

Fluctuations in viral load below the detection threshold of PCR may have caused some cases of persistent infection to be misclassified as transient due to false-negative test results. In addition, it is impossible to ascertain via HPV DNA detection alone if test positivity is equated with true (active, albeit latent) viral infection. On the other hand, it is reassuring to note that PCR typing alone may be sufficient for defining persistent infections. Longitudinal testing for molecular variants of HPVs 16 and 18, while providing an enhanced level of taxonomic detail for ascertaining true persistence, [*73*] indicated that persistently detected HPVs 16 or 18 were of the same molecular variant in each case. [*252*] This observation suggests that persistent detection of the same viral type may truly represent a persistent infection.

Misclassification of lesion outcome history is a noteworthy weakness since our results were based on cytological ascertainment, however carefully conducted in a reference laboratory following a strict quality control protocol. We opted for an intensive, expert cytological follow-up every 6 months of all SILs found in the study to avoid having to perform unnecessary biopsies, which would have interfered with the natural history of early lesions. It is conceivable, however, that the magnitude of the associations would have been much greater if we had used histological ascertainment of all lesions detected in the study, an observation that we will make at a later phase of the investigation after we are able to define HPV persistence using algorithms that encompass at least 2 or more years' worth of follow-up visits with complete HPV testing and after more lesions are documented during long-term follow-up. Differential misclassification is unlikely because all HPV and Papanicolaou tests were performed blindly with respect to each other and by different laboratories. Therefore, being nondifferential, it is unlikely that the putative lesion misclassification bias would have elevated the observed associations. [*72*]

The HPV measurements in our study were collected at several return visits over a period of 1 year, allowing us to be more stringent with our definition of exposure to a persistent infection. Previous studies have relied on only 2 points of measurement, sometimes over a period of several years until a diagnosis of SIL. [*138;255*] The restriction to prevalence measures in case-control studies produces similarly elevated risk associations for concurrent HPV infection and lesion development.[1] When we emulated this restricted approach by ascertaining persistence using HPV test results taken at enrollment and at the time of diagnosis for an incident SIL during the first year of follow-up, our ORs increased to 94.9 (95% CI, 27.6-325.7) for "persistent" oncogenic infections and to 56.3 (95% CI, 16.5-192.6) for "persistence" of nononcogenic types. We even observed an increase in OR for transient infections to 24.1 (95% CI, 10.1-57.7) when relying on such an algorithm for HPV exposure.

There are some underlying differences between our study and previous cohort studies with respect to defining HPV persistence. The first approach is that described above where HPV status was evaluated at 2 points in time, the first at enrollment and the second being a prevalence measure collected at the same moment as the outcome is diagnosed. [*138;255*] A second approach is that HPV was measured at repeated intervals before the onset of disease, [*53;106;131;171;174;262;268*] although Ho et al [*106*] used a time-dependent algorithm for exposure assessment similar to that of the above studies. A variation on this latter method is to take into account the transient nature of precursor lesions in the cause of cervical neoplasia allowing for a woman to contribute more than once to the analysis. [*107*] Potentially more efficient when multiple events occur, this approach must take into account the sequential interdependencies between repeated events within subjects as they are followed up over time. However, a repeated analysis approach does not lend itself to an evaluation of severe outcomes such as HSIL because of the need for intervention.

Among those studies that evaluated oncogenic HPV infection status through consecutive scheduled visits, [*106;131;171;174;262;268*] few investigated the RR association for incidence of cervical neoplasia following a repeated HPV infection. [*131;171;174;262*] Using a nonamplified hybridization method for typing HPVs, Koutsky et al [*131*] observed an RR of 26 (95% CI, 6.5-112) for incident HSIL among women with multiple positive visits for HPV. Relative risk associations for single-point infections with oncogenic HPV types 16 or 18 were lower (RR, 11; 95% CI, 4.6-26). We observed a similar dose-response relationship in ORs with level of oncogenicity for persistent infections, although our RR associations, extended over a longer period of time, were lower. This observation has been part of a trend in decreasing RRs with increasing interval period between HPV exposure and SIL incidence in our study (data not shown). In 2 studies on populations of adolescent women, Moscicki et al [*171*] and Woodman et al [*262*] also observed a decreasing cumulative risk with time since first exposure to HPV. It has been suggested that latent SIL events occurring several years following an HPV infection found at baseline may correlate better with more recent infections yet to be detected. [*255;268*] This remains to be confirmed by continued follow-up and HPV typing at repeated visits.

Persistence of HPV infection was monitored by detection of individual HPV types by PCR, which provides a much finer level of detail than that afforded by a commercially available HPV testing method such as the Hybrid Capture assay (Digene Corporation, Gaithersburg, Md), the only HPV assay approved by the US Food and Drug Administration. Studies that rely on the Hybrid Capture test are limited to testing for multiple oncogenic HPVs collectively without distinguishing among types. Among those with repeated positive test results, we found that persistent HPV infections for both nononcogenic and oncogenic types were associated with substantially elevated RRs for

SIL incidence and persistence. Such risk associations would have been missed if only the 13 oncogenic types had been tested for, in combined form. Furthermore, considering subjects with nononcogenic infections as negative would have diminished the strength of the associations between HPV infections and lesions. Interestingly, for women with repeated positivity for oncogenic types (one of the highest risk categories in our study), almost all harbored persistent infections with the same types. No instances of HSIL were observed among women with nonpersistent (implying different type) yet with repeatedly positive HPV test results.

Some studies have demonstrated elevated RRs associated with HPV types 16 and 18 detected 2 years [*131;142*] or more [*262;268*] prior to the development of a high-grade lesion. In this study we were able to give particular attention to persistence of HPV types 16 and 18 and its relationship to incidence of HSIL. We found that the RR associations intensified with repeated detection of HPV 16 and 18 both in Cox regression analyses and by actuarial analysis for all SIL events. Ellerbrock et al [*53*] observed similar increases in RR for persistent HPV type 16 or 18 infections (RR, 11.6; 95% CI, 2.7-50.7) after adjustment for HIV seropositivity.

We also looked at the risk of developing a persistent lesion after an HPV infection. We noted particularly high RRs for persistent infections by oncogenic types. Of the 29 women with persistent lesions, 13 (44.8%) involved a diagnosis of HSIL. Due to the small number of subjects with persistent lesions, we could not speculate on the relationship with HPV type 16 or 18 infections. Other studies have attempted to look for predictors of lesion persistence or regression [*107;123*] basing their comparisons on a control group of women with previously detected lesions. Ho et al [*107*] observed ORs above 1 for lesion persistence among women who were positive for HPV at 2 prior consecutive visits, although they were not able to differentiate between oncogenic and nononcogenic HPV types.

The increase in ORs for incidence of SIL in women harboring long-term oncogenic HPV infections adds to the evidence for HPV persistence as a key determinant of lesion development. Although few SIL events were available for analysis after restriction for HPV positivity at the first 3 scheduled visits, proportionally fewer lesions were found in women who eventually eliminated their infection within 8 months. No persistent SIL cases or incident HSIL events were detected among those who cleared their infection at the third visit in the study. In a study of adolescent women with HPV infection at enrollment, Moscicki et al [*174*] observed an OR of 14.1 (95% CI, 2.3-84.5) for the incidence of HSIL in women positive for oncogenic HPV at 3 of 4 preceding visits compared with those who lost their infection after enrollment into the study.

In our study, we do not know the proportion of HPV-positive women identified at enrollment who were already harboring persistent infections before entering the investigation. Loss of persistence at the third visit could merely indicate that the 2 previous positive visits were of a transient nature, whereas those with the initial 3 visits being positive for oncogenic HPVs might have represented true long-term persistence that began before they entered the study. As we extend typing to subsequent follow-up visits in the study, we will be able to more accurately evaluate the effects of long-term persistence and of loss of positivity to HPV on the subsequent development of SIL.

In conclusion, our study adds to the body of evidence strongly implicating persistent HPV infections, particularly with oncogenic types and more prominently with HPV types 16 and 18, in the cause of SIL. Using a longitudinal, repeated measurement cohort investigation we were able to assess more refined algorithms of cumulative HPV exposure with respect to their prognostic value in determining lesion outcome history. However, further analyses for repeated measures remain to be done to evaluate the transient nature of the disease and to investigate the long-term natural history of HPV infection and cervical intraepithelial neoplasia in our cohort. Our results, however, would support the proposal for the application of repeated type-specific HPV DNA testing in screening and for the potential use of vaccines for HPV types 16 and 18 to prevent the development of clinically relevant cervical lesions.

Table I-1. Distribution of selected characteristics among 1789 subjects with HPV test results at enrollment.

Demographic Variable	Categories	No. of subjects*	Percent
Age (years)	18-24	343	19.2
	25-34	700	39.1
	35-44	538	30.1
	≥ 45	208	11.6
Ethnicity	White	1158	64.8
	Non-white	630	35.2
Education	Less than elementary	412	23.1
	Elementary	1046	58.5
	High School	280	15.7
	College/University	49	2.7
Marital status	Single	167	9.3
	Married or living together	1468	82.1
	Divorced or widow	154	8.6
Smoking	Never	884	49.4
	Current	601	33.6
	Former	304	17.0
Age at first intercourse (years)	≤ 15	486	27.2
	16-17	456	25.5
	18-19	371	20.7
	≥ 20	476	26.6
Lifetime number of sexual partners	0-1	799	44.7
	2-3	624	34.9
	4+	365	20.4
Duration of oral contraceptive use	Never	281	15.7
	≤ 5 years	980	54.8
	> 5 years	528	29.5
Number of pregnancies	0-1	275	15.5
	2-3	769	43.3
	4-6	546	30.8
	≥ 7	184	10.4
Menopausal status	Pre- or perimenopausal	1724	96.4
	Post-menopausal	65	3.6
Condom use	Never	716	40.0
	Rarely	629	35.2
	Frequently or always	444	24.8
Douching	Never	1191	66.7
	Occasionally	413	23.1
	Frequently	182	10.2
History of sexually-transmitted diseases	Not reported	1372	77.0
	HPV-associated	87	4.9
	Other diseases	322	18.1

* Subjects with missing information excluded.

Table I-2. Frequencies of new cases developing cervical lesions after extended follow-up according to HPV infection assessed at the first two visits*

Definition of persistent HPV infection based on first 2 visits	No. of subjects	Women-months‡	Any SIL		High Grade SIL		Persistent SIL†	
			No. of events (%)	Rate/1000 women-months (95%CI)	No. of events (%)	Rate/1000 woman- months (95%CI)	No. of events (%)	Rate/1000 woman- months (95%CI)
Negative	1268	61948.7	44(3.5)	0.73 (0.5-0.9)	9 (0.7)	0.15 (0.1-0.2)	12 (0.9)	0.19 (0.1-0.3)
Positive either visit for any type§	198	9638.2	29 (14.6)	3.40 (2.2-4.6)	6 (3.0)	0.64 (0.1-1.1)	6 (3.0)	0.64 (0.1-1.2)
HPV types 16 or 18 once	36	1434.4	3 (8.3)	2.25 (0.0-4.8)	0 (0)	0.0 -	0 (0)	0.0 -
Positive on 2 visits for same non-oncogenic types	47	2212.5	7 (14.9)	3.52 (0.9-6.1)	0 (0)	0.0 -	3 (6.4)	1.39 (0.0-3.0)
Positive on 2 visits for same oncogenic types except 16 or 18	38	1867.6	12 (31.6)	8.27 (3.6-13.0)	2 (5.3)	1.08 (0.0-2.6)	6 (15.8)	3.51 (0.7-6.3)
HPV types 16 or 18 (same) both visits	24	1037.8	7 (29.2)	8.68 (2.3-15.1)	2 (8.3)	2.06 (0.0-4.9)	1 (4.2)	1.00 (0.0-3.0)

* Excluding subjects with squamous intraepithelial lesions (SIL) detected at enrollment or missing HPV test results.

† Two or more visits with SIL during follow-up allowing for one negative interval visit.

‡ Total study follow-up time for all enrolled women.

§ Includes women with two positive visits but different HPV types in each, except types 16 or 18.

Table I-3. RRs and corresponding 95% CIs of cervical lesions over five years of follow-up according to HPV positivity during the first two visits*

Definition of HPV infection status based on first 2 visits	No. of subjects	Any SIL		High Grade SIL		Persistent SIL†	
		RR‡	95% CI	RR‡	95% CI	RR‡	95% CI
Emphasis on same-type persistence							
Negative	1268	1.0	(referent)	1.0	(referent)	1.0	(referent)
Positive either visit for any type§	198	4.23	(2.6-6.8)	4.08	(1.4-11.7)	3.53	(1.3-9.5)
HPV types 16 or 18 once	36	2.69	(0.8-8.7)	3.85	(0.5-31.1)	0.00	-
Positive on 2 visits for same non-oncogenic types	47	4.49	(2.0-10.0)	0.00	-	7.93	(2.2-28.3)
Positive on 2 visits for same oncogenic types except 16 or 18	38	9.92	(5.2-18.9)	9.68	(2.6-36.2)	18.89	(7.0-51.1)
HPV types 16 or 18 (same) both visits	24	11.15	(5.0-24.9)	12.27	(2.6-57.5)	7.40	(1.0-57.3)
Persistence based on oncogenic types							
Negative	1268	1.0	(referent)	1.0	(referent)	1.0	(referent)
Positive at only 1 visit	197	2.92	(1.7-5.0)	2.64	(0.8-8.7)	2.78	(1.0-8.0)
Positive both visits for any type	53	4.89	(2.3-10.4)	2.48	(0.3-19.7)	7.26	(2.0-25.9)
Positive both visits with an oncogenic type in one	25	10.57	(5.0-22.5)	5.24	(0.7-41.8)	5.47	(0.7-42.4)
Positive both visits with different oncogenic types	4	24.55	(5.8-103.5)	0.0	-	0.0	-
Positive both visits for the same oncogenic types	64	10.19	(5.9-17.6)	11.67	(4.1-33.3)	14.92	(5.8-38.3)

* Excluding subjects with squamous intraepithelial lesions (SIL) detected at enrollment. Missing categories not shown.

† Two or more visits with SIL during follow-up allowing for one negative interval visit.

‡ RR estimates and 95% CI by Cox proportional hazards regression adjusting for age and ethnicity.

§ Includes women with two positive visits but different HPV types in each, except types 16 or 18.

Table I-4. ORs and corresponding 95% CIs of incident cervical lesions within two years following a persistent infection for HPV according to infection status at the third visit*

Assessment of HPV infection		Any SIL			Persistent SIL†		
Status at visits 1 and 2	Status at visit 3	Lesion /No lesion	OR‡	95% CI	Lesion /No lesion	OR‡	95% CI
Negative all three visits		25/1046	1.0	(referent)	8/1008	1.0	(referent)
Positive for 2 visits with same non-oncogenic types	Negative	1/12	3.55	(0.4-28.5)	0/10	-	
	Positive	2/25	3.25	(0.7-14.5)	2/21	12.03	(2.4-60.8)
Positive for 2 visits with same oncogenic types§	Negative	1/12	3.29	(0.4-26.5)	0/12	-	
	Positive	7/13	22.02	(8.1-60.2)	4/12	41.16	(10.7-158.3)
Positive for 2 visits for HPV 16/18	Negative	1/4	10.71	(1.1-101.2)	0/4	-	
	Positive	3/10	12.27	(3.2-47.6)	1/8	17.21	(1.9-158.4)

* Excluding subjects with squamous intraepithelial lesions (SIL) detected at enrollment or first follow-up visit.

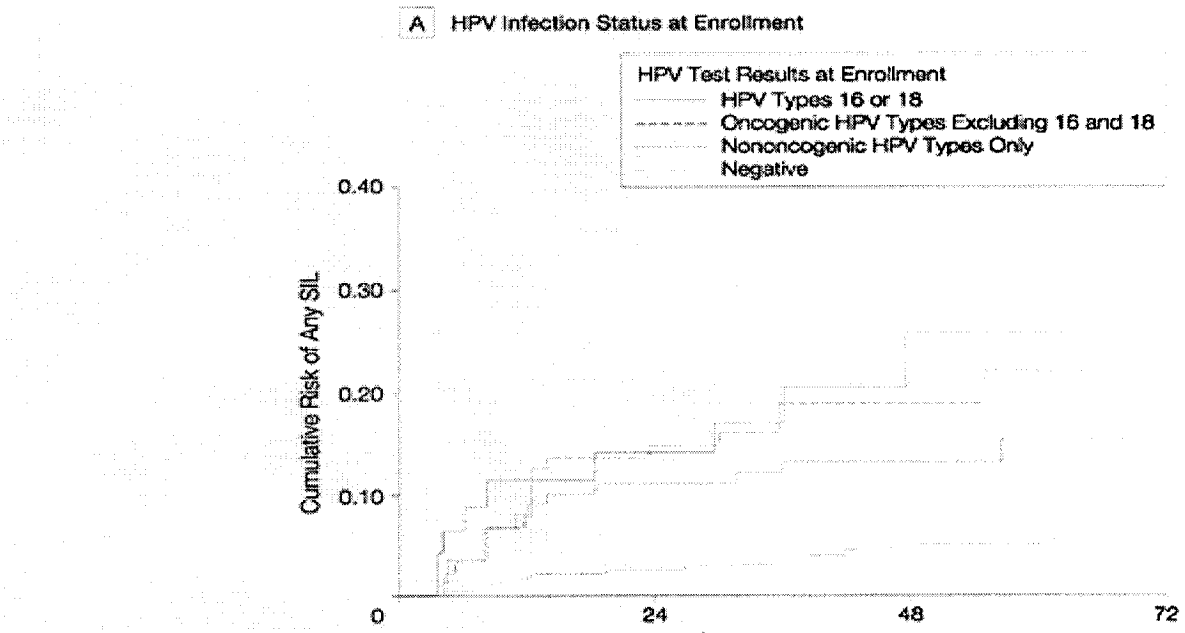
† Two or more visits with SIL during follow-up allowing for one negative interval visit.

‡ OR and 95% CI by logistic regression with analyses restricted to events occurring within 48 months of assessment of HPV persistence adjusting for age and ethnicity.

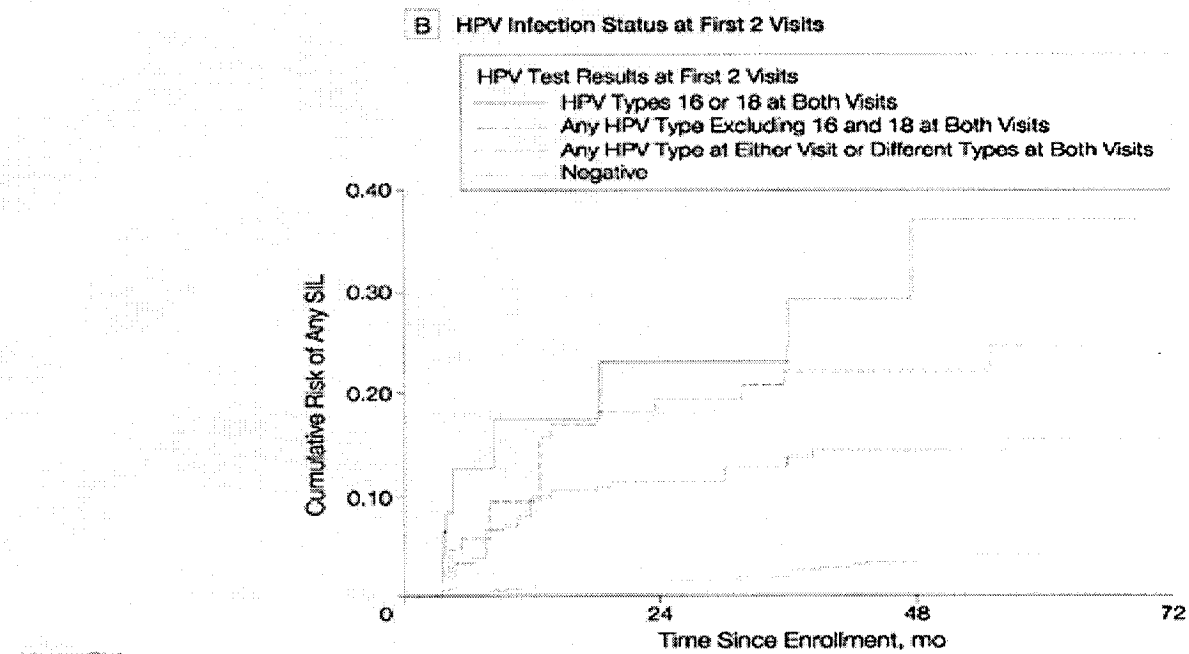
§ Excluding HPV types 16 or 18.

Figure I-1. Kaplan-Meier estimates of cumulative incidence of any grade of SILs

Legend Mutually exclusive categories for human papillomavirus (HPV) infection among women free of cervical lesions at enrollment are displayed. A, HPV positivity at enrollment by oncogenicity among 1746 women, and B, HPV persistence for the first 2 visits by oncogenicity among 1611 women. Women with lesions detected at enrollment or missing HPV test results are excluded from the analyses.



No. at Risk				
HPV Types 16 or 18	47	31	14	0
Oncogenic HPV Types Excluding 16 and 18	94	70	36	1
Nononcogenic HPV Types Only	109	87	58	4
Negative	1496	1210	839	73



No. at Risk				
HPV Types 16 or 18 at Both Visits	24	14	8	0
Any HPV Type Excluding 16 and 18 at Both Visits	87	63	37	0
Any HPV Type at Either Visit or Different Types at Both Visits	232	178	111	12
Negative	1268	1096	763	66

6.3. Preface to Manuscript II

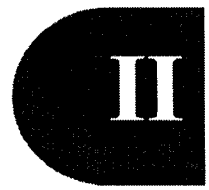
Most HPV infections are transient and only a small proportion of women infected develop cervical neoplasia and cancer. Persistence of HPV infections seems to be an essential intermediate event for cervical carcinogenesis. An association between high viral load and HPV persistence has also been observed [*25;63;105;107;142*]. When multiple measurements of the virus cannot be made, a single measure of viral burden may be sufficient to identify a subset of HPV infected women at higher risk of developing cervical cancer [*114;248*].

Unlike previous studies on this subject, I report in the next manuscript results using true viral burden as a precursor event in the genesis of cervical neoplasia among HPV positive women analyzed in a prospective fashion. Owing to the longitudinal, repeated measurement design and to the highly sensitive methods of HPV quantification of this cohort study I believe that the present manuscript provides critically important new data to the notion that only a subset of women infected with HPV will go on to develop cervical lesions. With the recent advances on HPV vaccine technology phase III trials are about to start in different populations and the discussion of using HPV testing as a screening tool for cervical cancer, this study provides baseline data on the definition of HPV viral burden and associated lesion risks that could be used for the planning of vaccine trials or screening programs. As per the underlying theme of this thesis I explored the time effects of viral load using models that accommodate the longitudinal nature of the data. These models allowed for the evaluation of the predictive effect of lesion risk by viral load at a previous visit.

6.4. Manuscript II: Viral load as a predictor of the risk of cervical intraepithelial neoplasia

(Running title: HPV load and CIN)

Nicolas F. Schlecht MSc[1], Andrea Trevisan MSc[2,3], Eliane Duarte-Franco MD MPH[1], Thomas E. Rohan MD PhD[4], Alex Ferenczy MD[5], Luisa L. Villa PhD[3], Eduardo L. Franco PhD[1]



Author Affiliations:

[1] Division of Cancer Epidemiology, McGill University, Montreal, Canada;

[2] Department of Microbiology, ICB, University of São Paulo;

[3] Ludwig Institute for Cancer Research, São Paulo, Brazil;

[4] Department of Epidemiology and Social Medicine, Albert Einstein College of Medicine, New York, USA;

[5] Department of Pathology, Sir Mortimer B. Davis Jewish Hospital, McGill University, Montreal, Canada.

ABSTRACT

HPV infections are believed to be a necessary cause of cervical cancer. Viral burden, as a surrogate indicator for persistence, may help predict risk of subsequent SIL. We used results of HPV test and cytology data repeated every 4-6 months in 2081 women participating in a longitudinal study of the natural history of HPV infection and cervical neoplasia in São Paulo, Brazil. 473 women were positive for HPV DNA by the MY09/11 PCR protocol during the first two visits. We retested all positive specimens by a quantitative, low-stringency PCR method to measure viral burden in cervical cells. Mean viral loads and 95% CIs were calculated using log-transformed data. RR and 95%CIs of incident SIL were calculated by proportional hazards models adjusting for age and HPV oncogenicity. The risk of incident lesions increased with viral load at enrollment. The mean number of viral copies per cell at enrollment was 2.6 for women with no incident lesions and increased (trend p value = 0.003) to 15.1 for women developing 3 or more SIL events over 6 years of follow-up. Compared to those with less than one copy per cell in specimens tested during the first two visits, RRs for incident SIL increased from 1.9 (95%CI: 0.8-4.2) for those with 1-10 copies per cell, to 4.5 (95%CI: 1.9-10.7) for >1000 copies per cell. The equivalent RR of HSIL for >1000 copies per cell was 2.6 (95%CI: 0.5-13.2). Viral burden seems to have an independent effect on SIL incidence. Measurement of viral load, as a surrogate for HPV persistence, may identify women at risk of developing cervical cancer precursors.

** This manuscript was accepted for publication in the Int J Cancer, October 2002*

INTRODUCTION

Establishment of productive infections by oncogenic types of HPV is a key early event in the natural history of cervical cancer. Given the positive relationship between viral load and the likelihood of persistent HPV infection [*211*] and the strong relationship between the latter and risk of cervical neoplasia [*42*], a single measurement of viral load in cervical specimens may be a suitable bio-marker of either a transient infection or the likelihood that an instance of HPV DNA positivity may represent a persistent infection or one that may become persistent over time and lead to the development of cervical SIL. In the latter case, little is known, however, about the relationship between level of viral burden and lesion incidence or lesion grade severity in the natural history of CIN. HPV DNA viral load has been identified as a biomarker for cervical lesions in women [*116;268*], primarily with minor grade cytological atypia [*228*]. However, only a few studies have evaluated the predictive effect of viral load on the incidence of cervical cancer precursors [*35;107;108;116;268*].

A few methods of viral load estimation have been evaluated in epidemiological studies. Although several cross-sectional studies have examined the association between the HPV viral load measured by HC and the presence of SIL [*40;44;89;226;237*], HC cannot accurately quantify HPV DNA in cervical specimens because of the variable exfoliated cell content of such samples. The correlation between HPV viral load and grade of lesions is dependent on HPV type [*240*], HC identifies multiple types altogether making quantification difficult to interpret. Other studies have used semi-quantitative PCR protocols to measure the HPV viral load [*44;89;107;108;127;226;261*]. However, 2 factors that influence estimates of viral load, namely PCR inhibition and the cell content of samples, were not taken into account in these investigations. Recently, Joseffson et al. [*114*] described an accurate measure of viral load by employing a real-time PCR protocol.

In 1993, we began a longitudinal investigation (Ludwig-McGill Cohort Study) of the natural history of HPV infection and cervical neoplasia in the city of São Paulo, Brazil, using a PCR based method for viral load quantification. The focus of the analyses reported here was to assess the long-term risk of incident CIN as a function of viral load, as well as to evaluate the potential use of viral load measurement as an aid to the clinical management of HPV positive women. The study participants were followed for up to 8 years, during which time they were screened concurrently for cervical lesions and HPV infection at scheduled return visits.

MATERIALS AND METHODS

Subject recruitment

Two nurses, specifically trained for the study, systematically selected women attending a comprehensive maternal and child health program catering to low income families in of São Paulo, Brazil. Subjects entered the study only after giving signed informed consent. The study protocol was approved by institutional ethical and research review boards of the participating institutions in Canada and in Brazil. A detailed description of the design and methods of the study has been published previously [*62*].

Women were eligible to participate if they: (i) were aged between 18 and 60 years; (ii) were permanent residents of São Paulo (city); (iii) were not currently pregnant and had no intention of becoming pregnant during the next 12 months; (iv) had an intact uterus and no current referral for hysterectomy; (v) reported no use of vaginal medication in the previous 2 days; and (vi) had not had treatment for cervical disease in the previous 6 months. A total of 2528 participants were recruited into the study between 1993 and 1997.

All participants were seen every 4 months in the first year (0, 4, 8 and 12 months), and twice yearly thereafter. Follow-up has been ongoing for over eight years with a response rate at each visit of approximately 70%. Cervical specimens were taken for Pap cytology and HPV testing at every visit.

Cervical cell specimens

An Accelon biosampler (Medscand, Inc., Holywood, FL) was used to collect ecto- and endocervical samples. After the cells were smeared on a glass slide and fixed for cytology, the sampler containing the residual exfoliated cells was immersed in a tube containing Tris-EDTA buffer (pH 7.4). The tube was agitated to release the cells from the sampler, which was then discarded. Samples were then sent to the Ludwig Institute for Cancer Research in São Paulo for storage and testing. The Pap smears were shipped to Montreal, where they were reread by one of the authors (A.F.). The cytopathology reports were based on the Bethesda system for cytological diagnoses [*97*] and were blinded to HPV results for the same samples.

HPV DNA detection

Cervical specimens were tested for the presence of HPV DNA by a previously described PCR protocol amplifying a highly conserved 450 bp segment in the *L1* viral gene (flanked by primers MY09/11) [*10;105*]. Typing of the amplified products was performed by hybridization with individual oligonucleotide probes specific for 27 HPV genital types whose nucleotide sequences for probes within the MY09/11 fragment have been published in the literature [6/11, 16, 18, 26, 31, 33, 35, 39, 40, 42, 45, 51, 52, 53, 54, 55, 56, 57, 58, 59, 66, 68, 73, 82, 83, and 84] [*105*].

Amplified products that hybridized with the generic probe but with none of the type-specific probes were further tested by RFLP analysis of the *L1* fragment [*12*] to distinguish among unknown HPVs. Use of the RFLP analysis extended the range of identifiable HPV types to include additional HPVs, i.e. 32, 34, 44, 62, 64, 67, 69, 70, 71, 72 and 81, as well as CP6108 plus other unknown types. To verify the specificity of the hybridizations, we included more than 30 type-specific positive controls in all membranes. In order to check the integrity of the host DNA material extracted from the specimens, assays also included an additional set of primers (GH20 and PC04) to amplify a 268 bp region of the *ß-globin* gene [*10*]. For the investigation reported here, HPV types were separated into 2 groups by presumed oncogenicity (i.e., oncogenic and non-oncogenic). Oncogenic types included HPVs 16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, 59, and 68; all others were considered non-oncogenic. All HPV assays were done on coded specimens blinded to any identification linking specimens from the same woman.

Measuring viral burden

We retested all cervical specimens found to be positive with the main PCR protocol (MY09/11) by a quantitative, low-stringency PCR to measure viral burden in exfoliated cervical cells [*29*]. The method uses general primers (GP5, GP6) from a well known PCR protocol that detects a broad spectrum of HPVs [*247*]. The quantitative PCR protocol employs low stringency conditions to co-amplify the specific HPV DNA fragment along with DNA sequences from the human genome present in the starting PCR mixture. Standards consisting of mixtures containing varying amounts of reference HPV 16 plasmid (corresponding to 0, 4, 20, 100, 500, and 2500 viral copies per cell) were included in duplicate in every assay added to a constant background of normal human DNA. In addition, control samples consisting of DNA from 2 cervical carcinoma cell lines with known quantities of HPV copies (HeLa, 20-40 copies/cell of HPV-18; Caski, 400-600 copies/cell of HPV-16) were included in duplicate in every assay. The silver-stained gel bands corresponding to the HPV and to the constant human genome fragments were quantified by densitometry [*29*]. The logarithm of the ratio between these 2 bands is directly proportional to the logarithm of the amount of HPV DNA in the individual samples. Proper quantification is obtained by linear interpolation in a standard curve constructed with the results from the control mixtures. The dose-response relationship is linear at concentrations as high as 5000 copies per cell and is independent of the amount of DNA present in the reaction mixture, of the number of PCR amplification cycles, of staining intensity, and of the choice of human genome bands used as reference [*29*]. Samples with viral loads near the limits of detection (5000 copies per cell) were diluted and re-run to ensure accuracy. Assessment of viral burden via this protocol represents a combined measure for any HPV types present in the specimen that are identifiable by the GP5/6 primer pair. Quality control data using HeLa and Caski controls from all test runs in this study

indicated good reproducibility and accuracy. Means were 52.5 (95% CI 40-66) and 801 (95% CI 633-969) copies per cell for HeLa and Caski, respectively.

Statistical Analysis

To evaluate the long-term risk of SIL/CIN in association with viral load we focused on 465 women positive for HPV at enrollment and after four months for whom cervical specimens were tested for viral load. Mean viral loads and 95% CIs across these visits were calculated using logarithm transformed data. We analyzed the risk of an incident SIL diagnosed by cytology occurring in women free of SIL at enrollment according to viral load at enrollment and during the first 2 follow-up visits. Histology was not performed for diagnostic ascertainment.

RRs of incident cervical lesions over time were modeled by Cox proportional hazards regression over the entire 8 years of follow-up for women free of SIL at enrollment. Time to event was measured from enrollment to the first instance of a lesion event or to the last recorded return visit date for censored subjects. Incidence of SIL was estimated by lesion severity (first confirmed SIL event of any grade and first occurrence of HSIL), and by further taking into account the different sub-categories of lesion grades. This corresponded to stratifying LSIL into those that showed predominantly cytopathic effects of productive HPV infections (LSIL/HPV) and those that showed squamous abnormalities (LSIL/SQ). Similarly, HSIL findings were also separated into those lesions consistent with moderate dysplastic changes (equivalent to CIN2) and those with more severe dysplastic features (equivalent to CIN3). All incident cases of lesions were compared with subjects with no detected lesions or with atypical squamous cells of undetermined significance (ASCUS) during the entire period of follow-up.

To assess the potential clinical value of quantitative viral measurements we examined the strength of the association between viral load and lesion risk at the same and in consecutive visits. Allowing for repeated assessments of HPV exposure, we evaluated the association between viral load and concurrent and subsequent cervical lesion status for 1441 HPV positive cervical specimens collected at six consecutive scheduled visits during the first 2 years of follow-up. We calculated average viral loads for each HPV positive subjects using results from the first 2 follow-up visits to reduce the level of misclassification in assessing viral burden. A total of 10863 person-visits and cervical specimens were recorded during this period. We used the GEE regression approach to estimate marginal ORs as estimates of the RR of any SIL or HSIL at the same visit and at the next visit according to viral load at any given point in time. This technique allowed us to account for the correlation between repeated events within subjects by adjusting for the intra-subject correlation using an exchangeable correlation matrix. This approach produced the most conservative estimates of effect for the OR estimates. Whenever an exchangeable

correlation matrix could not be estimated after stratification, repeated events were assumed to be independent. Prevalent cases of lesions detected at the previous visit were excluded from the analysis of lesion incidence at the next visit.

We adjusted *a priori* for age using 4 groups contrasted as dummy variables (18-24, 25-34, 35-44, and 45-60 years), and for HPV infection status at the index visit (only non-oncogenic types and any oncogenic types). We investigated the potential effect of confounding by marital status, education, smoking history, feminine hygiene practices, parity, oral contraceptive use, condom use, menopausal status, age at first intercourse, number of sexual partners, history of sexually transmitted diseases, and occurrence of other genital infections that could influence the development of SIL; no other variables were identified as confounders based on a change in estimate criterion [*148*]. We performed tests for trend by including categorized risk factors as ordinal variables in the multivariate models. Statistical analyses were performed using the statistical programs SPSS® version 11.0 (Chicago, IL) and STATA® version 6 (StataCorp, College Station, TX).

RESULTS

Of the 2081 women who provided cervical samples for cytological and HPV analysis, 473 (23%) women tested positive for HPV at least once at either of the first 2 visits. Of these, 40 subjects presented with prevalent lesions detected at enrollment, 1 Pap smear was inadequate for evaluation. Viral load estimation could not be made for a further 15 women positive for HPV at either of the first two follow-up visits, leaving 417 women available for inclusion in the analyses. The mean follow-up time until censoring or incidence of HSIL for HPV positive subjects with less than 1 copy per cell, low viral load (1-10 copy per cell), moderate (11-100 copies per cell), high (101-1000 copies per cell) and very high viral load (>1000 copies per cell) were 63.1, 61.1, 51.2, 58.4 and 58.6 months, respectively.

To investigate the role of viral load in the natural history of cervical lesion development, we analyzed viral load measured at either of the first 2 visits according to the number of subsequent positive Pap tests during follow-up (Table II-1). Compared to women who remained cytologically negative during the follow-up period, viral load was higher for women with at least 2 positive Pap tests at 6.2 copies per cell and increased with the number of positive Pap tests to 18.7 copies per cell in women with 3 or more subsequent SIL by cytology. The viral load in the latter women was significantly higher ($p = 0.001$) than in those with no positive Pap tests (mean = 2.2 copies per cell). A statistically significant weighted dose response relationship was observed between log viral load and increasing number of incident SILs (p for trend < 0.0005).

We examined the relative increase in risk of developing any SIL and HSIL over time with increasing viral load among women who were HPV positive at either of the first 2 visits (Table II-2). There was a dose-response trend for an increase in RRs with higher average viral copy frequency from less than 1 to over 1000 copies per cell. However, this trend was statistically significant only for incidence of any SIL ($p < 0.001$). A less strong relationship was seen when we restricted the viral load exposure assessment to the first visit only (data not shown). There was no dose-response trend in viral load - lesion risk relation for HSIL events among those with no prevalent HSIL at enrollment but this analysis was based on fewer cases ($n = 15$). Nevertheless, risk of HSIL was elevated for those in the highest viral load category.

We repeated the analysis shown in Table II-2 separately for women less than 35 and those 35 and over. A similar trend in increasing RR of any SIL was observed with viral load for younger women with point estimates increasing from 3.1 (95%CI 1.2-8.0) for women with specimens containing 1-10 viral copies per cell, 5.8 (95%CI 2.5-13.6) for 11-100 copies per cell, 4.3 (95%Ci 1.8-10.3) for 101-1000 copies per cell, and 9.9 (95%CI 3.8-25.9) for >1000 copies per cell. RRs of SIL for older women were lower but not significant due to the small number of subjects with high viral load. RRs of incident HSIL were also higher for younger women: 1.4 (95%CI 0.1-13.2) for 1-10 copies per cell, 4.4 (95%CI 0.9-21.8) for 11-100 copies per cell, 2.4 (95%CI 0.4-14.8) for 101-1000 copies per cell, and 5.9 (95%CI 0.9-36.9) for >1000 copies per cell. Due to small sample sizes, point estimates were not estimable for older women.

When we expanded the lesion categories to distinguish between different grades of LSIL (with and without koilocytosis) and HSIL (Table II-3) we observed consistently elevated RRs for women harboring higher mean viral loads at the first two visits after adjustment for age and HPV status. The RRs for those with >100 copies per cell decreased somewhat with increasing lesion severity from LSIL/SQ to HSIL/CIN2 and CIN3 although the RRs for the latter were based on fewer events and thus lacked precision.

We also evaluated the potential use of viral load measurement as an aid for clinical management of HPV positive women, by calculating RRs of lesions according to high, intermediate, or low viral loads assessed at the same visit, as a cross-sectional association, and at the 2 previous visits, as prospective associations (Table II-4). In all there were 1525 HPV positive cervical samples from the 2081 women tested during the first 2 years of follow-up. Of these, 81 HPV positive specimens were not quantifiable for viral load. 1444 valid samples were quantifiable for viral load and were included in the analyses. Increasing viral load correlated strongly with overall prevalence of LSIL and HSIL cross-sectionally. The association was weaker for HSIL than for LSIL with viral load levels above 100 copies per cell. The same pattern of effects was seen for incident LSIL events

occurring at the next visit (4 to 6 months after viral load assessment) as well as for events occurring 2 visits after viral load ascertainment (corresponding to an interval of 8 to 12 months). However, the magnitudes of effect were decreased somewhat compared to those for the cross-sectional associations. RR estimates for HSIL alone were less consistent but there was evidence of a stronger association, which was maintained for higher viral load levels, ascertained by 2 visits prior to the cytological assessment (8 to 12 months apart).

To determine if cytological status at the visit in which viral load was assessed affected the associations described above, we conducted analyses of any SIL events by restricting observations at the viral load assessment visit to women with either normal or benign cellular changes, as one stratum, and to those with ASCUS, as a second stratum (Table II-5). This stratified analysis indicated that regardless of the time of lesion ascertainment, either 4-6 or 8-12 months post viral load assessment; viral load was only predictive of SIL occurrence among women with smears within normal limits. There was no prognostic value in knowing viral load levels among women with ASCUS smears.

DISCUSSION

Relying merely on simple HPV type detection as a bio-marker of increased risk of subsequent cervical neoplasia suffers from several practical limitations. Fluctuations in viral load below the detection threshold of screening tests can lead to misclassification of some infected women as false negative. This can be a problem with non-PCR based techniques [*226*]. In addition, it is impossible to ascertain via HPV DNA detection alone if test positivity is equated with an active, persistent viral infection. For practical purposes, one can ascertain persistence via repeated testing 6 to 12 months apart, which improves substantially the ability to predict cervical lesion incidence [*224*]. However, assessment of persistent infections cannot be a solution for screening and management algorithms because it would require a high patient compliance with scheduled return visits. Poor compliance with repeat tests is one of the greatest pitfalls of current cervical cancer screening programs [*166*]. The amount of viral load could conceivably serve as a proxy for persistent infections. We therefore attempted to evaluate the independent effect of viral burden on the development of SIL in a large cohort of women using a PCR-based technique for viral load quantification. While our study is one of the largest currently evaluating the influence of viral burden on the incidence of CIN, we recognize that the design has some limitations.

Misclassification of lesion outcome history is a potential limitation in this study since our results were based on cytological ascertainment only. We opted for intensive, expert cytologic follow-up every four to six months of all ASCUS and LSILs found in the study to avoid having to perform unnecessary biopsies, which would have interfered with the natural history of cervical lesions.

Nevertheless, it is conceivable that the magnitude of the associations would have been greater if we had used histological ascertainment of all lesions detected in the study. Colposcopy-oriented biopsies in expert hands can miss 20% to 30% of HSILs [*163*]. Differential misclassification is unlikely because all HPV and Pap tests were performed blindly with respect to each other and the 2 tests were performed by different laboratories. Women were referred immediately for colposcopy whenever an HSIL result was detected by either the local or review cytology readings or by cervicography. The results of colposcopic follow-up and biopsies of such cases are currently being reviewed and will form the basis for a future analysis when the entire cohort completes a minimum of 5 years of follow-up.

Testing under a wide range of conditions has shown that the PCR-based technique used to measure viral load in this study is reproducible and adaptable to large scale testing in epidemiologic studies [*29*]. The method has also been demonstrated to be reliable over repeated samplings since women harbor similar levels of viral load from one visit to the next [*211*]. Abnormal cells found in CIN 2 or 3 lesions express fewer intercellular adhesion molecules than normal cells [*127*] and could possibly be sampled more readily than normal cells. Our method allows for this by normalizing the numbers of copies of viral DNA against the quantity of host DNA permitting the calculation of true viral load in terms of the number of copies per cell, thus eliminating the fluctuation due to variation in cell content among specimens from different subjects and from the same subject over time [*29*]. Furthermore, the results from the inclusion of cervical carcinoma cell lines as secondary controls (HeLa and Caski cells) in every testing batch indicate that the method is sufficiently accurate and precise in allowing a quantitative assessment of the number of HPV copies per host cell [*29*]. The dose-response relationship is linear at concentrations as high as 5000 copies per cell and is independent of the amount of DNA present in the reaction mixture, of the number of PCR amplification cycles, of staining intensity, and of the choice of human genome bands used as reference [*29*]. Unlike previous methods based on semi-quantitative assessment of PCR signals using external standards [*241*], this technique appropriately satisfies the criteria for quantitative measurement of viral load in cervical specimens. It is reproducible, has an adequate linear range in dose-response, and provides results standardized for cell content.

We further attempted to reduce the level of misclassification in assessing viral load by including measurements from the 2 initial follow-up visits. Expectedly, this seemed to improve the magnitude of the associations with lesion risk but in general, a single testing opportunity (e.g., the enrollment specimen or the ones at any other follow-up visit) carries substantial risk prediction value for both short- and long-term incident lesions.

Our results indicate that among HPV-positive women viral burden seems to have an effect on LSIL and HSIL incidence that is independent of that contributed by age and HPV types grouped by oncogenic potential. Women with multiple positive Pap tests during follow-up were also more likely to have harbored higher viral loads at the beginning of the study than those who remained lesion free throughout follow-up. A dose-response relationship with increasing viral load was also observed for the risk of an incident LSIL event over time, indicating that women with greater viral burdens are more likely to develop LSIL. While less significant, the presence of high viral load also appears to serve as a predictor of incidence of HSIL among women with no such lesions. Measurement of viral burden may help in the clinical setting to identify HPV positive women at greater risk of developing moderate or severe dysplasia but its greatest value is in predicting the initiation of the dysplastic process.

Few truly quantitative PCR assays have been developed for HPV viral load measurement. Other groups have utilized real-time PCR assays for HPV DNA load determinations [*114;116;127;240;248;268*]. In one study, cervical CIS occurred in women with consistently higher HPV-16 DNA load [*268*], and higher HPV16 viral load at baseline was found to predict the risk of CIS over a period of 7 years [*116*]. Van Duin et al. [*248*] found higher viral load for HPV16 to be associated with incidence and progression of CIN 2 and 3 in women with both normal and abnormal cytology at baseline. In the latter 2 studies, sample accuracy was assessed by measuring the relative expression of human *ß-actin* or *ß-globin* gene in samples to normalize HPV results for cellular content by expressing quantities in copies per cell [*116;127;248;268*]. Such correction is necessary to avoid introducing a bias in viral load measurement since more cellular material can be obtained from sampling HSILs, potentially generating artificially higher HPV viral load [*127*]. The measurements, however, were restricted to HPV types 16 or 18 without consideration of the overall burden of infection nor do they consider other HPV types of differing oncogenic potential. Viral load quantification in our study was undertaken irrespective of HPV type, and represents an average, overall viral burden. While we could not determine the relative contributions to the overall viral load of different HPV types in specimens harboring infections with multiple types, we attempted to control for such effects by adjusting for HPV types grouped by oncogenic potential in all analyses. Nonetheless, we realize that this adjustment is incomplete and estimates may have been attenuated. On the other hand, use of our technique removes the need for testing for viral burden with type-specific assays, which would make testing in screening conditions more time-consuming and expensive.

Cross-sectional studies using HC-II have shown mixed associations between viral load and lesion severity, particularly for HSIL samples [*228;237*]. Others [*35*] have found it unreliable in predicting HSIL up to 3 years later. Using a semi-quantitative PCR technique, Schneider et al. [*226*] found

viral loads in HPV-positive smears of prevalent CIN 2/3 to be significantly higher than in incident CIN 2/3 lesions (p = 0.0005).

Our results show that the risk prediction ability of viral load measurements is less for incident HSIL than it is for LSIL, at least within the duration of follow-up accrued in our study to date. This may be indicative of a correlation between the degree of proliferation associated with the underlying HPV infection and lesion grade, with the reductions in magnitude of the associations between high copy number and HSIL being consistent with a tendency for viral integration to become a dominant feature as high-grade lesions develop. Our analysis of sub-categories of lesion grades seemed to corroborate these findings. Sherman et al. [*228*] observed a similar decline in viral load measured by HC-II with increasing lesion severity defined by colposcopy or histopathology.

We further observed that while viral load has substantial value as a risk stratifier for incident LSIL and less for HSIL, the short-term (4 to 12 months) predictive value is influenced by the underlying cytological abnormalities. The associations were present for those with smears within normal limits and notably absent for those with ASCUS in their specimens. The lack of prediction by HPV viral load among women with ASCUS results by cytology may be because some of these classifications already represent an initiation of a lesion disguised as ASCUS. Therefore, these results may be a reflection of the aforementioned hypothesized relationship, that while viral burden may increase the likelihood of initiation of a dysplastic event, it may not play as strong a role in the subsequent malignant transformation of infected cells into neoplasia [*260*].

In conclusion, in women who do not already exhibit morphological changes by cytology, measurement of viral load, as a surrogate for HPV persistence, may identify women at increased risk of developing cervical lesions. However, the clinical utility for measures of viral burden remains to be confirmed by histological evaluation in screening studies.

Table II-1 – Mean viral load and 95% CI in HPV-positive cervical specimens according to incident SIL persistence over 8 years of follow-up[1]

Incident SIL	Number of women	Copy frequency[2]	95% CI	p^3
No occurrence	352	2.2	1.7-2.8	
Only 1 positive Pap test	44	8.9	3.9-20.6	0.0003
Two positive Pap tests	8	6.2	1.1-33.2	0.2117
Three or more positive Pap tests	13	18.7	3.1-112.6	0.0014

[1] SIL ascertained by cytology at return visits over 8 years, excluding prevalent cases of SIL at enrollment.

[2] Viral load (copies/cell) calculated as the average between viral loads at enrollment and first follow-up visit.

[3] T-test statistical significance for pooled data with comparison to the mean viral load in women with no SIL events.

Table II-2 – RR and 95% CI of any incident SIL during 8 years of follow-up according to average viral load in 2 initial specimens[1]

Average viral load[2] (copies/cell)	Any SIL			Only HSIL		
	Lesion event/ No lesion	RR	95% CI	Lesion event/ No lesion	RR	95% CI
<1	23/224	1.0	Referent	7/243	1.0	Referent
1-10	8/39	1.87	0.8-4.2	1/48	0.67	0.1-5.4
11-100	15/40	3.38	1.8-6.5	3/55	1.71	0.4-6.7
101-1000	12/38	2.94	1.5-5.9	2/57	1.16	0.2-5.6
>1000	7/11	4.47	1.9-10.7	2/21	2.59	0.5-13.2
p for trend		0.000			0.309	

[1] RR and 95% CI derived by Cox regression excluding prevalent at enrollment cases of SIL and human papillomavirus (HPV) negative women. Adjusted for HPV types grouped by oncogenic potential and age.

[2] Calculated as the average between viral loads at enrollment and first follow-up visit.

Table II-3 – RR and 95% CI of different sub-categories[1] of SIL during 8 years of follow-up according to average viral load in the 2 initial cervical specimens[2]

Average viral load[3] (copies/cell)	LSIL/HPV			LSIL/SQ			HSIL/CIN2			HSIL/CIN3		
	Number of. samples	RR	95% CI	Number of samples	RR	95% CI	Number of samples	RR	95% CI	Number of samples	RR	95% CI
<1	226	1.0	Referent	229	1.0	Referent	230	1.0	Referent	230	1.0	Referent
1-100	82	2.32	1.3-4.1	82	1.79	1.0-3.2	86	1.38	0.7-2.8	89	1.23	0.6-2.5
>100	57	2.48	1.2-5.3	63	4.17	2.2-7.7	69	2.97	1.2-7.2	71	2.23	0.9-5.8

[1] According to the dominant cytological feature: cytopathic effects (LSIL/HPV), mild squamous dysplastic component present (LSIL/SQ), moderate dysplasia (HSIL/CIN2), and severe dysplasia or worse abnormalities (HSIL/CIN3)

[2] RR and 95% CI adjusted for HPV types grouped by oncogenic potential and age by Cox regression excluding prevalent at enrollment cases of SIL and HPV negative women.

[3] Calculated as the average between enrollment and first follow-up visit viral loads.

Table II-4 – Marginal OR and 95% CI for the prevalence and incidence of LSIL and HSIL according to viral load assessed at the same or at 1 or 2 previous visits[1]

Timing of viral load measurement	Viral load (copies/cell)	LSIL			HSIL		
		Number of samples	OR	95% CI	Number of samples	OR	95% CI
At same visit	<1	820	1.0	Referent	809	1.0	Referent
	1-100	272	3.91	2.2-7.1	264	5.73	2.4-13.9
	>100	248	9.32	5.2-16.7	198	3.90	1.5-10.2
At previous visit[2]	<1	653	1.0	referent	658	1.0	referent
	1-100	194	3.71	1.7-8.1	197	2.26	0.6-8.7
	>100	161	5.56	2.6-11.7	178	1.27	0.2-7.1
Two visits prior[3]	<1	682	1.0	referent	678	1.0	referent
	1-100	201	1.53	0.7-3.3	207	4.50	0.8-26.7
	>100	156	2.74	1.2-6.0	174	7.62	1.3-43.6

[1] OR and 95% CI adjusted for HPV types grouped by oncogenic potential and age using GEE regression models correcting for intra-subject correlation among HPV positive women (see text for details).

[2] HPV load at previous visit, excluding prevalent cases of SIL at index visit for exposure assessment.

[3] HPV load 2 visits prior, excluding prevalent cases of SIL at index visit for exposure assessment.

Table II-5 – Marginal OR and 95% CI for the incidence of any SIL according to viral load: results stratified by cytological status at the viral load assessment visit[1]

Timing of viral load measurement	Viral load (copies/cell)	Cytology status at viral load measurement visit[2]: Within normal limits — Number of. samples	OR	95% CI	ASCUS — Number of samples	OR	95% CI
At previous visit[3]	<1	637	1.0	Referent	18	1.0	Referent
	1-100	180	3.13	1.4-6.8	17	1.94	0.3-11.7
	>100	144	5.02	2.4-10.4	18	0.89	0.1-8.3
Two visits prior[4]	<1	665	1.0	referent	18	1.0	referent
	1-100	190	1.76	0.8-3.9	12	0.38	0.04-4.0
	>100	145	3.72	1.7-8.0	14	0.18	0.01-2.1

[1] OR and 95% CI adjusted for HPV types grouped by oncogenic potential and age using GEE regression models correcting for intra-subject correlation among HPV positive women (see text for details).

[2] Restricted to subjects with corresponding cytological classification at index visit for exposure assessment: 'within normal limits' includes normal smears and benign cellular changes; ASCUS: atypical squamous cells of undetermined significance.

[3] HPV load at previous visit, excluding prevalent cases of SIL at index visit for exposure assessment.

[4] HPV load 2 visits prior, excluding prevalent cases of SIL at index visit for exposure assessment.

6.5. Preface to Manuscript III

Most epidemiologic research on the natural history of HPV infection and cervical cancer has been based on only one measurement of exposure to the virus and its determinants or cofactors and on one measurement of cervical lesion end points. Case-control and cohort investigations have relied on different approaches to determining the baseline status for HPV and other factors and lesion outcomes, simultaneously, retrospectively, or prospectively. Statistical modeling by logistic and proportional hazards regression methods has enhanced the ability to probe associations in epidemiologic datasets by allowing control of confounding, assessment of interaction among variables. However, it is only with multiple repeated assessments that methodological errors in sampling and cytology testing, and effects of temporal fluctuations in detectability of HPV during the course of infection, can be corrected for.

It has been demonstrated that, in traditional epidemiologic designs, incomplete measurement of cumulative exposure to HPV may make it unreliable to use the magnitude of the RR estimates for the association between HPV and cervical neoplasia to evaluate the degree with which HPV acts as a necessary cause of the disease [72]. Sources of misclassification of HPV can arise in two ways: 1) methodologic or laboratory error in assessing the true cervical HPV infection status of a given cohort participant, due to problems arising from cervical sampling, specimen processing, or accuracy of testing, or 2) erroneous measurement of cumulative exposure to HPV infection on the basis of testing of a single specimen only. Many previous studies on HPV and CIN incidence have been retrospective in design and did not collect information on exposure to HPV on more than one occasion. As a result such studies are unable to examine consistently the magnitude of effect of HPV on CIN, nor can they determine when during the natural history of the disease the virus plays an active role in the carcinogenesis of the squamous epithelium.

In this paper I investigated the pitfalls of traditional epidemiologic approaches to measuring the role HPV infection in cervical neoplasia by exploring in detail the time-dependent characteristics of this association. In particular I try to show that although the observed differences in magnitude of RR observed in previous cohort studies can be due to the use of imperfect methods of outcome classification, the effect of mismeasurement of CIN may not be the main culprit in cohort or case-control studies. Even with highly sensitive PCR-based methods of measuring the presence of HPV DNA in cervical specimens, the delay with which exposure assessments are made with respect to the time of disease ascertainment can affect considerably the magnitude of effect by HPV. This problem becomes more serious when one relies on single measurements of exposure.

6.6. Manuscript III: Modelling the time dependence of the association between human papillomavirus infection and cervical cancer precursor lesions

(Running title: Time-dependent association between HPV and cervical neoplasia)

Nicolas F. Schlecht [1,2], Robert W. Platt [2,3], Abdissa Negassa [5], Eliane Duarte-Franco [1], Thomas E. Rohan [5], Alex Ferenczy [4], Luisa L. Villa [6], Eduardo L. Franco[1,2]

Author Affiliations:

[1] Departments of Oncology, [2] Epidemiology & Biostatistics, [3] Pediatrics and [4] Pathology, McGill University, Montreal, Canada;

[5] Department of Epidemiology and Social Medicine, Albert Einstein College of Medicine, New York, USA;

[6] Ludwig Institute for Cancer Research, São Paulo, Brazil.

III

ABSTRACT

We studied the time-dependent association between human papillomavirus (HPV) infections and squamous intraepithelial lesions (SIL) among women enrolled in a cohort study using repeated measurements for SIL by cytology and HPV testing by polymerase chain reaction. We investigated different Cox regression modelling approaches to assess the effect of varying interval times by mimicking alternative cohort study designs embedded into the repeated measurements investigation. Associations between HPV status and early lesion events were of high magnitude. The age-adjusted hazard ratios (HRs) for the association between HPV status at enrollment and any-grade SIL decreased gradually with time until 72 months for both oncogenic (HR=4.12, 95% CI: 2.7-6.3) and for non-oncogenic HPV types (HR=2.39, 95% CI: 1.4-4.1). The HR for incident high grade SIL (HSIL) remained constant over time ranging from 7.15 (95% CI: 2.0-25.1) at 12 months to 6.26 (95% CI: 2.7-14.5) at 72 months for oncogenic HPV types. With HPV as a time dependent predictor, the HR for the association of oncogenic types with incident SIL and HSIL events were 14.2 (95% CI: 8.7-23.1) and 32.7 (95% CI: 8.4-127.3), respectively. We may underestimate the prognostic value of HPV detection using designs that rely on HPV ascertainment at a single point in time. The waning in HRs should be considered in the implementation of screening programs based on HPV detection.

** This manuscript was submitted for publication to the Amer J Epidemiol, November 2002*

INTRODUCTION

The association between HPV DNA and cervical cancer has been well documented worldwide [*16*]. In the natural history of the disease, preinvasive lesions of cervical neoplasia can be transient and reoccur over time. Women may develop low and high grade squamous intraepithelial lesions (SILs); LSIL may progress on to HSIL or may regress back to a normal state [*109*]. Human papillomavirus (HPV) infection is the main etiological agent in the initiation of this process [*16*]. While the likelihood of a lesion persisting or progressing to cancer can be dependent on the characteristics of HPV infections [*142;185;248;252*], the strength of the observed association with HPV can be influenced by methodological issues such as the measures of cumulative exposure that are used in the study design and analysis [*67;72*].

To date, cohort studies of HPV and SIL have differed in the frequency with which subjects are re-evaluated over time for exposure and outcome status [*106;107;131;138;171;174;185;255;262*], as well as in the time elapsed between these two measurements [*99;268*]. While all of these studies support a causal relationship between HPV and incidence of SIL, there is a lack of consensus on the magnitude of the effect and little is known about the possible waning of the risk association over time. Furthermore, studies that rely on single measures of HPV infection may be more susceptible to misclassification of viral exposure given the transient nature of many HPV infections [*106*].

Longitudinal studies with repeated measures present a unique challenge for the statistical analysis of observational data and the investigation of disease associations due to the inherent correlation between measures [*52;243*]. The newer statistical approaches that have been developed for the analysis of such data contrast with the simplistic approach, which involves collapsing the information on repeated events into one or two summary measures [*81*]. The latter approach does not make use of all available data collected at repeated intervals for each individual and ignores the time dependence of the epidemiological association between events.

In the study reported here, we analyzed data from an epidemiological investigation involving repeated measurements of HPV infection status and cervical lesion outcomes over time among women attending a comprehensive maternal and child health program catering to low income families in the northern sector of the city of São Paulo, Brazil. Using algorithms that mimic different study layouts embedded in our cohort investigation we analyzed the longitudinal relationship between HPV and cervical neoplasia to examine the strength of this association overall and as a function of time between exposure assessment and outcome events.

METHODS AND MATERIALS

Subject recruitment

A detailed description of the design and methods of the Ludwig-McGill cohort study and characteristics of subjects have been published previously [*62;224*]. From November 1993 to March 1997, two study nurses selected a systematic sample of women to be approached for interview from daily lists of outpatients in the family medicine, gynecology, and family planning clinics at the "Vila Nova Cachoeirinha" municipal hospital. Eligible women were given a detailed overview of the study and invited to participate. Eligibility criteria included: (i) age between 18 and 60 years; (ii) permanent residence of São Paulo (city); (iii) not currently pregnant and no intention of becoming pregnant during the next 12 months; (iv) having an intact uterus and no current referral for hysterectomy; (v) no use of vaginal medication in the previous 2 days; and (vi) not treated for cervical disease in the previous 6 months. In addition to these criteria, women were considered ineligible if they were not interested in complying with all scheduled returns, at least for the subsequent 2 years.

Subjects entered the study only after giving signed informed consent. The study protocol was approved by institutional ethical and research review boards of the participating institutions in Canada and in Brazil. All participants were seen every 4 months in the first year (0, 4, 8 and 12 months), and twice yearly thereafter. Cervical specimens were taken for Pap cytology and HPV testing at every visit. The study nurses also performed a detailed interview to collect information on sociodemographic factors, reproductive health, sexual activity, smoking, and diet at enrollment. Information on sexual activity and reproductive health was also collected at each return visit during the first 12 months and at subsequent yearly returns.

Cervical cell specimens

An Accelon biosampler (Medscand, Hollywood, FL) was used to collect an ecto- and endocervical sample at each of the visits. After the cells were smeared on a glass slide and fixed for cytology, the sampler containing the residual exfoliated cells was immersed in a tube containing Tris-EDTA buffer pH 7.4. Samples were then sent to the Ludwig Institute for Cancer Research for storage and testing. The Pap smears were shipped to Montreal for coding and classification by an expert cytopathologist (AF). The cytopathology reports were based on the 1992 Bethesda system for cytological diagnoses [*97*]. Readings were blinded to previous cytology outcomes and to HPV results for the same women.

HPV DNA detection

Cervical specimens were tested for the presence of HPV DNA by a standardized polymerase chain reaction (PCR) protocol [*10;105*]. Typing of the amplified products was performed by hybridization with individual oligonucleotide probes specific for all 27 HPV genital types whose

nucleotide sequences for probes within the MY09/11 fragment have been published in the literature [*105*]. Amplified products that hybridized with the generic probe but with none of the type-specific probes were tested further by restriction fragment length polymorphism (RFLP) analysis [*12*] to increase the number of identifiable HPV types. To verify the specificity of the hybridizations, we included more than 30 type-specific positive controls in all membranes. In order to check the integrity of the host DNA material extracted from the specimens, assays also included an additional set of primers to amplify the ß-globin gene [*10*]. HPV types were separated into two groups by presumed level of oncogenicity. Oncogenic types included HPVs 16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, 59, and 68; non-oncogenic types included 6/11, 26, 32, 34, 40, 42, 44, 53, 54, 55, 57, 62, 64, 66, 67, 69, 70, 71, 72, 73, 81, 82, 83, 84, CP6108, plus other unknown types. All HPV assays were done on coded specimens with no identification linking specimens from the same woman. Appropriate precautions were taken to reduce the possibility of specimen contamination.

Statistical Analyses

For the analyses reported here, follow-up continued until March 2002, the development of neoplasia requiring treatment, death, or loss to follow-up, whichever occurred first. Time to event was measured from the date of enrollment to the date of first occurrence of a lesion (as defined below) or to the last recorded return visit date for censored subjects.

In order to assess the time dependence of the association between HPV and lesion outcomes we evaluated different analysis approaches based on distinct design layouts that were conceivable as part of the repeated measurement dataset of the Ludwig-McGill cohort (Figure III-1). We first looked at the association between HPV infections at enrollment and the occurrence of cervical lesions that could be documented exclusively within a specified period of time, with follow-up ending (i.e., ignored) after that date (Figure III-1, Model A). A second approach involved initiating follow-up for outcome events later in the study, after a period of delay had elapsed, and associating HPV infection status at enrollment with the occurrence of cervical lesions occurring only after a specified period of time (i.e., outcomes occurring before that time were ignored) (Figure III-1, Model B). Models A and B assume traditional cohort analyses of single point assessment of exposure at enrollment and a first documented instance of outcome based on the respective layout restrictions (model B addresses the possibility that the outcome was present at the time that exposure was assessed). The traditional Cox regression model was used for this purpose.

In order to produce a cumulative estimate of association over time while accurately representing the transient nature of HPV infections, this model was extended to incorporate a time-varying

measure of HPV infection by modeling the hazard for changing HPV status at each visit (Figure III-1, Model C). The constraint with all of the previous three models is that instances of SIL after the first event are not considered. Therefore, a fourth model approach (Figure III-1, Model D) was used, which involved correlating HPV infection status and lesion incidence at different, fixed follow-up returns. This was performed using two points of assessment; one for exposure corresponding to time t_0 and the other for outcome assessed a specified number of months later at time t_0'. In this layout, subjects could contribute multiple periods of observation for each time span combination, in which both the exposure and outcome assessment visits were allowed to vary to make up strata of equivalent follow-up duration. The diagram in Figure 1 depicts an example of model D with a period of observation of 12 months. Two approaches, generalized estimating equations (GEE) [*271*] and marginal hazards regression [*243*], were used to estimate relative risks (RRs) of SIL while taking into account the clustering within each individual implied in model D. In the GEE approach, correlations between outcome events are treated as nuisance parameters, thereby allowing for inference based on the coefficients for the covariates in the model that can be either time-dependent or time-independent – in this case, HPV infections and age at enrollment, respectively. All models incorporated an exchangeable or equal correlation pattern for the repeated events. We used the marginal hazards regression model to verify the consistency of estimates resulting from the GEE approach. We adapted this model [*243*] to the analysis of pairwise comparisons between HPV results at any given visit and occurrence of SIL at different interval returns (Model D). Stratification in this model was performed to allow the baseline hazard to vary with each period of observation defined by the index visit where HPV was tested. The hazard function [$h(t,X[t])$] at time t for an individual with the vector of explanatory variables $X(t)$ is described by the following formula:

$$h(t, X[t]) = h_{oj}(t) \times e^{\sum_{i=1}^{p} \beta i Xi(t)}$$

where $=h_{0j}(t)$ is an arbitrary and unspecified baseline hazard function for each period of observation (j); βi is the regression parameter associated with the i^{th} explanatory variable, and $Xi[t]$ is the i^{th} explanatory variable, i=1,...,p (i.e. HPV status at time t).

Different restriction criteria for excluding lesions at baseline visits were assessed in Model D analyses using the above approach. In the least restrictive analyses all subjects were included, regardless of baseline status by cytology. A second set of analyses used a subset of subjects after removing prevalent cases of SIL at enrollment. The most restrictive analyses included only pairwise arrangements in which the baseline visit did not reveal cytological abnormalities. A robust variance was estimated in all models, clustering on each individual.

For all models (except GEEs), we estimated the RR of lesion occurrence given HPV infection status by computing hazard ratios (HR) and respective 95 percent confidence intervals (CI). We

examined proportionality of the hazards for the traditional Cox regression models using Schoenfeld residuals, and fit models for non-proportional hazards using the extended Cox model [*243*]. Statistical analyses were performed using the statistical programs SPSS® version 11.0 (SPSS, Chicago, IL), STATA® versions 6 and 7 (StataCorp, College Station, TX) and S-PLUS® for Windows version 6.0 (Insightful, Seattle, WA).

RESULTS

Between November 1993 and March 1997, 2528 women were enrolled into the study, corresponding to a 70 percent response rate. Women who were found not to fit the eligibility criteria after recruitment (N=66) were excluded after enrollment. The remaining 2462 women participating in the study were followed up at repeated scheduled returns for a period of up to 8 years. Fifty-one (2.1 percent) women presenting with a prevalent lesion at enrollment and six women with inconclusive cytology results were excluded from the analysis. The total follow-up time for subjects with no lesions at enrollment was 128,129 women-months. Taking into account time-to-event follow-up data used in the Cox regression models, the total follow-up time decreased to 122,299 women-months. After excluding only cases of HSIL at enrollment, the total follow-up time was 128,624 women-months.

A decrease in the effect of HPV over time was evident when we looked at the association between HPV infections at enrollment and the occurrence of SIL over time (Table III-1). This approach is analogous to performing different prospective cohort studies with varying periods of follow-up from 4 months to 72 months in which multiple cervical samplings were made within the assumed study duration (Figure III-1, Model A). When we consider all incident SIL events that occur over time after enrollment until the indicated time has elapsed, we see a decreasing trend in HR with time from 6.49 for non-oncogenic HPV types after an interval of 4 months to HR=2.39 after 72 months. A similar trend is observed for oncogenic HPV types, with HRs decreasing from 7.15 to 4.12, over the same time intervals. Point estimates for different periods of follow-up return were higher for the association with incidence of HSIL over time. Although the initial association for oncogenic HPVs was very high over the first 8 months of follow-up, the HR then decreased and remained between 6.26 and 7.83 for study periods reaching from 12 months to 72 months follow-up. While HRs for non-oncogenic HPV were lower, a similar pattern of risk was observed with increasing follow-up time.

A second cohort analysis approach involves initiating follow-up of subjects later in the study after a period of delay (Table III-2). This approach mimics a study design with delayed initiation of follow-up whereby occurrences of SIL are associated with an HPV test result taken several months or years earlier (Model B). The magnitude of the association between HPV infection at

enrollment and SIL is lower than that seen in equivalent intervals in table 1 and decreases as model B analyses lengthen the interval between exposure and first assessment of outcome. When we looked at the risk effect of a persistent oncogenic infection, defined as being positive for the same oncogenic types at both enrollment and first follow-up visits (data not shown), the HRs decreased from 8.35 (95% CI: 5.0-13.9) after 4 months to 5.60 (95% CI: 1.6-19.8) after 48 months. The effect was weaker for persistent non-oncogenic infections (HR=3.14, 95% CI: 1.6-3.9 and HR=1.87, 95% CI: 0.2-14.3 for 4-month and 48-month delays, respectively). For the occurrence of HSIL, no decreasing trend in HR was demonstrated with increasing interval delay. While the stability of the point estimates decreased with interval period, the HR for oncogenic HPV remained high at 5.4 (95% CI: 2.3-12.6) after a delay of only 4 months, and stayed at the same level (HR=5.6; 95% CI: 0.5-59.4) even after a delay of 60 months.

Cox proportional hazards models with time-dependent covariates incorporate the transient nature of HPV infections by updating the HPV results over the course of follow-up to reflect the latest infection status observed at the previous visit (Model C). The maximum follow-up that occurred was 3 years because HPV testing results were available only for the first two years of follow-up. Compared with the previous analyses using a single HPV assessment, table 3 shows substantially higher HRs for SIL events, with HRs of 5.7 for non-oncogenic types and 14.2 for oncogenic types. The equivalent associations with HSIL were of much greater magnitude.

Model D analyses considered multiple measurements of HPV infection and lesions within individuals. Among those with no indication of a prevalent lesion by cytology at enrollment, 38 had two or more SIL events during the period of follow-up. The number of subjects contributing multiple HSIL events was lower (8 people) and events were concentrated in only a few people (of the 8, 6 had two events, one had 4, and one had 7). Although correlations between multiple events are unspecified we allowed for a varying baseline hazard for each observation period depending on which follow-up visit is considered. This approach was judged to be statistically more conservative than stratifying on event order. We observed associations of similar magnitude for the occurrence of cervical lesions 4 or 6 months after a positive HPV result (Table 4) to those observed by the traditional analyses for first occurrence of SIL. Effect estimates by the two regression models (GEE and marginal hazards) were generally comparable although they differed in definition, i.e., odds ratios for GEEs and HRs for Cox regression for marginal data. RRs begin somewhat higher for oncogenic HPV types than for non-oncogenic types after an average period of 4 months followed by a drop in RR at 8 months. A similar drop is seen at an average follow-up time of 6 months after testing for non-oncogenic HPV types. More importantly, table 4 also shows a trend of decreasing RRs for both non-oncogenic and oncogenic HPV types as cohort interval time increases. We also repeated all models in table 4 by using more or less

stringent criteria for restricting subjects with lesions at baseline. Point estimates using both GEEs and marginal hazards were comparable to the ones shown in table 4, regardless of whether subjects with prevalent lesions were included or when all baseline lesions were excluded at the beginning of each observation period (data not shown). The latter method was the most restrictive since it only included in the analyses time segments of exposure-outcome observations in which no lesions existed at the time of exposure assessment, effectively forcing repeated events to occur only after an interval period without SIL incidence.

DISCUSSION

Using the repeated measurements made over time in our study, we were able to analyze the dynamic status of cervical lesion outcomes with respect to similar changes in status over time for HPV testing results. Previous studies of the natural history of HPV and cervical cancer have also used statistical methods for longitudinal data, to a limited extent, to analyze either HPV infections [*1*] or precursor lesions [*107*] as outcomes [*141;262*]. Although potentially more powerful, such statistical approaches have some additional conditions and assumptions that are not required in traditional cohort study designs [*33*].

The objective of this study was to evaluate the time-dependence of the association between HPV infections and the risk of SIL incidence by decomposing our cohort data to mimic different cohort study designs. First, we replicated traditional cohort designs of baseline assessment of exposure and delayed outcome incidence with increasing durations of follow-up that either assumed a termination date (outcomes after that date being ignored, i.e., Model A) or imposed a specified period of time during which no outcome ascertainment was made (outcomes before this period of time being ignored, i.e., Model B). These two approaches are based on a single time assessment of HPV exposure (i.e., at enrollment) and first incident lesion (at a given interval restriction), and thus they ignore most of the exposure and lesion information collected during the study. We then resorted to models that supplemented the enrollment HPV status with the post-enrollment HPV testing data and correlated the resulting combined information with risk of incident lesions (i.e., Model C). Finally, we used longitudinal designs with repeated assessments of HPV and SIL over time and varying time intervals between exposure and outcome (Model D).

The hallmark of our findings is the decline in the magnitude of the association between baseline HPV and risk of subsequent SIL with increasing duration of follow-up. RRs might be attenuated by the occurrence of remote SIL events (i.e., events later in the study) as a result of new HPV infections occurring after enrollment. As a result, traditional cohort analyses that rely on single measures of exposure may underestimate the HR when we try to associate prevalent HPV infections with SIL events several years later (Model A). The hazard associated with initial HPV

status is high early on but declines with time, suggesting that HPV status at baseline is closely associated with early events but almost unassociated with later events. The slight association with exclusively distant events (Model B) could be ascribed to the correlation between current and new HPV infections after the primary visit. We cannot exclude, however, the possibility that the point estimates may be influenced by false negative HPV tests as HPV infections are transient, for the most part.

When we evaluate the pairwise associations between a measurement for HPV and SIL at different points in time following an HPV measurement (Model D), we see the magnitude of the association drop at 6 to 8 months after HPV assessment for non-oncogenic HPV types and at 8 months after HPV assessment for oncogenic types. A return to initial levels of association does not occur until after 12 months follow-up for either type. The model D approach was also useful in testing the hypothesis that the quality of the HPV testing in our study did not change over time. Although we used a standard PCR protocol for all tests, small technical variants have been introduced over the long time span (nearly 10 years) involving laboratory work in this project. Enrollment HPV results were generated 1-4 years before HPV tests done in later follow-up specimens. The fact that model D results largely replicated the findings from models A and B suggest that the quality of our viral testing did not drift over time.

Traditional cohort analyses that rely on a single measure of HPV exposure obtained on enrollment present a design situation analogous to nested-case-control studies carried out on registries of passive screening data derived from a passive follow-up (Model A). Two such studies by Wallin et al. [*255*] and Liaw et al. [*138*] observed odds ratios (ORs) of 15.0 (95% CI: 0.8-1541) for the incidence of cervical cancer an average of 6 years after an HPV-DNA positive normal Pap smear and 12.7 (95% CI: 6.2-25.9) for incident HSIL within 5 years of follow-up after an HPV positive normal Pap smear, respectively. Taking an alternative follow-up approach, we witness a similar decline in HR associations as the time interval between positive HPV results and cervical lesion incidence becomes longer. That is, the likelihood of a new SIL or a repeat LSIL event decreases as the opportunity for detecting such events by cytology is delayed. This method reproduces the situation encountered in studies with passive [*35;268*] (Model D) or delayed follow-up of subjects [*142*] (Model B). We observed decreasing HRs associated with SIL occurring later during follow-up for both oncogenic and non-oncogenic HPV infections at enrollment. Considering persistent type-specific oncogenic HPV infections for the first four months of follow-up, the association with incident SIL or repeat SIL events continues for a longer period of time with appreciably higher HRs being observed for incident events 48 months later. This is not seen for non-oncogenic persistent infections (data not shown).

A second important observation in this study is the divergent pattern of risk association observed for HSIL compared to any SIL. While point estimates for HSIL were less stable due to the small number of events, they were indicative of a constant risk relation over time for both HPV infection variables. The majority of incident HSIL events were preceded by an LSIL or ASCUS smear (18/27). The magnitude and persistence of the HR over time is consistent with a causal hypothesis. Other studies [*262*] have observed a decline in risk of HSIL by cytology in relation to time since first detection of HPV from less than 6 months (RR: 25.33, 95% CI: 8.81-72.83) to over 12 months (RR: 6.42, 95% CI: 2.10-19.65). Ylitalo et al. [*268*] showed decreasing ORs across levels of HPV 16 viral load measured over one to nine years before diagnosis of CIS from histological samples. Even though previous studies have shown oncogenic HPV types to be more likely to persist than other types [*105;210*] we cannot discount the role of false positive and negative test results in these associations and those of previous studies.

Studies have tried to address this attenuation in effect with time by incorporating time-varying measures of HPV status into their statistical models of cervical neoplasia (Model C) [*106;141;171*]. Use of statistical methods that allow time-dependent covariates (e.g., GEEs) in these studies demonstrate a clinical utility to HPV testing for predicting the incidence of new or recurrent cervical lesion following a recent HPV infection. In similar cohort studies by Ho et al. [*106*] and Moscicki et al [*174*], the authors observed elevated ORs for SIL following a positive infection for oncogenic HPV types detected 6 months and 4 months earlier, respectively.

A recommendation of GEE and marginal models is the assumption that the data conform to a particular balanced structure. These models lose efficiency if, as follow-up continues, more individuals are measured irregularly. Furthermore, while the correlation matrix can differ from subject to subject, missing data must be missing at random [*52*]. The occurrence of missing values has been minimized as much as possible in this study as delays in returning for a given appointment were permitted although this created longer interval periods for some subjects. We observed similar point estimates by both regression approaches, i.e., GEE and marginal hazards, which gives credence to the validity of the findings by either model.

Although outcome assessment was carefully conducted in a reference laboratory following a strict quality control protocol and without knowledge of exposure status, cytological misclassification may have resulted in an attenuation of the RR estimates. We opted for an intensive, expert cytological follow-up every four to six months of all Pap smears collected in the study to avoid having to perform unnecessary biopsies, which may interfere with the early natural history of lesions. The inclusion of intensive repeated cytological testing over time also permitted us to evaluate recurrence of lesion

events. However, we cannot be sure that a repeat detection of an event represents a recurrence or persistence of SIL, even if no lesion is detected at interval assessments by cytology.

In epidemiological studies such as this we rarely know the disease history of women following a first exposure to HPV in cohort studies of SIL. As a result, we and others often exclude prevalent cases detected at enrollment from the analysis [*30;106;131;138;174;224;266*] and assume that women with negative cytology lesions at enrollment are at similar risks of lesion incidence. The majority of the SIL events in our study began as LSIL. While the progression of LSIL to HSIL to cancer may depend on other factors in conjunction with HPV, the detection of SIL could simply fluctuate with HPV status. The short-term association observed between HPV and LSIL may result from the relatively easy recognition by cytology of the cytopathic effects induced by HPV [*83;229;232*]. However, a relationship between HPV and incidence of SIL over an extended period of time has also been observed [*171*] suggesting that there exists a temporal risk effect. When estimating the association between HPV and SIL events, we can gain precision by taking into account the transient nature of the disease. The incorporation of repeat SIL events in our analyses, achievable due to the length of follow-up in this study, marginally improved the model power (as indicated by the narrowing of the CI) without changing the estimation of effects (data not shown).

In conclusion, we observed a decrease in the magnitude of the association between baseline HPV detection and SIL over time. However, allowing for the dynamic nature of both the exposure and outcome yielded larger RR estimates for corresponding time points. As a result, use of traditional cohort design and analysis techniques relying on ascertainment of exposure at a single point in time may result in an underestimation of the effects of HPV, even when one is testing only for high risk, oncogenic viral types. Future assessments of the empirical effects of HPV infection in long term follow-up studies may require more complex algorithms to quantify past exposure to the virus (e.g., inclusion of viral load, long term persistence and type-specific information). In addition, future implementation of screening programs based on HPV testing will have to consider the waning of the predictive value of viral testing over time in detecting clinically significant lesions.

Table III-1: HRs and 95 percent CIs for the association between non-oncogenic and oncogenic HPV infection *at enrollment* and any grade SIL detected within specified periods of follow-up*

SIL detected		Non-oncogenic HPV			Oncogenic HPV	
within	No. events SIL/ HSIL	Any SIL	HSIL	No. events SIL/ HSIL	Any SIL	HSIL
Enrollment†	3/0	5.29 (1.4-20.3)	-	35/16	46.27 (20.9-102.4)	49.5 (16.1-152.2)
4 months	5/1	6.49 (2.1-19.8)	-	9/3	7.15 (2.7-18.8)	-
8 months	9/1	4.46 (2.1-9.7)	10.65 (0.7-172.8)	14/3	5.14 (2.6-10.0)	23.53 (2.3-238.5)
12 months	11/1	3.77 (1.9-7.5)	1.92 (0.2-16.5)	19/5	5.46 (3.1-9.7)	7.15 (2.0-25.1)
24 months	13/2	3.82 (2.0-7.2)	3.00 (0.6-14.5)	22/7	5.41 (3.2-9.2)	7.83 (2.7-22.8)
36 months	14/3	2.78 (1.5-5.0)	3.78 (1.0-14.4)	27/8	4.65 (2.9-7.4)	7.57 (2.8-20.5)
48 months	14/3	2.41 (1.3-4.3)	2.85 (0.8-10.3)	29/10	4.32 (2.8-6.7)	7.15 (3.0-17.1)
60 months	16/3	2.45 (1.4-4.2)	2.67 (0.7-9.5)	30/10	4.06 (2.7-6.2)	6.69 (2.9-15.7)
72 months	16/3	2.39 (1.4-4.1)	2.50 (0.7-8.8)	31/10	4.12 (2.7-6.3)	6.26 (2.7-14.5)

* Traditional Cox Proportional Hazards model with robust variance for events occurring by the scheduled return visit adjusted for age, excluding prevalent cases at enrollment. Events up to and including the indicated scheduled visit considered according to model A (see text for details).

† Logistic regression for the cross-sectional relationship between HPV and cytology determined at enrollment (adjusted for age).

Table III-2: HRs and 95 percent CIs for the association between non-oncogenic and oncogenic HPV infection *at enrollment* and any grade SIL detected after specified periods of postponement in follow-up*

SIL detected		Non-oncogenic HPV			Oncogenic HPV	
after	No. events SIL/ HSIL	Any SIL	HSIL	No. events SIL/ HSIL	Any SIL	HSIL
4 months	14/3	2.11 (1.2-3.7)	2.42 (0.7-8.6)	28/9	3.76 (2.4-5.8)	5.35 (2.3-12.6)
8 months	12/3	2.15 (1.2-4.0)	2.64 (0.7-9.4)	21/8	3.45 (2.1-5.7)	5.23 (2.1-12.9)
12 months	7/2	1.49 (0.7-3.3)	1.95 (0.4-8.8)	18/8	3.55 (2.1-6.1)	5.81 (2.3-14.6)
24 months	4/2	1.05 (0.4-2.9)	2.78 (0.6-13.2)	9/4	2.27 (1.1-4.7)	4.30 (1.3-14.5)
36 months	3/1	1.02 (0.3-3.3)	2.20 (0.3-19.0)	7/4	2.34 (1.0-5.3)	6.96 (1.8-26.4)
48 months	2/0	1.45 (0.3-6.3)	-	3/3	2.28 (0.7-7.9)	14.89 (2.4-91.8)
60 months	1/0	1.12 (0.1-8.7)	-	1/2	1.15 (0.1-9.0)	5.61 (0.5-59.4)

* Traditional Cox Proportional Hazards model with robust variance for events occurring only after the indicated follow-up time adjusted for age, excluding prevalent cases at enrollment. Interval SIL events after enrollment but before the indicated period of follow-up were not considered according to model B (see text for details).

Table III-3: HRs and 95 percent CIs for the association between current status for non-oncogenic and oncogenic HPV infection *at a given visit* and first instance of SIL over the first 3 years of follow-up

HPV status at given visit	Traditional time-dependent Cox model*			
	First SIL event		First HSIL event	
Negative	1.0	(reference)	1.0	(reference)
Only non-oncogenic type	5.68	(3.0-10.7)	11.33	(2.2-57.3)
Any oncogenic type	14.20	(8.7-23.1)	32.69	(8.4-127.3)

* Cox Proportional Hazards model for the first lesion event at any scheduled return visit with time-dependent variables for HPV status adjusted for age, excluding prevalent cases at enrollment according to model C (see text for details).

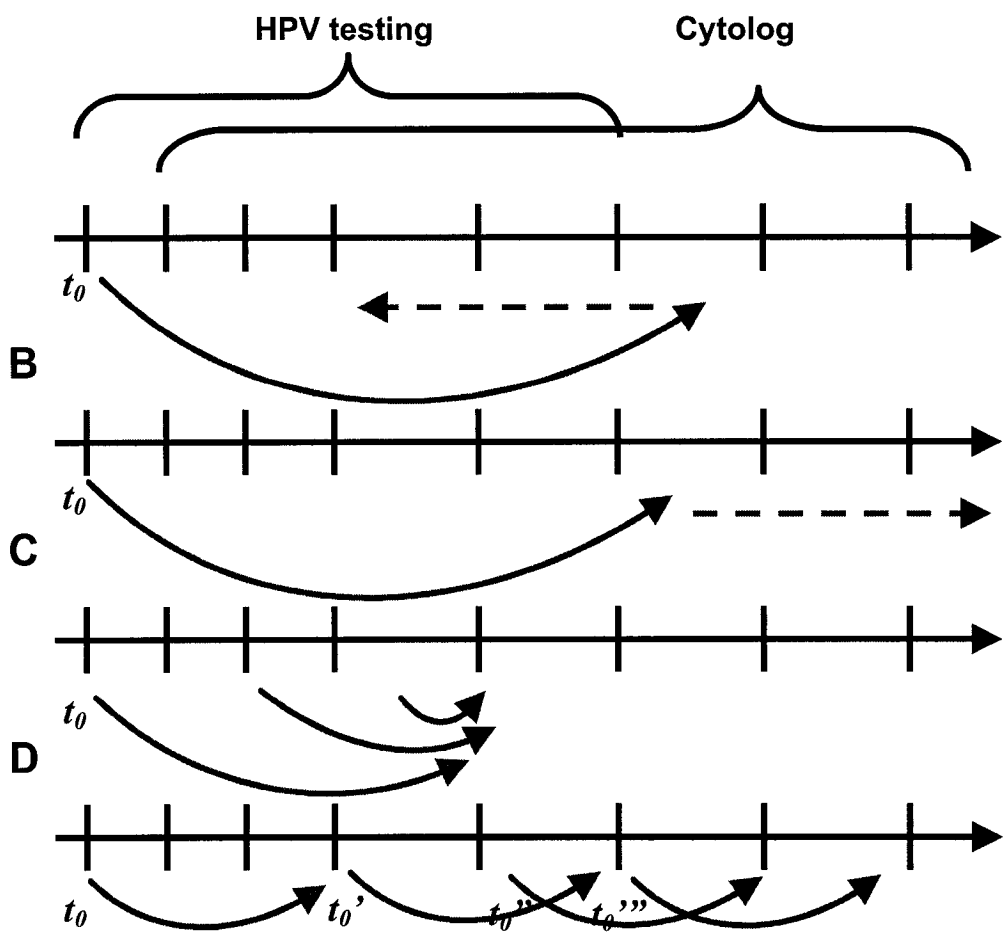
Table III-4: RR estimates and 95 percent CIs for the association between non-oncogenic and oncogenic HPV infection *at a given visit* and any grade SIL detected at scheduled follow-up returns*

SIL detected *at*	Marginal hazards model*				GEE regression model†			
	Non-oncogenic HPV		Oncogenic HPV		Non-oncogenic HPV		Oncogenic HPV	
	HR	(95% CI)	HR	(95% CI)	OR	(95% CI)	OR	(95% CI)
0 months (same visit)	NA		NA		31.09	14.2-67.8	48.57	23.7-99.5
4 months	6.14	(3.1-12.1)	13.03	(7.6-22.5)	6.60	(3.2-13.5)	15.76	(9.0-27.5)
6 months	4.39	(0.7-25.8)	14.32	(4.6-44.9)	1.92	(0.1-42.1)	10.60	(3.2-35.5)
8 months	3.08	(1.3-7.4)	5.33	(2.8-10.2)	3.13	(1.2-8.0)	5.74	(2.7-12.1)
12 months	5.56	(2.5-12.2)	11.81	(6.4-21.9)	4.10	(1.8-9.3)	10.60	(5.6-20.1)
18 months	6.97	(2.8-17.2)	8.95	(4.0-20.1)	6.78	(2.7-16.9)	9.94	(4.5-22.2)
24 months	3.85	(1.6-9.5)	7.34	(3.6-15.1)	3.19	(1.3-7.7)	5.58	(2.5-12.4)
30 months	1.90	(0.4-8.9)	7.78	(3.8-16.0)	1.46	(0.2-8.7)	8.29	(4.0-17.3)
36 months	2.62	(1.0-6.7)	3.35	(1.6-7.0)	2.24	(0.8-6.1)	2.57	(1.1-6.1)
42 months	2.47	(0.8-7.4)	5.69	(2.9-11.3)	2.48	(0.8-7.7)	5.78	(2.8-11.9)
48 months	2.63	(0.5-13.0)	4.40	(1.6-11.8)	2.32	(0.5-10.6)	4.47	(1.7-11.9)
54 months	4.23	(1.3-13.6)	1.98	(0.5-7.4)	5.07	(1.8-14.2)	1.82	(0.5-6.7)
60 months	0.73	(0.1-6.3)	3.37	(1.0-11.1)	1.13	(0.2-7.2)	2.30	(0.5-10.8)

* Marginal hazards Cox regression analysis for repeated outcomes over unequal intervals of time with robust variance adjusted for age, excluding prevalent cases of SIL at enrollment. Interval SIL events before or after the indicated return visit are not considered according to model D (see text for details).

† Generalized estimating equation (GEE) regression analysis with logit link for binary outcomes adjusted for age (exchangeable correlation between repeated events), excluding prevalent cases of SIL at enrollment.

Figure III-1: Graphical representation of different potential cohort analyses for varied study designs with staggered and repeated assessments of HPV infection and SIL outcomes by cytology



6.7. Preface to Manuscript IV

In the natural history of cervical cancer, women can progress from a normal state where no neoplastic or pre-neoplastic changes are detected in the squamous epithelium, to varying states of cellular abnormalities in the cervical epithelium including CIS [*154*]. Women may develop SILs of low and high grade and progress on to CIS or may regress back to a normal state [*109*]. While some research on the rates of progression and regression of cervical neoplasia has been done [*109;152;155;162;164;187*], to date no one has evaluated sojourn time (preclinical duration of preinvasive neoplasia), or regression and progression of precursor lesions according to HPV infection status.

The analyses in the previous manuscript demonstrated a predictive effect of the presence of HPV DNA in cervical specimens prior to incidence of LSIL and HSIL. HPV persistence and high viral load increases the risk of developing a lesion over time. However, the association with HPV is reduced once the epithelium begins to show signs of koilocytosis or dysplasia. The time dependence of the association between HPV infection and SIL incidence indicates a progressive relationship between onset of infection and development of neoplasia over time. I therefore attempted to measure the differences in time to progression and regression according to HPV status prior to or on the date of detection of a cytological abnormality for subjects followed in the Ludwig-McGill cohort study. In context to the other manuscripts in this thesis where I focused on the magnitude of the various time-dependent relations of interest, the present article deals with actual time measurements obtained by Kaplan-Meier and life table analysis.

6.8. Manuscript IV: Human papillomavirus infection as a predictor of progression and regression of cervical cancer precursor lesions

(Running title: HPV and progression of SIL)

Nicolas F. Schlecht[1,2], Robert W. Platt[2,3], Eliane Duarte-Franco[1], Thomas E. Rohan[5], Alex Ferenczy[4], Luisa L. Villa[6], Eduardo L. Franco[1,2]

Author Affiliations:

[1] Departments of Oncology, [2] Epidemiology & Biostatistics, [3] Pediatrics and [4] Pathology McGill University, Montreal, Canada;

[5] Department of Epidemiology and Social Medicine, Albert Einstein College of Medicine, New York, USA;

[6] Ludwig Institute for Cancer Research, São Paulo, Brazil.

IV

ABSTRACT

Background: The duration of cervical cancer precursor lesions according to human papillomavirus (HPV) infection status is unknown. We estimated the rates of progression and regression and the sojourn time of cervical squamous intraepithelial lesions (SILs) according to HPV status.

Methods: We used data from a longitudinal study of the natural history of HPV infection and cervical neoplasia in the city of São Paulo, Brazil. Cervical specimens were taken for Pap cytology and HPV testing every 4-6 months over a period of 8 years. We employed actuarial analysis to measure the differences in time to progression and regression and to investigate the rates of disease progression and regression according to HPV infection status in index lesions.

Results: The study included 2404 women. Among those with no lesions at enrollment, respectively 119, 24, and 174, incident LSIL, HSIL and ASCUS events were detected. Of those with incident LSIL, 11 (9.3%) progressed to HSIL or worse with a mean time to progression of 85.7 woman-months. Time to progression for those with oncogenic HPV DNA in the index Pap smear was shorter compared with those with no infection (70.3 versus 83.5 women-months). Half of all LSILs regressed within 6 months. Mean time to regression to a normal result was longer for those with oncogenic HPVs compared with women with no HPV infection (17.3 versus 8.7 women-months). Similarly, a longer time to regression was observed for women with HPV infection and ASCUS or HSIL smears.

Conclusion: Precursor lesions of the cervix persist longer and are more likely to progress in women with oncogenic HPV infections. Using PCR testing for oncogenic HPVs may help in identifying those lesions that progress more rapidly.

INTRODUCTION

The natural history of cervical cancer involves changes in the cervical tissue from a normal state where no neoplastic changes are detected in the squamous epithelium to varying states of cellular abnormalities leading ultimately to cervical cancer [*154*]. This sequence forms the premise on which cytological screening for cervical cancer is based and corresponds to an underlying multistep carcinogenic process in the development of cervical intraepithelial neoplasia (CIN) [*162*]. Women with low-grade squamous intraepithelial lesions (LSILs) may progress to high-grade SIL (HSIL) and invasive cervical cancer or may regress back to a normal state [*109*]. Few studies of cervical neoplasia have evaluated lesion recurrence [*107;262*] or disease progression over time [*109*].

Identification of a biomarker that can influence the rate of progression or regression and the duration of the preinvasive stages of cervical cancer could be used to devise a strategy for targeted screening or chemoprevention. While human papillomaviruses (HPV) are the main etiological agents in the initiation of cervical neoplasia [*16*], to date no study has evaluated the regression or progression of cervical neoplasia precursor lesions and their duration as a function of HPV infection status.

Beginning in 1993, we initiated a cohort study involving repeated measurements of HPV infection and cervical cytology in women attending a comprehensive maternal and child health program catering to low income families living in neighborhoods located in the northern sector of the city of São Paulo, Brazil. By restricting ourselves to the detection and follow-up of precursor lesions for which treatment is generally not prescribed in this population, we were able to evaluate prospectively the occurrence of SIL events at regular intervals over time. In particular, we sought to measure the rate of progression, regression and duration of precursor lesions (sojourn time) in cervical cancer according to HPV status.

METHODS AND MATERIALS

Subject recruitment

Using patient rosters two study nurses selected a systematic sample of 4990 women to be approached for interview from daily lists of outpatients in the family medicine, gynecology, and family planning clinics at the "Vila Nova Cachoeirinha" municipal hospital. Women who were potentially eligible were given a detailed overview of the study and invited to participate.

Women were eligible to participate if they: (i) were aged between 18 and 60 years; (ii) were permanent residents of São Paulo (city); (iii) were not currently pregnant and had no intention of becoming pregnant during the next 12 months; (iv) had an intact uterus and no current referral for

hysterectomy; (v) reported no use of vaginal medication in the previous 2 days; and (vi) had not had treatment for cervical disease by electrocoagulation, cryotherapy, or conization in the previous 6 months. In addition to these criteria, women were considered ineligible if they were not interested in complying with all scheduled returns, at least for the subsequent 2 years.

Subjects entered the study only after giving signed informed consent. The study protocol was approved by institutional ethical and research review boards of the participating institutions in Canada and in Brazil. A detailed description of the design and methods of the study has been published previously [*62*]. All participants were seen every 4 months in the first year (0, 4, 8 and 12 months), and twice yearly thereafter. Delays in returning for a given appointment were allowed; the visit sequence was maintained even where subjects returned for their follow-up after the scheduled date with information and specimens collected being assigned to the pending follow-up return to retain the same number of scheduled visits, which precluded the occurrence of missing interval visits. Cervical specimens were taken for Pap cytology and HPV testing at every visit. The study nurses also performed a detailed interview to collect information on sociodemographic factors, reproductive health, sexual activity and smoking at enrollment. Information on sexual activity and reproductive health was also collected at each return visit during the first 12 months and at yearly returns after that.

Cervical cell specimens

An Accelon biosampler (Medscand, Inc.) was used to collect an ecto- and endocervical sample. After the cells were smeared on a glass slide and fixed for cytology, the sampler containing the residual exfoliated cells was immersed in a tube containing Tris-EDTA buffer pH 7.4. Samples were then sent to the Ludwig Institute for Cancer Research in São Paulo for storage and testing. The Pap smears were shipped to Montreal for coding and classification by an expert cytopathologist (AF) who was blinded to previous cytology outcomes and to HPV results for the same and previous samples. The cytopathology reports were based on the 1992 Bethesda system for cytological diagnoses [*97*].

The progression and regression states were classified by lesion severity (LSILs and HSILs), and by further taking into account the different sub-categories within these lesion grades. This was done by stratifying LSILs into lesions that showed predominantly koilocytotic atypia or effects of a productive HPV infection (LSIL/HPV) and those that showed signs of squamous abnormality (LSIL/SQ) equivalent to mild dysplasia or CIN1, as judged by blind review cytology. Similarly, cytological HSIL findings were also separated into those lesions consistent with moderate dysplastic changes (equivalent to CIN2) and those with more severe dysplastic features (equivalent to CIN3).

HPV DNA detection

Cervical specimens were tested for the presence of HPV DNA by a standardized polymerase chain reaction (PCR) protocol [*10;105*]. Typing of the amplified products was performed by hybridization with individual oligonucleotide probes specific for 27 HPV genital types whose nucleotide sequences within the MY09/11 fragment have been published in the literature [*105*]. Amplified products that hybridized with the generic probe but with none of the type-specific probes were further tested by restriction fragment length polymorphism (RFLP) analysis [*12*] to extend the range of identifiable HPV types. To verify the specificity of the hybridizations, we included more than 30 type-specific positive controls in all membranes. In order to check the integrity of the host DNA material extracted from the specimens, assays also included an additional set of primers to amplify the ß-globin gene [*10*]. HPV types were separated into two groups by level of oncogenicity. Oncogenic types included HPVs 16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, 59, and 68; non-oncogenic types included 6/11, 26, 32, 34, 40, 42, 44, 53, 54, 55, 57, 62, 64, 66, 67, 69, 70, 71, 72, 73, 81, 82, 83, 84, CP6108 plus other unknown types. All HPV assays were done on coded specimens with no identification linking specimens from the same woman. Appropriate precautions were taken to reduce the possibility of specimen contamination.

Statistical Analyses

For the analyses reported here, follow-up continued until March 2002. Subjects with an incident cytologically abnormal result (ASCUS, LSIL, or HSIL) were eligible for inclusion in each risk set at the time of their first detected result. Prevalent cases at enrollment were excluded from the risk set at baseline. Time to event in the analyses was measured from the first instance of a non-normal cytological result (date of index visit) to the first detected transition to a more or less severe state (for estimates of progression or regression, respectively), or to the last recorded return visit date for censored subjects. This represents interval-censored data as the exact date of HPV infection and SIL incidence are not known. Subjects who received biopsies were censored at the time of their biopsy if no transition event had occurred prior to the biopsy date. Women who dropped out of the study were censored at their last visit date. The time to a regression event from HSIL or LSIL was defined as the time until the first follow-up visit when a subject presented with a LSIL, ASCUS or normal Pap smear (depending in the regression endpoint of interest), whether or not a worse cytological event was detected during that period. Incident cases of progression or regression were evaluated for the overall group at risk, by age and ethnic group, and stratified by HPV status at the last available visit (either at the same visit in which the index lesion was first detected or the next earliest valid HPV result prior to the index event).

The estimates of incidence rates for a given lesion event included only women at risk of acquiring a lesion. Cumulative probability of remaining in the same lesion stage or progressing to the next was estimated by actuarial analysis using Kaplan-Meier curves [*121*] as a function of the length of follow-up, among women who had no lesion detected at enrollment. The life table method was also used to estimate the proportion of women who remained positive for a precursor lesion detected by cytology according to their HPV status at the index visit when the lesion was first detected [*38*]. 95% confidence intervals (CIs) for the actuarial estimates were calculated directly by using the standard error of the cumulative probability at the end of a particular interval where an event occurred [*197*]. Statistical comparison of lesion sojourn times between HPV positive subjects and HPV negative subjects was performed by log-rank test. Statistical analyses were performed using SPSS® version 11.0 (SPSS, Chicago, IL).

In addition to the latter actuarial estimates of mean duration of preinvasive lesions we used a standard formula based on the epidemiologic tenet that, within a stationary population and in the absence of migration, the prevalence proportion (P) is a function of the incidence rate (I) and of the mean duration (D) of the condition. Therefore, the average duration can be estimated with the general formula D=P/[I x (1-P)], where prevalence was calculated as a weighted average of the point prevalence over time for each lesion grade. This formula holds provided that the point prevalence within each stratum is less than 0.1 [*76*], a condition that was met for all individual lesion grades analyzed in the study. The incidence rate was derived for subjects with normal cytology at enrollment. These estimates were stratified by the cumulative HPV status over the first year of follow-up as per the following hierarchical categories: (i) subjects with no HPV DNA detected at all 4 visits, (ii) those with only non-oncogenic types in any of the visits, (iii) those with any oncogenic HPV types except HPV 16, in any of the visits, and (iv) subjects with HPV 16 at any of the visits.

RESULTS

3589 women initially met the eligibility criteria and were invited to participate. Between November 1993 and March 1997, 2528 women were enrolled into the study, representing a response rate of 70.4%. A further 66 women who were found not to fit the eligibility criteria were excluded after enrollment. Fifty-one (2.1%) women presented with prevalent lesions at enrollment and 7 women had inconclusive cytology results. Among the 2404 remaining women, there were 131 (5.4%) women with incident SIL events, of which 24 were HSIL, and 119 were LSIL. There were also 174 incident reports of ASCUS within the 8 years of follow-up. The prevalence rates for oncogenic and non-oncogenic HPV types were similar across visits and varied between 7.8% to 9.2% for oncogenic types and 6.2% to 7.6% for non-oncogenic types.

The majority of ASCUS and LSIL events detected by cytology regressed to normal (Table IV-1). Actuarial analyses showed that half of these events regressed within 6 months of first detection. The overall mean duration of a cytological state was shorter for ASCUS than for LSIL. The mean duration of incident LSIL events was also shorter when regression was assumed to either ASCUS or a normal result than when the regression definition was restricted to a normal smear, although the difference in estimates was small. In general, the mean duration of an index abnormality regardless of grade was longer for lesions with oncogenic HPV types than for those with non-oncogenic HPVs or negative specimens. However, the differences in regression times were not always statistically significant and the median durations differed little. The mean time to regression for abnormalities with oncogenic HPV types was generally longer than those with non-oncogenic types or HPV negative specimens. There were no appreciable differences in regression rates for the subcategories of LSIL across levels of HPV status.

We also calculated mean duration based on the non-actuarial formula described above stratified by cumulative HPV status over the first year of follow-up. The mean duration for ASCUS smears for women HPV negative at all four visits was 7.9 months. The mean durations increased to 10.5 months for subjects with only non-oncogenic HPV types, 15.4 months for those with any oncogenic types excluding HPV 16, and 13.4 months for those positive for HPV 16 at any visit. The equivalent mean durations were similar for LSILs: 8.9, 10.3, 12.2 and 13.4 months for HPV negative, non-oncogenic, oncogenic and HPV 16 positive subjects, respectively. For HSILs the mean durations were 7.6, 5.7, 15.6 and 57.0 months, as per the aforementioned HPV categories.

With respect to progression to a higher preinvasive lesion grade, we observed the reverse increase in trend with respect to HPV infection status (Table IV-2). That is, subjects with no HPV detected in the index specimen took longer to progress than those with oncogenic HPV. Few HPV 16 positive ASCUS or LSIL specimens progressed to a higher grade. The mean sojourn time for LSILs showing signs of squamous abnormality (LSIL/SQ) equivalent to CIN 1 (time to progression to HSIL or worse) was 89.2 (95%CI 84.2-94.2) and 86.4 months (95%CI 81.9-90.9) for LSIL smears with only koilocytotic atypia (LSIL/HPV). Rates of progression for all groups were low, however, and differences in mean time to progression were not significant.

We estimated the differential rates of progression separately for younger and older women. On average, women 31 years of age and over progressed to HSIL from an incident LSIL smear earlier (mean time to progression=77.9 mos, 95%CI 71.1-84.6) than younger women (88.4 mos, 95%CI 82.6-94.1). Little difference in mean time to progression to HSIL was observed between older and younger women for ASCUS smears (90.4 mos, 95%CI 87.7-93.2 and 87.8 mos, 95%CI 77.9-85.6, respectively). When stratified by HPV status, older women with oncogenic infections

had a higher cumulative risk of progression to HSIL regardless of baseline abnormality (Figure IV-1). Mean time to progression from LSIL to HSIL or worse for women with oncogenic infections was 68.4 mos (95%CI 55.1-81.7) in women aged 31 and over, and 75.6 mos (95%CI 65.3-85.8) in women aged less than 31 years. A similar difference in mean time was observed for progression from ASCUS smears.

We also evaluated the rate of regression and progression for subjects with respect to ethnic origin (data not shown). Due to the small number of non-white subjects whose ethnic background was different than African Brazilian, ethnicity in this study was classified as white (n = 1542) and non-white (n = 856). The regression density rates of incident ASCUS, LSIL and HSIL Pap smears to normal for whites versus non-whites were: 10.8 (95% CI 8.7-13.4) vs. 11.4 (95% CI 8.9-14.5) for ASCUS smears; 10.2 (95% CI 7.8-13.0) vs. 7.8 (95% CI 5.7-10.4) for LSIL; and 3.9 (95% CI 1.6-8.1) vs. 8.6 (95% CI 0.4-42.4) for HSIL per 100 women-months, respectively. As observed for rates of regression, we saw little difference in rates of progression to HSIL or worse between ethnic groups: 0.07 (95% CI 0.02-0.2) vs. 0.06 (95% CI 0.01-0.2) for ASCUS smears; and 0.23 (95% CI 0.1-0.5) vs. 0.18 (95% CI 0.06-0.4) for LSIL per 100 women-months, respectively.

Table IV-3 shows the estimated rates of regression as detected by cytology for subjects with incident HSIL. Referral to colposcopy was based either on the results of local or review cytology reports, or on the basis of the cervicography, which was performed once every two years. Biopsies were done at a colposcopy referral visit if lesional tissue could be visualized. During the follow-up of the women with incident HSIL included in the regression analyses, 23 received a biopsy, which were performed an average of 27.9 months after first detection of HSIL by cytology. As the study progressed, recommendations for biopsy became more aggressive, but some lesions appeared to regress during follow-up before biopsy could be performed. We therefore evaluated rates of regression among women separated into two groups depending on whether the date of biopsy occurred before or after regression was determined by cytology. While the number of lesions with (N=12) or without (N=12) possible biopsy interference was small, those receiving a biopsy before regression retained their lesion longer than those whose biopsy occurred after regression. The mean duration of lesion persistence after receiving the biopsy (calculated by subtracting time to biopsy from time to regression) was 5.6 months.

DISCUSSION

The cumulative evidence to date on the natural history of HPV and cervical neoplasia suggests that a relationship between the likelihood of precursor lesions persisting and progressing to cancer is dependent on the characteristics of HPV infections [*16;107;142;224*]. However, to date, few studies have investigated rates of progression and regression of preinvasive lesions with

respect to risk factors for cervical cancer [*109;152;155;162;187*]. Following a large cohort of women in Brazil at repeated intervals, we found that precursor lesions of the cervix detected by cytology persisted longer and were more likely to progress in women with oncogenic HPV infections. Before we consider the study's implications there are a number of strengths and weaknesses that should be mentioned.

Using aggressive repeated screening by cytology we were able to evaluate the regression and progression of cervical lesions over time on a more systematic fashion than previous studies based on registry data and screening programs, which are based on passive data collection. However, given that our outcome ascertainment was based on cytological analysis, misclassification of lesion outcome history is a potential limitation of the study, even though the cytological assessments were carefully conducted in a reference laboratory following a strict quality control protocol. We opted for an intensive, expert cytologic review every 4 to 6 months of all subjects in the study, and referred all instances of HSIL for colposcopy. This approach reduced to likelihood of unnecessary biopsies, which would have interfered with the natural history of early lesions [*206*]. Nonetheless, the occurrence of false negative tests by cytology may have resulted in overestimation of regression time. Alternatively, false negative Pap smears could have either increased or decreased time to progression depending on whether these occurred at lesion outset or during the sojourn period. In a prospective study of cases of carcinoma *in situ* (CIS) diagnosed by histology, McIndoe et al. [*152*] found 60% of the cases had normal cytology after biopsy with only one case developing invasive carcinoma within four years. We therefore censored subjects at the time of their biopsy, anticipating that the procedure could influence the rates of disease determined by cytology. Misclassification of HPV status was less likely as we employed a highly sensitive PCR-based testing method. We tested for HPV DNA in the same specimen used in to perform Pap cytology with the assumption that the HPV finding in incident lesions is a proxy for prior HPV infection states that led to the lesion.

Studies using both histology and cytology to follow the natural history of cervical neoplasia have observed no effect of superficial sampling by biopsy on the short-term course of dysplasia [*269*]. In a review of 27 studies of CIN, Mitchell et al. [*162*] observed similar probabilities of regression, persistence, and progression based on biopsy evaluation versus cytology. For those subjects identified as having HSIL, we investigated the effect of time of biopsy on the rate of regression to ASCUS or normal. We found that HSILs persisted on average 5.6 months following a biopsy. Although it is more likely for persistent lesions to be biopsied prior to regressing than those of short duration, it was not evident if administration of a biopsy interrupted the natural history of HSIL in this study. While more aggressive standards for biopsy were adopted later in the study by the local colposcopists following our recommendations, we cannot exclude the possibility that the

biopsy procedures performed earlier were as a result of visualization of a more severe lesion at colposcopy that would have taken longer to regress regardless of the biopsy sampling.

We calculated the mean duration or sojourn time of incident cytological lesions using two methods. Estimates of mean duration based on actuarial probability estimates indicated that lesions with oncogenic HPV infections appeared to be more persistent than those with non-oncogenic infections. The average duration estimates based on the non-actuarial formula were similar although mean duration of HSIL was longer for subjects with HPV 16. The prevalence-incidence relation method is appropriate to estimate average duration of incident illnesses, such as cytological abnormalities consistent with ASCUS, LSIL or HSIL, which rarely go over 10% in most clinical settings [*76*]. ASCUS diagnoses were relatively rare in our cohort but overcall of lesion status by cytology may have decreased estimates of the rate of regression while overestimating time to progression. On the other hand, the non-actuarial formula does not account for censored data (i.e., incomplete observations due to non-cleared lesions at study closing date or losses to follow-up). In addition to the small number of observations, it is also possible that our incongruous results showing a shorter mean time to regression for HPV 16 lesions was due to insufficient follow-up to reveal the actual time to regression. In such cases, actuarial estimates of median time to regression, where obtainable by Kaplan-Meier analysis, may provide more appropriate estimates of duration. For our evaluation of time to progression, estimates were restricted to the longest follow-up time regardless of event status. As a result mean times were underestimated where the largest observed analysis time was censored.

It is conceivable that a more intensive cytologic follow-up with shorter screening intervals could have revealed more closely the actual duration of the lesion sojourn times. In this study, between-test intervals were increased to 6 months after the first year in order to make follow-up consistent with that of cytology-based screening programs. As a result, the reduction in follow-up frequency may have increased the observed sojourn time. However, this effect would have been equal across HPV groups as all cytology and HPV evaluations were carried out blindly with respect to previous results for the same subject, making comparisons valid on a relative scale. Furthermore, even though visits were scheduled according to the study design, the interval delays between subsequent visits for participants were varied, ranging from 2.9 to 81.3 months apart.

A few review and meta-analysis studies have attempted to summarize rates of progression and regression along the continuum of cervical dysplastic changes using different criteria for selecting investigations and for combining natural history data. Östor [*187*] observed that the average probabilities of regression were 57% for CIN 1, 43% for CIN 2, and 32% for CIN 3 using histological data. Probabilities of progression to CIS were 11% for CIN 1 and 22% for CIN 2.

Mitchell et al. [*162*] observed probabilities of regression, persistence, and progression to any higher grade lesion of 34%, 41% and 25%, respectively. Regarding the latter progression figure, 10% of the lesions progressed to CIS and 1% to invasive cancer. Melnikow et al. [*155*] calculated weighted average rates of progression to HSIL at 24 months according to baseline cytological abnormality: these were 7.1%, 20.8%, 23.4% for ASCUS, LSIL, and HSIL (persistence). The cumulative proportions of progression to invasive cancer at 24 months were 0.3% for ASCUS, 0.2% for LSIL, and 1.4% for HSIL. Average rates of regression to a normal Pap smear were 68.2% for ASCUS, 47.4% for LSIL, and 35.0% for HSIL. Using mild dysplasia taken as referent, Holowaty et al. [*109*] found RRs of CIS of 8.1 for women with moderate dysplasia and 22.7 for those with severe dysplasia within a 2-year period. The equivalent RRs of invasive cancer were 4.5 and 20.7 for the latter lesion grades, respectively.

Holowaty et al. [*109*] observed an increase in rate of progression during the first two years following a positive Pap smear relative to the rate of progression in subsequent years. This was interpreted to be a result of undercalling of the original smear. The recommendation was therefore that the initial repeat smears should be done within 6 months of a first smear rather than one year. We concur with this recommendation concerning the appropriate interval between repeat Pap smears for this and for additional reasons. Specifically, half of all mild to moderate dysplasias in our study regressed to normal within 6 months. Repeat screening with a shorter delay would therefore detect a substantial proportion of lesions that would regress spontaneously. This would be a concern particularly in health care settings that recommend colposcopy referral and biopsy after two LSILs. However, given a mean time to progression observed for LSILs of 85.7 months, most repeat cytology screenings before 1 year would not detect whether a lesion is more likely to progress.

In this study, presence of oncogenic HPV in LSIL smears seemed to provide some discrimination of whether a lesion will progress faster. Holowaty et al. [*109*] examined the influence of parity, age, OC use and number of dysplastic smears but found these did not alter the relative risk of progression for LSILs. While the mean time to progression was not statistically significantly different between age groups, we found that the effect of harboring oncogenic HPV types was stronger in older women over the age of 31. This finding would support the recommendation for aggressive screening of older women at risk of cervical cancer [*136;157*]. We also observed a shorter time to progression for specimens of ASCUS with non-oncogenic and oncogenic HPV relative to HPV negative specimens. This provides further evidence of the importance of following up women with ASCUS smears that have been found to be HPV positive for oncogenic types [*233*].

In conclusion, using screening tests for oncogenic HPVs may help in identifying those lesions that progress faster to more advanced stages. HPV testing of abnormal Pap smears may therefore help to identify women who might benefit from aggressive follow-up and identify candidates for chemopreventive treatment or therapy. The ability to identify subjects whose lesions will take longer to progress could also be cost-saving by reducing clinical follow-up and morbidity resulting from potentially unnecessary invasive therapeutic and diagnostic procedures.

Table IV-1. Actuarial estimates of time to regression of first incident cervical abnormality events stratified by HPV status.

Baseline lesion and regression event	HPV status in baseline sample*	No. events /Total†	Women months of follow-up	Regression density rate (per 100 women–months) (95%CI)	Median time to regression‡ (95%CI) in months	Mean time to regression‡ (95%CI) in months	Proportion§ (%) remaining with abnormality at (SE)			
							6 months	12 months	18 months	24 months
ASCUS to normal										
	Overall	147/174	1325.9	11.2 (9.5-13.1)	6.1 (5.9-6.2)	9.2 (7.4-11.0)	57.1 (3.9)	19.3 (3.3)	6.7 (2.1)	2.7 (1.4)
	Negative	105/117	827.5	12.7 (10.4-15.3)	6.1 (6.0-6.2)	7.6 (6.9-8.4)	58.0 (4.7)	14.8 (3.5)	4.2 (2.0)	0 (0.0)
	Non-oncogenic	19/24	151.3	12.6 (7.8-19.3)	5.8 (5.3-6.4)	7.7 (5.2-10.2)	45.5 (10.6)	15.2 (8.6)	7.6 (6.9)	7.6 (6.9)
	Oncogenic¶	17/23	270.1	6.3 (3.8-9.5)	6.2 (5.8-6.7)	17.0 (6.6-27.4)‡‡	70.7 (10.1)	43.5 (11.4)	16.3 (8.6)	10.9 (7.2)
	HPV 16	6/8	71.2	8.4 (3.4-17.5)	5.8 (5.4-6.2)	10.5 (5.6-15.5)	33.3 (17.2)	33.3 (17.2)	16.7(14.6)	16.7(14.6)
LSIL (any) to normal										
	Overall	102/119	1137.9	9.0 (7.4-10.8)	6.0 (5.8-6.2)	11.5 (8.5-14.1)	52.4 (4.7)	26.9 (4.3)	12.6 (3.4)	8.8 (3.0)
	Negative	37/44	346.9	10.7 (7.6-14.5)	6.1 (5.8-6.3)	8.7 (7.4-10.1)	57.7 (7.6)	23.6 (6.8)	6.7 (4.5)	0 (0.0)
	Non-oncogenic	27/28	208.7	12.9 (8.7-18.5)	5.3 (3.7-7.0)	7.8 (5.3-10.2)	32.1 (8.8)	17.0 (7.2)	8.5 (5.6)	4.3 (4.1)
	Oncogenic¶	30/38	503.5	6.0 (4.1-8.4)	6.3 (0.0-13.7)	17.3 (10.6-23.9)	61.6 (8.1)	38.2 (8.2)	24.9 (7.6)	21.3 (7.3)
	HPV 16	7/7	67.2	10.4 (4.6-20.6)	11.8 (0.0-27.1)	9.6 (6.3-12.9)	57.1 (18.7)	28.6 (17.1)	0 (0.0)	-
LSIL (any) to ASCUS or normal										
	Overall	104/119	1057.8	9.8 (8.1-11.9)	6.0 (5.9-6.1)	10.5 (8.1-12.9)	48.9 (4.7)	22.3 (4.0)	10.0 (3.1)	7.5 (2.8)
	Negative	37/44	304.5	12.2 (8.7-16.6)	6.0 (5.9-6.1)	7.6 (6.4-8.7)	50.6 (7.7)	16.0 (5.9)	4.0 (3.8)	-
	Non-oncogenic	27/28	208.7	12.9 (8.7-18.5)	5.3 (3.7-7.0)	7.8 (5.3-10.2)	32.1 (8.8)	17.0 (7.2)	8.5 (5.6)	4.3 (4.1)
	Oncogenic¶	32/38	471.5	6.8 (4.7-9.5)	6.1 (4.7-7.4)	14.9 (8.9-20.8)	58.9 (8.1)	32.4 (7.9)	19.4 (6.9)	16.2 (6.5)
	HPV 16	7/7	61.4	11.4 (5.0-22.6)	6.0 (5.4-6.6)	8.8 (5.5-12.1)	57.1 (18.7)	28.6 (17.1)	0 (0.0)	-
LSIL/HPV** to ASCUS or normal										
	Overall	47/56	547.5	8.6 (6.4-11.3)	6.1 (5.7.6.6)	12.1 (8.2-15.9)	54.1 (6.8)	25.5 (6.1)	15.3 (5.4)	12.3 (5.1)
	Negative	15/20	136.0	11.0 (6.4-17.8)	6.0 (5.8-6.2)	7.8 (5.9-9.7)	52.6 (11.5)	15.5 (8.9)	15.5 (8.9)	-
	Non-oncogenic	14/15	137.2	10.2 (5.8-16.7)	6.5 (3.1-9.8)	9.7(5.3-14.2)	53.3 (12.9)	32.0(12.3)	16.0 (10.1)	8.0 (7.6)
	Oncogenic¶	14/17	238.0	5.9 (3.4-9.6)	7.1 (0.0-17.9)	16.5 (7.2-25.9)	57.6 (12.2)	30.5 (11.8)	22.9 (11.0)	22.9 (11.0)
	HPV 16	3/3	30.4	9.9 (2.5-26.9)	11.8 (2.1-21.4)	10.1 (5.8-14.4)	66.7 (27.2)	33.3 (27.2)	-	-
LSIL/SQ** to ASCUS or normal										
	Overall	63/72	598.5	6.9 (4.3-10.5)	6.0 (5.9-6.1)	9.9 (6.9-12.8)	49.4 (6.0)	19.1 (4.9)	8.7 (3.6)	5.2 (2.9)
	Negative	25/29	198.5	12.6 (8.4-18.3)	6.0(5.9-6.1)	7.8 (6.2-8.6)	50.0 (9.5)	16.7 (7.5)	0 (0.0)	-

	Non-oncogenic	15/15	86.6	17.3 (10.1-27.9)	5.8 (5.2-6.3)	5.7 (4.8-6.7)‡‡	26.7 (11.4)	0 (0.0)	-	-
	Oncogenic¶	19/23	276.5	6.9 (4.3-10.5)	6.1 (5.7-6.4)	15.3 (7.2-23.3)	62.8 (10.4)	33.8 (10.3)	24.2 (9.4)	14.5 (7.7)
	HPV 16	4/4	31.1	12.9 (4.1-31.1)	5.8 (3.6-8.0)	7.8 (2.7-12.9)	50.0 (25.0)	25.0 (21.7)	0 (0.0)	-
HSIL/CIN2** to ASCUS or normal										
	Overall	13/16	152.1	8.5 (4.8-14.2)	6.1 (6.0-6.1)	11.6 (5.7-17.5)	66.7 (12.2)	14.8 (9.6)	14.8 (9.6)	14.8 (9.6)
	Negative	6/6	41.8	14.3 (5.8-29.8)	6.0 (5.7-6.3)	7.0 (5.0-8.9)	66.7 (19.3)	0 (0.0)	-	-
	Non-oncogenic	2/2	17.9	11.2 (1.9-37.0)	6.1 (-)	8.9 (3.3-14.6)	100 (0.0)	0 (0.0)	-	-
	Oncogenic††	5/8	92.4	5.4 (2.0-12.0)	6.0 (0.0-13.2)	17.1 (4.1-30.1)	57.1 (18.7)	38.1 (19.9)	38.1(19.9)	38.1(19.9)

* Baseline sample defined as the first detected event of the stated cytological abnormality.

† Number lesions regressed / total number of index lesions. HPV stratum specific number of samples may not add up to the overall number if valid HPV results were unavailable.

‡ Estimates from actuarial analysis using the Kaplan-Meier technique. Prevalent cases at enrollment were excluded.

§ Proportion remaining with lesion at end of interval period (6, 12, 18, 24 months) derived by life table analysis corresponding to survival at time t [S(t)].

¶ Oncogenic HPV types excluding HPV 16 in baseline smear.

** LSIL/HPV: LSIL with koilocytotic atypia induced by a productive HPV infection; LSIL/SQ: LSIL showing squamous effects equivalent to cervical intraepithelial neoplasia grade 1; HSIL/CIN2: HSIL with moderate dysplasia equivalent to cervical intraepithelial neoplasia grade 2.

†† No HPV 16 positive specimens were identified.

‡‡ Significant for log rank test ($p<0.05$) using HPV negative as comparison group.

Table IV-2. Actuarial estimates of time to progression of first incident cervical abnormality events stratified by HPV status.

Baseline lesion and progression event	HPV status in baseline sample*	No. events /Total†	Women months of follow-up	Progression density rate (per 100 women–months) (95%CI)	Mean time to progression‡ (95%CI) in months	Proportion§ (%) progressing to worse abnormality at (SE) 6 months	12 months	18 months
ASCUS to LSIL								
	Overall	15/174	6879.5	0.2 (0.1-0.4)	84.4 (80.3-88.4)	3.7 (1.5)	7.2 (2.1)	7.9 (2.2)
	Negative	6/117	4992.2	0.12 (0.05-0.3)	88.0 (84.3-91.7)	0.9 (0.9)	3.9 (1.9)	3.9 (1.9)
	Non-oncogenic	5/24	756.5	0.7 (0.2-1.5)	67.9 (51.9-83.9)‡‡	9.8 (6.6)	21.5 (9.6)	21.5 (9.6)
	Oncogenic¶	4/23	954.1	0.4 (0.1-1.0)	71.9 (58.9-84.9)‡‡	13.6 (7.3)	13.6 (7.3)	18.4 (8.3)
	HPV 16	0/8	170.8	0.0 (0.0-1.8)	-	-	-	-
ASCUS to LSIL or worse								
	Overall	18/174	6867.3	0.3 (0.2-0.4)	82.7 (78.4-87.1)	3.7 (1.5)	7.9 (2.2)	8.6 (2.3)
	Negative	6/117	4992.2	0.12 (0.05-0.3)	88.0 (84.3-91.7)	0.9 (0.9)	3.9 (1.9)	3.9 (1.9)
	Non-oncogenic	5/24	756.5	0.7 (0.2-1.5)	67.9 (51.9-83.9)‡‡	9.8 (6.6)	21.5 (9.6)	21.5 (9.6)
	Oncogenic¶	6/23	954.1	0.6 (0.3-1.3)	66.2 (52.2-80.1)‡‡	13.6 (7.3)	13.6 (7.3)	18.4 (8.3)
	HPV 16	1/8	158.7	0.6 (0.03-3.1)	42.7 (29.3-56.2)	0 (0.0)	18.2 (16.5)	18.2 (16.5)
LSIL (any) to any HSIL or worse								
	Overall	11/119	5365.5	0.2 (0.1-0.4)	85.7 (80.8-90.6)	1.7 (1.2)	3.6 (1.8)	6.5 (2.4)
	Negative	2/44	1676.3	0.12 (0.02-0.4)	83.5 (78.0-89.1)	0 (0.0)	0 (0.0)	2.8 (2.8)
	Non-oncogenic	1/28	1563.8	0.06 (0.0-0.3)	91.3 (85.1-97.4)‡‡	0 (0.0)	3.6 (3.6)	3.6 (3.6)
	Oncogenic¶	8/38	1729.5	0.4 (0.2-0.9)	70.8 (60.9-80.6)	5.4 (3.7)	8.3 (4.6)	14.3 (5.9)
	HPV 16	0/7	378.2	0.0 (0.0-0.4)	-	-	-	-
LSIL/SQ to HSIL/CIN3**								
	Overall††	4/72	3482.6	0.11 (0.04-0.3)	89.2 (84.2-94.2)	1.0 (1.0)	1.4 (1.4)	3.0 (2.1)
LSIL/HPV to HSIL/CIN3**								
	Overall††	2/56	2266.5	0.09 (0.01-0.3)	86.4 (81.9-90.9)	1.8 (1.8)	1.8 (1.8)	1.8 (1.8)

* Baseline sample defined as the first detected event of the stated cytological abnormality.

† Number lesions progressed / total number of index lesions. HPV stratum specific number of samples may not add up to the overall number if valid HPV results were unavailable.

‡ Estimates from actuarial analysis using the Kaplan-Meier technique. Prevalent cases at enrollment were excluded. Median time to progression was not estimated as less than 50% of the index lesions progressed. Mean times were restricted to the longest follow-up time regardless of event status and may be underestimated where the largest observed analysis time were censored.

§ Proportion progressing at end of interval period (6, 12, 18 months) derived by life table analysis corresponding to one minus survival at time t [1-S(t)].

¶ Oncogenic HPV types excluding HPV 16 in baseline smear.

** LSIL/HPV: LSIL with koilocytotic atypia induced by a productive HPV infection; LSIL/SQ:LSIL showing squamous effects equivalent to cervical intraepithelial neoplasia grade 1; HSIL/CIN3: HSIL with severa dysplasia equivalent to cervical intraepithelial neoplasia grade 3.

†† All baseline (index) LSIL events were oncogenic HPV positive.

‡‡ Significant for log rank test ($p<0.05$) using HPV negative as comparison group.

Table IV-3. Actuarial estimates of time to regression of incident HSIL events to an ASCUS or normal state detected by cytology stratified by HPV and biopsy status.

Date biopsy performed	HPV status in index sample*	No. events /Total†	Women months of follow-up	Regression density rate (per 100 women–months) (95%CI)	Median time to regression‡ (95%CI) in months	Mean time to regression‡ (95%CI) in months	Proportion§ (%) remaining with lesion (SE) at 6 months	1 year	18 months
After date of regression									
	Overall	11/12	99.4	11.1 (5.9-19.2)	6.1 (6.0-6.1)	8.3 (5.8-10.8)	66.7 (13.6)	16.7 (10.8)	5.6 (7.4)
	Negative	5/5	35.9	13.9 (5.1-30.9)	6.1 (6.0-6.1)	7.2 (4.9-9.4)	80.0 (17.9)	0 (0.0)	-
	Non-oncogenic	2/2	17.9	11.2 (1.9-37.0)	6.05 (-)	8.9 (3.3-14.6)	100	0. (0.0)	0 (0.0)
	Oncogenic¶	4/4	41.5	9.6 (3.1-23.2)	6.0 (0.0-13.8)	10.4 (3.8-17.1)	50.0 (25.0)	50.0 (25.0)	16.7 (21.0)
	HPV 16	1/1	4.1	24.2 (1.2-119.1)	-	-	-	-	-
Before date of regression									
	Overall	8/12	159.2	5.0 (2.3-9.5)	11.9 (11.8-12.1)	16.9 (7.9-25.9)	81.0 (12.1)	34.7 (16.0)	23.1 (14.3)
	Negative	2/2	17.7	11.3 (1.9-37.3)	5.8 (-)	8.9 (2.8-15.0)	50.0 (35.4)	0.0 (0.0)	-
	Non-oncogenic	0/0	0.0	-	-	-	-	-	-
	Oncogenic¶	4/7	118.5	3.4 (1.1-8.1)	11.9 (11.7-12.2)	23.2(9.4-37.0)	83.3 (15.2)	41.7 (22.2)	41.7 (22.2)
	HPV 16	2/3	23.0	8.7 (1.5-28.7)	6.7 (-)	9.4 (4.1-14.7)	1 (0.0)	50.0 (35.4)	0 (0)
Irrespective of biopsy status									
	Overall	19/24	258.7	7.4 (4.6-11.3)	6.7 (0.0-15.1)	12.8 (7.6-18.0)	73.3 (9.3)	24.4 (9.5)	13.6 (7.8)
	Negative	7/7	53.6	13.1 (5.7-25.8)	6.1 (6.0-6.1)	7.7 (5.5-9.8)	71.4 (17.1)	0 (0.0)	-
	Non-oncogenic	2/2	17.9	11.2 (1.9-37.0)	6.1 (-)	8.9 (3.3-14.6)	100	0 (0.0)	0 (0.0)
	Oncogenic¶	7/11	160.0	4.4 (1.9-8.7)	11.9 (7.0-16.9)	19.4 (9.0-29.9)	70.0 (14.5)	46.7 (16.6)	33.3 (16.4)
	HPV 16	3/4	27.1	11.1 (2.8-30.1)	6.7 (2.9-10.6)	8.1 (3.7-12.5)	71.4 (17.1)	0 (0.0)	-

* Index sample defined as the first detected high-grade squamous intra-epithelial lesion (HSIL) event.

† Number HSIL regressed / total number of index HSIL. HPV stratum specific number of samples may not add up to the overall number if valid HPV results were unavailable.

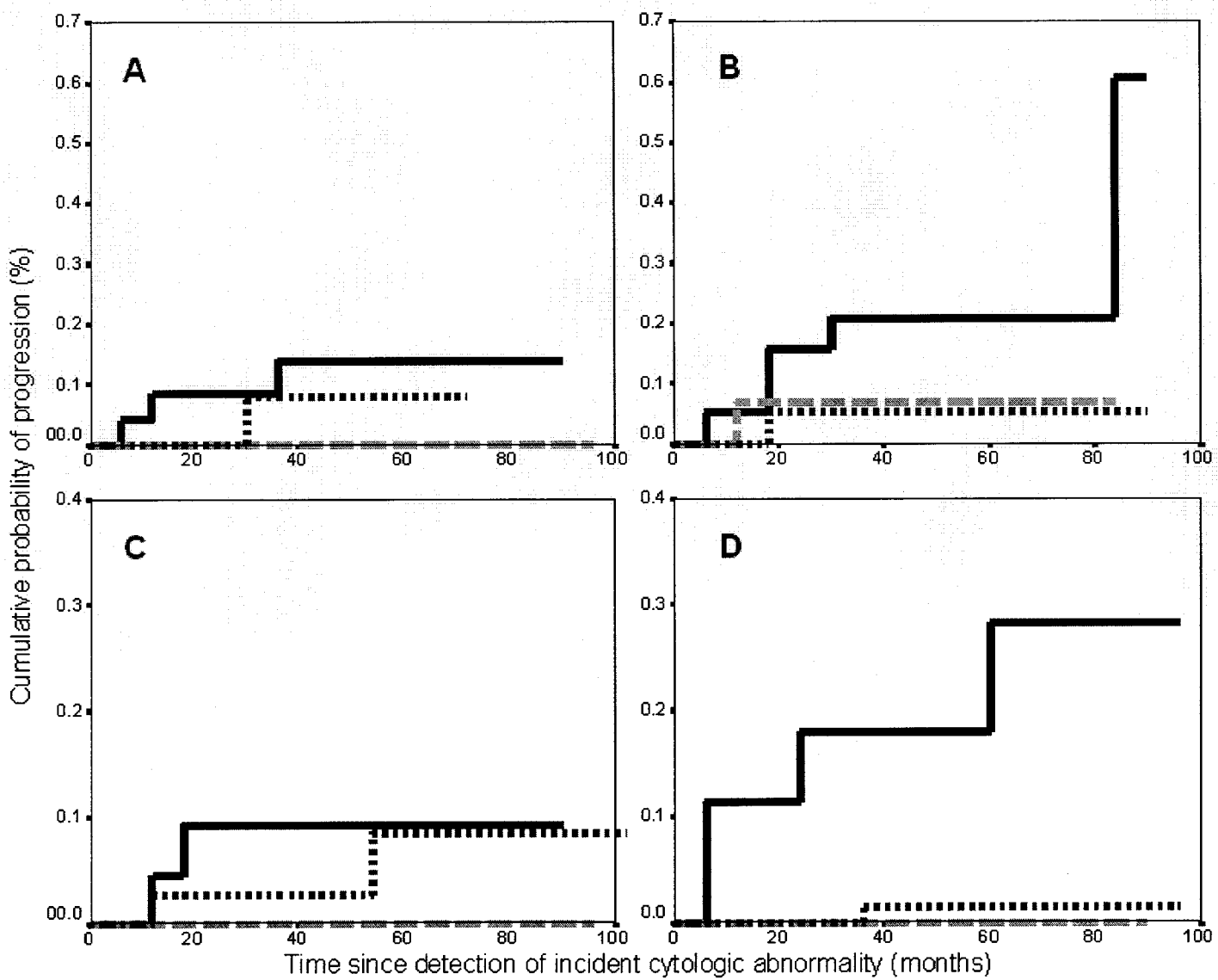
‡ Estimates from actuarial analysis using the Kaplan-Meier technique.

§ Proportion remaining with lesion at end of interval period (6, 12, 18 months) derived by life table analysis corresponding to survival at time t [S(t)]. Prevalent cases at enrollment were excluded.

¶ Oncogenic HPV types excluding HPV 16 in index smear.

Figure IV-1. Kaplan-Meier graphs for time to progression to HSIL (in months) of incident cervical abnormality detected by cytology according to HPV oncogenicity at baseline and age group

Graph A, Progression of LSIL to HSIL in women 16-30 years; Graph B, Progression of LSIL to HSIL in women 31-65 years; Graph C, Progression of ASCUS to HSIL in women 16-30 years; Graph D, Progression of ASCUS to HSIL in women 31-65 years. HPV status: Negative (dotted line), Non-oncogenic (dashed line), Oncogenic (solid).


A
B
C
D
Cumulative probability of progression (%)
Time since detection of incident cytologic abnormality (months)
0.7
0.6
0.5
0.4
0.3
0.2
0.1
00.0
0.0
0
20
40
60
80
100

7. IMPLICATIONS AND FUTURE RESEARCH

7.1. Public health impact

The results from this project could have the following public health implications:

Pap test screening guidelines vary widely across countries. The WHO recommends that every woman should have a Pap test between ages 35 and 40 years. If the resources are available, frequency of screening should be increased to every 10 years starting at 35 years of age and then every 5 years for women aged 40-55. In Canada this screening interval is considerably shorter (once every 1-2 years), with more aggressive follow-up if an abnormal test is detected. Given that the introduction of screening in developed countries has reduced cancer incidence 80% [*159*], knowledge about the etiology of the disease and its progression can help policy makers and public health practitioners develop more efficient screening programs. This could occur: i) by allowing for appropriate screening intervals; ii) by targeting high-risk groups with more intensive screening and follow-up. One scenario that would be supported by the results in this thesis would be the introduction of different screening intervals for women with oncogenic and non-oncogenic HPV infections given the differences in lesion progression rates seen for these HPV types.

Furthermore, recognition that infection with certain types of HPV [*110*] are the central cause of cervical neoplasia has created new research fronts in primary prevention (i.e. vaccination against HPV infection) [*222*] and in secondary prevention of this disease (i.e. HPV testing as a screening tool) [*43*]. Therefore, understanding the remote causes of persistent infection with the clinically relevant HPV types and the dynamics of lesion development is an important first step in the direction of implementing more effective public health programs aiming at risk reduction. Evidence on the time-dependent relationship between acquired HPV infections and the risk of preinvasive lesions presented in this thesis could be used in such evaluations of the effectiveness of screening programs using HPV testing and in defining intermediate outcomes to be measured in clinical trials of vaccines.

7.2. Implications for future research

7.2.1. Measures of viral integration

Some recent studies have proposed using viral load quantification methods to evaluate integration of HPV into the human genome as a marker of progression of cervical neoplasia [*126*]. The procedures suggested have targeted two HPV genes in particular to try and measure this process (E6 andE2). While E6 is similarly a good target for quantification as is L1, E2 can be disrupted as the lesion progresses, and could be used as a marker of integration in conjunction

with E6 measurements. Integration can occur at random sites, the disruption in the genome usually occurs at the E1/E2 sites involved in viral replication and transcription. Building on this characteristic of the HPV virus, some studies have suggested trying to correlate cytology results with physical status defined by the ratios E2/E6 and E2/integratedE6 [*181;196*].

There are some additional characteristics to consider when comparing methods measuring viral burden. First, the specimens are obtained from cervical smears, which may contain very few cells representative of a high-grade lesion amidst a large number of normal or even productively infected cells, in which episomes (i.e., extra chromosomal, noon-integrated forms of the virus) are expected to predominate over integrated viral genomes, and replicate actively. Moreover, at least in the case of ASCUS and CIN, misclassification of lesion outcomes can occur, additionally complicating the interpretation. Quantification based on the L1 gene, however, will likely give a more reliable measure of viral load, since this gene is preserved even after the virus integrates into the host genome, with few exceptions.

7.2.2. Misclassification of HPV

The observed reductions in association with time since HPV exposure assessment may be reflective of misclassification with respect to exposure, i.e., HPV status, as HPV infection is, for most part, transient. It is not as if the effect of HPV exposure at baseline is decaying overtime. Ignoring HPV status in the intervening time leads to misclassification hence dilution of effect as "recent" HPV status is more relevant for the accrual of viral exposure information, even though it may be less biologically pertinent with respect to the occurrence of SIL.

Likewise, clinical trials of HPV vaccines may fail to demonstrate an effect from vaccination if HPV testing is done without discrimination for type, with using simplistic measures of viral burden, or at the point in time where the virus initiates carcinogenesis, i.e. at integration, using probes for non-conserved regions such as E2.

7.2.3. HPV type persistence and co-factors

Inflammation due to co-infections and other factors may also play a role in the promotion of persistent HPV infections and cervical carcinogenesis through inhibition of cell-mediated immunity [*30;90*]. Therefore, future studies could evaluate if this translates into differential rates of disease development at incidence of HPV infection between subjects with single or multiple STDs.

In addition, other factors such as OC use and presence of genetic polymorphisms in tumor suppressor genes [*147*] could be investigated as effect modifiers of the rate of progression or regression of precursor lesions of the cervix.

7.2.4. Repeat and adjunct testing in cervical cancer screening

In addition to new technologies, alternatives to traditional screening strategies involving repeated testing by cytology have also been proposed to improve the test performance [*70*]. In a longitudinal study using histology to confirm the results from repeat cytology, Mitchell et al., [*161*] found no improvement in detection of HSIL by early repeat testing of Pap smears that initially lacked endocervical cells. The question of whether HPV testing should be adopted in cervical cancer screening has also been raised in several consensus and conference meetings, where researchers have concluded that there is enough justification to evaluate HPV testing as an adjunct to Pap smear screening.

For adjunct diagnostic techniques to be effective screening tools, there must be a demonstration of an improvement in sensitivity and specificity above what would occur by chance [*70*]. That is, it is inevitable that complementary tests used in tandem will improve the sensitivity of a screening program. Therefore, for such an approach to be effective, there should be a demonstration that the specificity is improved as well, and that both sensitivity and specificity of the screening test combination are better than either test augmented by a random adjunct test. Screening with a combination of cervicography and cytology in a randomized control trial reduced the incidence of new lesions by 30% one year later compared to what was observed in the cytology arm only [*7*]. The result was not significant, however, indicating perhaps that cervicography may be detecting more transient lesions than cytology.

8. CONCLUSION

In Summary, the following conclusions can be made from the results of the above manuscripts:

- The incidence rates for SIL and HSIL are higher in subjects harboring a persistent oncogenic HPV infection measured by type-specific PCR.

- Relative to women consistently negative for HPV, the risk of risk of SIL, HSIL and persistent SIL was substantially higher in women with persistent HPV 16 or 18 infections.

- Women with high viral load are at higher risk of incident SIL compared to those with less than one viral copy per cell in cervical specimens.

- The predictive effect of a viral load 4 to 6 months earlier decreases in subjects also presenting with an abnormal (ASCUS) cytology at the same visit.

- The magnitude of RRs of incident SIL decreases as the time interval between a positive HPV result and cervical lesion incidence increases.

- Relative risks of incident HSIL are higher for older women. Women 31 years of age and over also progress to HSIL from an incident LSIL smear earlier than younger women.

- The observed time to progression of precursor lesions is shorter for women with oncogenic HPV infections compared to those with no HPV infection.

- Although the majority of LSIL and HSIL lesions regress with time, the mean duration of lesions is longer in women with oncogenic HPV and HPV 16 infections.

9. BIBLIOGRAPHY

1. Ahdieh, L; Klein, RS; Burk, R; Cu-Uvin, S; Schuman, P; Duerr, A; Safaeian, M; Astemborski, J; Daniel, R; Shah, K. Prevalence, incidence, and type-specific persistence of human papillomavirus in human immunodeficiency virus (HIV)-positive and HIV- negative women. J Infect Dis 2001;184(6):682-90.

2. ALTS Group. Human papillomavirus testing for triage of women with cytologic evidence of low-grade squamous intraepithelial lesions: baseline data from a randomized trial. The Atypical Squamous Cells of Undetermined Significance/Low-Grade Squamous Intraepithelial Lesions Triage Study (ALTS) Group. J Natl Cancer Inst 2000;92(5):397-402.

3. Andersen, PK; Gill, RD. Cox's Regression Model for Counting Processes: A Large Sample Study. Annals of Statistics 1982;10(4):1100-20.

4. Apple, RJ; Becker, TM; Wheeler, CM; Erlich, HA. Comparison of human leukocyte antigen DR-DQ disease associations found with cervical dysplasia and invasive cervical carcinoma. J Natl Cancer Inst 1995;87(6):427-36.

5. Astori, G; Arzese, A; Pipan, C; de Villiers, EM; Botta, GA. Characterization of a putative new HPV genomic sequence from a cervical lesion using L1 consensus primers and restriction fragment length polymorphism. Virus Res 1997;50(1):57-63.

6. Austin, RM; Ramzy, I. Increased detection of epithelial cell abnormalities by liquid-based gynecologic cytology preparations. A review of accumulated data. Acta Cytol 1998;42(1):178-84.

7. Autier, P; Coibion, M; De Sutter, P; Wayemberg, M. Cytology alone versus cytology and cervicography for cervical and cervicography for cervical cancer screening: a randomized study. European Society for Oncological Research. Obstet Gynecol 1999;93(3):353-8.

8. Barron, BA; Richart, RM. A statistical model of the natural history of cervical carcinoma based on a prospective study of 557 cases. J Natl Cancer Inst 1968;41(6):1343-53.

9. Bauer, HM; Hildesheim, A; Schiffman, MH; Glass, AG; Rush, BB; Scott, DR; Cadell, DM; Kurman, RJ; Manos, MM. Determinants of genital human papillomavirus infection in low-risk women in Portland, Oregon. Sex Transm Dis 1993;20(5):274-8.

10. Bauer, HM; Ting, Y; Greer, CE; Chambers, JC; Tashiro, CJ; Chimera, J; Reingold, A; Manos, MM. Genital human papillomavirus infection in female university students as determined by a PCR-based method. JAMA 1991;265(4):472-7.

11. Becker, TM; Wheeler, CM; Mcgough, NS; Parmenter, CA; Jordan, SW; Stidley, CA; Mcpherson, RS; Dorin, MH. Sexually transmitted diseases and other risk factors for cervical dysplasia among Southwestern Hispanic and non-Hispanic White women. JAMA 1994;271:1181-8.

12. Bernard, HU; Chan, SY; Manos, MM; Ong, CK; Villa, LL; Delius, H; Peyton, CL; Bauer, HM; Wheeler, CM. Identification and assessment of known and novel human papillomaviruses by polymerase chain reaction amplification, restriction fragment length polymorphisms, nucleotide sequence, and phylogenetic algorithms. J Infect Dis 1994;170(5):1077-85.

13. Borresen, AL; Helland, A; Nesland, J; Holm, R; Trope, C; Kaern, J. Papillomaviruses, p53, and cervical cancer. Lancet 1992;339(8805):1350-1.

14. Bos, AB; van Ballegooijen, M; van Oortmarssen, GJ; van Marle, ME; Habbema, JD; Lynge, E. Non-progression of cervical intraepithelial neoplasia estimated from population-screening data. Br J Cancer 1997;75(1):124-30.

15. Bosch, FX; de Sanjose, S. Human papillomavirus in cervical cancer. Curr Oncol Rep 2002;4(2):175-83.

16. Bosch, FX; Lorincz, A; Munoz, N; Meijer, CJ; Shah, KV. The causal relation between human papillomavirus and cervical cancer. J Clin Pathol 2002;55(4):244-65.

17. Bosch, FX; Manos, MM; Munoz, N; Sherman, M; Jansen, AM; Peto, J; Schiffman, MH; Moreno, V; Kurman, R; Shah, KV. Prevalence of human papillomavirus in cervical cancer: a worldwide perspective. International biological study on cervical cancer (IBSCC) Study Group. J Natl Cancer Inst 1995;87(11):796-802.

18. Bosch, FX; Munoz, N; Shah, KV; Meheus, A. Second International Workshop on the Epidemiology of Cervical Cancer and Human Papillomavirus: Meeting Held in Brussels, Belgium, November 25-28, 1991. Int J Cancer 1992;52:171-3.

19. Breslow, NE. Analysis of survival data under the proportional hazards model. International Statistical Review 1975;43:45-58.

20. Brinton, LA. Epidemiology of cervical cancer--overview. IARC Sci Publ 1992;119:3-23.

21. Brinton, LA; Fraumeni, JF Jr. Epidemiology of uterine cervical cancer. J Chronic Dis 1986;39(12):1051-65.

22. Brinton, LA; Hamman, RF; Huggins, GR; Lehman, HF; Levine, RS; Mallin, K; Fraumeni, JF Jr. Sexual and reproductive risk factors for invasive squamous cell cervical cancer. J Natl Cancer Inst 1987;79(1):23-30.

23. Brinton, LA; Reeves, WC; Brenes, MM; Herrero, R; Britton, RC; Gaitan, E; Tenorio, F; Garcia, M; Rawls, WE. Parity as a risk factor for cervical cancer. Am J Epidemiol 1989;130:486-96.

24. Brinton, LA; Reeves, WC; Brenes, MM; Herrero, R; Gaitan, E; Tenorio, F; de Britton, RC; Garcia, M; Rawls, WE. The male factor in the etiology of cervical cancer among sexually monogamous women. Int J Cancer 1989;44(2):199-203.

25. Brisson, J; Morin, C; Fortier, M; Roy, M; Bouchard, C; Leclerc, J; Christen, A; Guimont, C; Penault, F; Meisels, A. Risk factors for cervical intraepithelial neoplasia: differences between low- and high-grade lesions. Am J Epidemiol 1994;140(8):700-10.

26. Burger, MP; Hollema, H; Pieters, WJ; Schroder, FP; Quint, WG. Epidemiological evidence of cervical intraepithelial neoplasia without the presence of human papillomavirus. Br J Cancer 1996;73(6):831-6.

27. Burk, RD; Ho, GY; Beardsley, L; Lempa, M; Peters, M; Bierman, R. Sexual behavior and partner characteristics are the predominant risk factors for genital human papillomavirus infection in young women. J Infect Dis 1996;174(4):679-89.

28. Burmer, GC; Parker, JD; Bates, J; East, K; Kulander, BG. Comparative analysis of human

papillomavirus detection by polymerase chain reaction and Virapap/Viratype kits. Am J Clin Pathol 1990;94(5):554-60.

29. Caballero, OL; Villa, LL; Simpson, AJ. Low stringency-PCR (LS-PCR) allows entirely internally standardized DNA quantitation. Nucleic Acids Res 1995;23(1):192-3.

30. Castle, PE; Hillier, SL; Rabe, LK; Hildesheim, A; Herrero, R; Bratti, MC; Sherman, ME; Burk, RD; Rodriguez, AC; Alfaro, M; Hutchinson, ML; Morales, J; Schiffman, M. An association of cervical inflammation with high-grade cervical neoplasia in women infected with oncogenic human papillomavirus (HPV). Cancer Epidemiol Biomarkers Prev 2001;10(10):1021-7.

31. Castle, PE; Shields, T; Kirnbauer, R; Manos, MM; Burk, RD; Glass, AG; Scott, DR; Sherman, ME; Schiffman, M. Sexual behavior, human papillomavirus type 16 (HPV 16) infection, and HPV 16 seropositivity. Sex Transm Dis 2002;29(3):182-7.

32. CDC. Results from the National Breast and Cervical Cancer Early Detection Program, October 31, 1991-September 30, 1993. MMWR Morb Mortal Wkly Rep 1994;43(29):530-4.

33. Chang, S-H; Wang, M-C. Conditional regression analysis for recurrence time data. JASA 1999;94(448):1221-30.

34. Chaouki, N; Bosch, FX; Munoz, N; Meijer, CJ; El Gueddari, B; El Ghazi, A; Deacon, J; Castellsague, X; Walboomers, JM. The viral origin of cervical cancer in Rabat, Morocco. Int J Cancer 1998;75(4):546-54.

35. Clavel, C; Masure, M; Bory, JP; Putaud, I; Mangeonjean, C; Lorenzato, M; Nazeyrollas, P; Gabriel, R; Quereux, C; Birembaut, P. Human papillomavirus testing in primary screening for the detection of high-grade cervical lesions: a study of 7932 women. Br J Cancer 2001;84(12):1616-23.

36. Coutlee, F; Gravitt, P; Richardson, H; Hankins, C; Franco, E; Lapointe, N; Voyer, H. Nonisotopic detection and typing of human papillomavirus DNA in genital samples by the line blot assay. The Canadian Women's HIV study group. J Clin Microbiol 1999;37(6):1852-7.

37. Cox, DR. Partial likelihood. Biometrika 1975;62:269-76.

38. Cox, DR. Regression Models and Life Tables. J Royal Stat Soc 1972;34:187-220.

39. Cox, DR; Oakes, D. Analysis of survival data. New York: Chapman and Hall; 1984.

40. Cox, JT; Lorincz, AT; Schiffman, MH; Sherman, ME; Cullen, A; Kurman, RJ. Human papillomavirus testing by hybrid capture appears to be useful in triaging women with a cytologic diagnosis of atypical squamous cells of undetermined significance. Am J Obstet Gynecol 1995;172(3):946-54.

41. Crum, CP; Nagai, N; Levine, RU; Silverstein, S. In situ hybridization analysis of HPV 16 DNA sequences in early cervical neoplasia. Am J Pathol 1986;123(1):174-82.

42. Cuzick, J. Viral load as a surrogate for persistence in cervical human papillomavirus infection. Franco, E; Monsonego, J, (eds.). New Developments in Cervical Cancer Screening and Prevention.Blackwell Science Ltd.; 1997. pp. 373-8.

43. Cuzick, J; Sasieni, P; Davies, P; Adams, J; Normand, C; Frater, A; van Ballegooijen, M; van den Akker-van Marle, E. A systematic review of the role of human papilloma virus (HPV) testing within a cervical screening programme: summary and conclusions. Br J Cancer 2000;83(5):561-5.

44. Cuzick, J; Terry, G; Ho, L; Hollingworth, T; Anderson, M. Type-specific human papillomavirus DNA in abnormal smears as a predictor of high-grade cervical intraepithelial neoplasia. Br J Cancer 1994;69:167-71.

45. de Roda Husman, AM; Walboomers, JM; van den Brule, AJ; Meijer, CJ; Snijders, PJ. The use of general primers GP5 and GP6 elongated at their 3' ends with adjacent highly conserved sequences improves human papillomavirus detection by PCR. J Gen Virol 1995;76 (Pt 4):1057-62.

46. de Sanjose, S; Hamsikova, E; Munoz, N; Bosch, FX; Hofmannova, V; Gili, M; Izarzugaza, I; Viladiu, P; Tormo, MJ; Moreo, P; Munoz, MT; Ascunce, N; Tafur, L; Shah, KV; Vonka, V. Serological response to HPV16 in CIN-III and cervical cancer patients. Case-control studies in Spain and Colombia. Int J Cancer 1996;66(1):70-4.

47. de Sanjose, S; Munoz, N; Bosch, FX; Reimann, K; Pedersen, NS; Orfila, J; Ascunce, N; Gonzalez, LC; Tafur, L; Gili, M; et, al. Sexually transmitted agents and cervical neoplasia in Colombia and Spain. Int J Cancer 1994;56(3):358-63.

48. de Vet, HC; Knipschild, PG; Grol, ME; Schouten, HJ; Sturmans, F. The role of beta-carotene and other dietary factors in the aetiology of cervical dysplasia: results of a case-control study. Int J Epidemiol 1991;20(3):603-10.

49. de Villiers, EM. Hybridization methods other than PCR: an update. IARC Sci Publ 1992;119:111-9.

50. Delius, H; Saegling, B; Bergmann, K; Shamanin, V; de Villiers, EM. The genomes of three of four novel HPV types, defined by differences of their L1 genes, show high conservation of the E7 gene and the URR. Virology 1998;240(2):359-65.

51. Denny, L; Kuhn, L; Pollack, A; Wainwright, H; Wright, TC Jr. Evaluation of alternative methods of cervical cancer screening for resource-poor settings. Cancer 2000;89(4):826-33.

52. Diggle, PJ; Liang, K-Y; Zeger, SL. Analysis of Longitudinal Data. New York: Oxford University Press Inc.; 1996.

53. Ellerbrock, TV; Chiasson, MA; Bush, TJ; Sun, XW; Sawo, D; Brudney, K; Wright, TC Jr. Incidence of cervical squamous intraepithelial lesions in HIV-infected women. JAMA 2000;283(8):1031-7.

54. Fahey, MT; Irwig, L; Macaskill, P. Meta-analysis of Pap test accuracy. Am J Epidemiol 1995;141(7):680-9.

55. Fahs, MC; Plichta, SB; Mandelblatt, JS. Cost-effective policies for cervical cancer screening. An international review. Pharmacoeconomics 1996;9(3):211-30.

56. Fairley, CK; Chen, S; Tabrizi, SN; McNeil, J; Becker, G; Walker, R; Atkins, RC; Thomson, N; Allan, P; Woodburn, C; et, al. Prevalence of HPV DNA in cervical specimens in women with renal transplants: a comparison with dialysis-dependent patients and patients with renal impairment. Nephrol Dial Transplant 1994;9(4):416-20.

57. Ferenczy, A; Franco, E. Persistent human papillomavirus infection and cervical neoplasia. Lancet Oncol 2002;3(1):11-6.

58. Ferenczy, A; Robitaille, J; Franco, E; Arseneau, J; Richart, RM; Wright, TC. Conventional cervical cytologic smears vs. ThinPrep smears. A paired comparison study on cervical cytology. Acta Cytol 1996;40(6):1136-42.

59. Ferenczy, A; Winkler, B. Cervical intraepithelial neoplasia and condyloma. In: Kurman, RJ, (ed.). Blaustein's Pathology of the Female Genital Tract. 3 ed. New York: Springer-Verlag; 1987. pp. pp. 184-91.

60. Francis, DA; Schmid, SI; Howley, PM. Repression of the integrated papillomavirus E6/E7 promoter is required for growth suppression of cervical cancer cells. J Virol 2000;74(6):2679-86.

61. Franco, E; Syrjanen, K; de Wolf, C; Patnick, J; Ferenczy, A; McGoogan, E; Bosch, X; Singer, A; Munoz, N; Meheus, A; Monsonego, J. New developments in cervical cancer screening and prevention. Geneva, Switzerland, June 17-19 1996. Workshop. Cancer Epidemiol Biomarkers Prev 1996;5(10):853-6.

62. Franco, E; Villa, L; Rohan, T; Ferenczy, A; Petzl-Erler, M; Matlashewski, G. Design and methods of the Ludwig-McGill longitudinal study of the natural history of human papillomavirus infection and cervical neoplasia in Brazil. Ludwig-McGill Study Group. Rev Panam Salud Publica 1999;6(4):223-33.

63. Franco, E; Villa, LL; Richardson, H; Rohan, T; Ferenczy, A. Epidemiology of cervical human papillomavirus infection. In: Franco & Monsonego, (eds.). New Developments in Cervical Cancer Screening and Prevention. Osney Mead, Oxford: Blackwell Science; 1997. pp. 14-22.

64. Franco, EL. Cancer causes revisited: human papillomavirus and cervical neoplasia. J Natl Cancer Inst 1995;87(11):779-80.

65. Franco, EL. Measurement errors in epidemiological studies of human papillomavirus and cervical cancer. IARC Sci Publ 1992;119:181-97.

66. Franco, EL. Prognostic value of human papillomavirus in the survival of cervical cancer patients: an overview of the evidence. Cancer Epidemiol Biomarkers Prev 1992;1(6):499-504.

67. Franco, EL. The sexually transmitted disease model for cervical cancer: incoherent epidemiologic findings and the role of misclassification of human papillomavirus infection. Epidemiology 1991;2(2):98-106.

68. Franco, EL. Statistical issues in human papillomavirus testing and screening. Clin Lab Med 2000;20(2):345-67.

69. Franco, EL; Duarte-Franco, E; Ferenczy, A. Cervical cancer: epidemiology, prevention and the role of human papillomavirus infection. CMAJ 2001;164(7):1017-25.

70. Franco, EL; Ferenczy, A. Assessing gains in diagnostic utility when human papillomavirus testing is used as an adjunct to papanicolaou smear in the triage of women with cervical cytologic abnormalities. Am J Obstet Gynecol 1999;181(2):382-6.

71. Franco, EL; Ferenczy A. Cervix. Franco, EL; Rohan, TE. Precursor Lesions: Epidemiology,

Detection, and Prevention. New York: Springer-Verlag; 2002. pp. 249-86.

72. Franco, EL; Rohan, TE; Villa, LL. Epidemiologic evidence and human papillomavirus infection as a necessary cause of cervical cancer. J Natl Cancer Inst 1999;91(6):506-11.

73. Franco, EL; Villa, LL; Rahal, P; Ruiz, A. Molecular variant analysis as an epidemiological tool to study persistence of cervical human papillomavirus infection. J Natl Cancer Inst 1994;86(20):1558-9.

74. Franco, EL; Villa, LL; Ruiz, A; Costa, MC. Transmission of cervical human papillomavirus infection by sexual activity: differences between low and high oncogenic risk types. J Infect Dis 1995;172(3):756-63.

75. Franco, EL; Villa, LL; Sobrinho, JP; Prado, JM; Rousseau, MC; Desy, M; Rohan, TE. Epidemiology of acquisition and clearance of cervical human papillomavirus infection in women from a high-risk area for cervical cancer. J Infect Dis 1999;180(5):1415-23.

76. Freeman, J; Hutchison, GB. Prevalence, incidence and duration. Am J Epidemiol 1980;112(5):707-23.

77. Gay, JD; Donaldson, LD; Goellner, JR. False-negative results in cervical cytologic studies. Acta Cytol 1985;29(6):1043-6.

78. Glynn, RJ; Buring, JE. Ways of measuring rates of recurrent events. BMJ 1996;312(7027):364-7.

79. Goldberg, GL; Vermund, SH; Schiffman, MH; Ritter, DB; Spitzer, C; Burk, RD. Comparison of Cytobrush and cervicovaginal lavage sampling methods for the detection of genital human papillomavirus. Am J Obstet Gynecol 1989;161(6 Pt 1):1669-72.

80. Goldie, SJ; Kuhn, L; Denny, L; Pollack, A; Wright, TC. Policy analysis of cervical cancer screening strategies in low-resource settings: clinical benefits and cost-effectiveness. JAMA 2001;285(24):3107-15.

81. Goldstein, H. Multilevel Statistical Models. New York: Halstead Press; 1995.

82. Goldstein, H; Healy, MJ; Rasbash, J. Multilevel time series models with applications to repeated measures data. Stat Med 1994;13(16):1643-55.

83. Gram, IT; Macaluso, M; Churchill, J; Stalsberg, H. Trichomonas vaginalis (TV) and human papillomavirus (HPV) infection and the incidence of cervical intraepithelial neoplasia (CIN) grade III. Cancer Causes Control 1992;3(3):231-6.

84. Gravitt, PE; Manos, MM. Polymerase chain reaction-based methods for the detection of human papillomavirus DNA. IARC Sci Publ 1992;119:121-33.

85. Gravitt, PE; Peyton, CL; Alessi, TQ; Wheeler, CM; Coutlee, F; Hildesheim, A; Schiffman, MH; Scott, DR; Apple, RJ. Improved amplification of genital human papillomaviruses. J Clin Microbiol 2000;38(1):357-61.

86. Greenland, S. Introduction to Regression Modeling. In: Rothman, KJ; Greenland, S, (eds.). Modern Epidemiology. second ed. Philadelphia: Lippincott-Raven; 1998. pp. 401-32.

87. Greenland, S. Introduction to Regression Models. In: Rothman, KJ; Greenland, S, (eds.). Modern Epidemiology. second ed. Philadelphia: Lippincott-Raven; 1998. pp. 359-99.

88. Guerrero, E; Daniel, RW; Bosch, FX; Castellsague, X; Munoz, N; Gili, M; Viladiu, P; Navarro, C; Zubiri, ML; Ascunce, N; et, al. Comparison of ViraPap, Southern hybridization, and polymerase chain reaction methods for human papillomavirus identification in an epidemiological investigation of cervical cancer. J Clin Microbiol 1992;30(11):2951-9.

89. Hall, S; Lorincz, A; Shah, F; Sherman, ME; Abbas, F; Paull, G; Kurman, RJ; Shah, KV. Human papillomavirus DNA detection in cervical specimens by hybrid capture: correlation with cytologic and histologic diagnoses of squamous intraepithelial lesions of the cervix. Gynecol Oncol 1996;62(3):353-9.

90. Harizi, H; Juzan, M; Pitard, V; Moreau, JF; Gualde, N. Cyclooxygenase-2-issued prostaglandin e(2) enhances the production of endogenous IL-10, which down-regulates dendritic cell functions. J Immunol 2002;168(5):2255-63.

91. Harris, RW; Brinton, LA; Cowdell, RH; Skegg, DC; Smith, PG; Vessey, MP; Doll, R. Characteristics of women with dysplasia or carcinoma in situ of the cervix uteri. Br J Cancer 1980;42(3):359-69.

92. Heard, I; Schmitz, V; Costagliola, D; Orth, G; Kazatchkine, MD. Early regression of cervical lesions in HIV-seropositive women receiving highly active antiretroviral therapy. AIDS 1998;12(12):1459-64.

93. Heinzl, H; Kaider, A; Zlabinger, G. Assessing interactions of binary time-dependent covariates with time in cox proportional hazards regression models using cubic spline functions. Stat Med 1996;15(23):2589-601.

94. Helland, A; Borresen, AL; Kaern, J; Ronningen, KS; Thorsby, E. HLA antigens and cervical carcinoma. Nature 1992;356(6364):23.

95. Helland, A; Langerod, A; Johnsen, H; Olsen, AO; Skovlund, E; Borresen-Dale, AL. p53 polymorphism and risk of cervical cancer. Nature 1998;396(6711):530-1; discussion 532.

96. Helland, A; Olsen, AO; Gjoen, K; Akselsen, HE; Sauer, T; Magnus, P; Borresen-Dale, AL; Ronningen, KS. An increased risk of cervical intra-epithelial neoplasia grade II-III among human papillomavirus positive patients with the HLA-DQA1*0102-DQB1*0602 haplotype: a population-based case-control study of Norwegian women. Int J Cancer 1998;76(1):19-24.

97. Herbst, AL. The Bethesda system for cervical/vaginal cytologic diagnoses. Clin Obstet Gynecol 1992;35(1):22-7.

98. Herrero, R; Brinton, LA; Reeves, WC; Brenes, MM; de Britton, RC; Gaitan, E; Tenorio, F. Screening for cervical cancer in Latin America: a case-control study. Int J Epidemiol 1992;21(6):1050-6.

99. Herrero, R; Hildesheim, A; Bratti, C; Sherman, ME; Hutchinson, M; Morales, J; Balmaceda, I; Greenberg, MD; Alfaro, M; Burk, RD; Wacholder, S; Plummer, M; Schiffman, M. Population-based study of human papillomavirus infection and cervical neoplasia in rural Costa Rica. J Natl Cancer Inst 2000;92(6):464-74.

100. Herrero, R; Potischman, N; Brinton, LA; Reeves, WC; Brenes, MM; Tenorio, F; Debritton, RC; Gaitan, E. A case-control study of nutrient status and invasive cervical cancer .I. dietary indicators. Am J Epidemiol 1991;134:1335-46.

101. Herrero, R; Schiffman, MH; Bratti, C; Hildesheim, A; Balmaceda, I; Sherman, ME; Greenberg, M; Cardenas, F; Gomez, V; Helgesen, K; Morales, J; Hutchinson, M; Mango, L; Alfaro, M; Potischman, NW; Wacholder, S; Swanson, C; Brinton, LA. Design and methods of a population-based natural history study of cervical neoplasia in a rural province of Costa Rica: the Guanacaste Project. Rev Panam Salud Publica 1997;1(5):362-75.

102. Hildesheim, A; Schiffman, M; Brinton, LA; Fraumeni, JF Jr; Herrero, R; Bratti, MC; Schwartz, P; Mortel, R; Barnes, W; Greenberg, M; McGowan, L; Scott, DR; Martin, M; Herrera, JE; Carrington, M. p53 polymorphism and risk of cervical cancer. Nature 1998;396(6711):531-2.

103. Hildesheim, A; Schiffman, M; Bromley, C; Wacholder, S; Herrero, R; Rodriguez, A; Bratti, MC; Sherman, ME; Scarpidis, U; Lin, QQ; Terai, M; Bromley, RL; Buetow, K; Apple, RJ; Burk, RD. Human papillomavirus type 16 variants and risk of cervical cancer. J Natl Cancer Inst 2001;93(4):315-8.

104. Hildesheim, A; Schiffman, M; Scott, DR; Marti, D; Kissner, T; Sherman, ME; Glass, AG; Manos, MM; Lorincz, AT; Kurman, RJ; Buckland, J; Rush, BB; Carrington, M. Human leukocyte antigen class I/II alleles and development of human papillomavirus-related cervical neoplasia: results from a case-control study conducted in the United States. Cancer Epidemiol Biomarkers Prev 1998;7(11):1035-41.

105. Hildesheim, A; Schiffman, MH; Gravitt, PE; Glass, AG; Greer, CE; Zhang, T; Scott, DR; Rush, BB; Lawler, P; Sherman, ME; Kurman, RJ; Manos, MM. Persistence of type-specific human papillomavirus infection among cytologically normal women. J Infect Dis 1994;169(2):235-40.

106. Ho, GY; Bierman, R; Beardsley, L; Chang, CJ; Burk, RD. Natural history of cervicovaginal papillomavirus infection in young women. N Engl J Med 1998;338(7):423-8.

107. Ho, GY; Burk, RD; Klein, S; Kadish, AS; Chang, CJ; Palan, P; Basu, J; Tachezy, R; Lewis, R; Romney, S. Persistent genital human papillomavirus infection as a risk factor for persistent cervical dysplasia. J Natl Cancer Inst 1995;87(18):1365-71.

108. Ho, GY; Palan, PR; Basu, J; Romney, SL; Kadish, AS; Mikhail, M; Wassertheil-Smoller, S; Runowicz, C; Burk, RD. Viral characteristics of human papillomavirus infection and antioxidant levels as risk factors for cervical dysplasia. Int J Cancer 1998;78(5):594-9.

109. Holowaty, P; Miller, AB; Rohan, T; To, T. Natural history of dysplasia of the uterine cervix. J Natl Cancer Inst 1999;91(3):252-8.

110. IARC Working Group. Human papillomaviruses. IARC Monographs on the evalutaion of carcinogenic risks to humans, Vol. 64. Lyon: International Agency for Research on Cancer; 1995.

111. Jacobs, MV; Snijders, PJ; Voorhorst, FJ; Dillner, J; Forslund, O; Johansson, B; von Knebel Doeberitz, M; Meijer, CJ; Meyer, T; Nindl, I; Pfister, H; Stockfleth, E; Strand, A; Wadell, G; Walboomers, JM. Reliable high risk HPV DNA testing by polymerase

chain reaction: an intermethod and intramethod comparison. J Clin Pathol 1999;52(7):498-503.

112. Jha, PK; Beral, V; Peto, J; Hack, S; Hermon, C; Deacon, J; Mant, D; Chilvers, C; Vessey, MP; Pike, MC; et, al. Antibodies to human papillomavirus and to other genital infectious agents and invasive cervical cancer risk. Lancet 1993;341(8853):1116-8.

113. Jones, CJ; Brinton, LA; Hamman, RF; Stolley, PD; Lehman, HF; Levine, RS; Mallin, K. Risk factors for insitu cervical cancer: results from a case-control study. Cancer Res 1990;50:3657-62.

114. Josefsson, A; Livak, K; Gyllensten, U. Detection and quantitation of human papillomavirus by using the fluorescent 5' exonuclease assay. J Clin Microbiol 1999;37(3):490-6.

115. Josefsson, AM; Magnusson, PK; Ylitalo, N; Quarforth-Tubbin, P; Ponten, J; Adami, HO; Gyllensten, UB. p53 polymorphism and risk of cervical cancer. Nature 1998;396(6711):531; discussion 532.

116. Josefsson, AM; Magnusson, PK; Ylitalo, N; Sorensen, P; Qwarforth-Tubbin, P; Andersen, PK; Melbye, M; Adami, HO; Gyllensten, UB. Viral load of human papilloma virus 16 as a determinant for development of cervical carcinoma in situ: a nested case-control study. Lancet 2000;355(9222):2189-93.

117. Kadish, AS; Ho, GY; Burk, RD; Wang, Y; Romney, SL; Ledwidge, R; Angeletti, RH. Lymphoproliferative responses to human papillomavirus (HPV) type 16 proteins E6 and E7: outcome of HPV infection and associated neoplasia. J Natl Cancer Inst 1997;89(17):1285-93.

118. Kadish, AS; Timmins, P; Wang, Y; Ho, GY; Burk, RD; Ketz, J; He, W; Romney, SL; Johnson, A; Angeletti, R; Abadi, M. Regression of cervical intraepithelial neoplasia and loss of human papillomavirus (HPV) infection is associated with cell-mediated immune responses to an HPV type 16 E7 peptide. Cancer Epidemiol Biomarkers Prev 2002;11(5):483-8.

119. Kahn, JA; Rosenthal, SL; Succop, PA; Ho, GY; Burk, RD. Mediators of the association between age of first sexual intercourse and subsequent human papillomavirus infection. Pediatrics 2002;109(1):E5.

120. Kainz, C; Gitsch, G; Heinzl, H; Breitenecker, G. Incidence of cervical smears indicating dysplasia among Austrian women during the 1980s. Br J Obstet Gynaecol 1995;102(7):541-4.

121. Kaplan, EL; Meier, P. Nonparametric estimation for incomplete observations. J Amer Stat Assoc 1958;53:457-81.

122. Karaloglu, D; Yazici, H; Alatli, C; Yaman, F; Aslay, I; Onat, H; Dalay, N. Detection of HPV 16 and HPV 18 infection in patients with cervical neoplasia. Eur J Gynaecol Oncol 1996;17(4):296-8.

123. Kataja, V; Syrjanen, K; Mantyjarvi, R; Vayrynen, M; Syrjanen, S; Saarikoski, S; Parkinen, S; Yliskoski, M; Salonen, JT; Castren, O. Prospective follow-up of cervical HPV infections: life table analysis of histopathological, cytological and colposcopic data. Eur J Epidemiol 1989;5:1-7.

124. Kaufman, RH; Adam, E; Icenogle, J; Lawson, H; Lee, N; Reeves, KO; Irwin, J; Simon, T;

Press, M; Uhler, R; Entman, C; Reeves, WC. Relevance of human papillomavirus screening in management of cervical intraepithelial neoplasia. Am J Obstet Gynecol 1997;176(1 Pt 1):87-92.

125. Kjaer, SK; Teisen, C; Haugaard, BJ; Lynge, E; Christensen, RB; Moller, KA; Jensen, H; Poll, P; Vestergaard, BF; DeVilliers, E-M; Jensen, OM. Risk factors for cervical cancer in Greenland and Denmark: a population-based cross-sectional study. Int J Cancer 1989;44:40-7.

126. Klaes, R; Woerner, SM; Ridder, R; Wentzensen, N; Duerst, M; Schneider, A; Lotz, B; Melsheimer, P; von Knebel Doeberitz, M. Detection of high-risk cervical intraepithelial neoplasia and cervical cancer by amplification of transcripts derived from integrated papillomavirus oncogenes. Cancer Res 1999;59(24):6132-6.

127. Klein, RS; Ho, GY; Vermund, SH; Fleming, I; Burk, RD. Risk factors for squamous intraepithelial lesions on Pap smear in women at risk for human immunodeficiency virus infection. J Infect Dis 1994;170(6):1404-9.

128. Kleinbaum, DG. Survival Analysis: A Self-learning Text. New York: Springer-Verlag Inc.; 1996.

129. Konya, J; Veress, G; Juhasz, A; Szarka, K; Sapy, T; Hernadi, Z; Gergely, L. Additional human papillomavirus types detected by the hybrid capture tube test among samples from women with cytological and colposcopical atypia. J Clin Microbiol 2000;38(1):408-11.

130. Kornegay, JR; Shepard, AP; Hankins, C; Franco, E; Lapointe, N; Richardson, H; Coutlee, F. Nonisotopic detection of human papillomavirus DNA in clinical specimens using a consensus PCR and a generic probe mix in an enzyme-linked immunosorbent assay format. J Clin Microbiol 2001;39(10):3530-6.

131. Koutsky, LA; Holmes, KK; Critchlow, CW; Stevens, CE; Paavonen, J; Beckmann, AM; DeRouen, TA; Galloway, DA; Vernon, D; Kiviat, NB. A cohort study of the risk of cervical intraepithelial neoplasia grade 2 or 3 in relation to papillomavirus infection. N Engl J Med 1992;327(18):1272-8.

132. La Vecchia, C; Decarli, A; Fasoli, M; Parazzini, F; Franceschi, S; Gentile, A; Negri, E. Dietary vitamin A and the risk of intraepithelial and invasive cervical neoplasia. Gynecol Oncol 1988;30(2):187-95.

133. Lacey, JV Jr; Brinton, LA; Abbas, FM; Barnes, WA; Gravitt, PE; Greenberg, MD; Greene, SM; Hadjimichael, OC; McGowan, L; Mortel, R; Schwartz, PE; Silverberg, SG; Hildesheim, A. Oral contraceptives as risk factors for cervical adenocarcinomas and squamous cell carcinomas. Cancer Epidemiol Biomarkers Prev 1999;8(12):1079-85.

134. Lacey, JV Jr; Frisch, M; Brinton, LA; Abbas, FM; Barnes, WA; Gravitt, PE; Greenberg, MD; Greene, SM; Hadjimichael, OC; McGowan, L; Mortel, R; Schwartz, PE; Zaino, RJ; Hildesheim, A. Associations between smoking and adenocarcinomas and squamous cell carcinomas of the uterine cervix (United States). Cancer Causes Control 2001;12(2):153-61.

135. Lamarcq, L; Deeds, J; Ginzinger, D; Perry, J; Padmanabha, S; Smith-McCune, K. Measurements of human papillomavirus transcripts by real time quantitative reverse transcription-polymerase chain reaction in samples collected for cervical

cancer screening. J Mol Diagn 2002;4(2):97-102.

136. Lawson, HW; Lee, NC; Thames, SF; Henson, R; Miller, DS. Cervical cancer screening among low-income women: results of a national screening program, 1991-1995. Obstet Gynecol 1998;92(5):745-52.

137. Lehtinen, M; Dillner, J; Knekt, P; Luostarinen, T; Aromaa, A; Kirnbauer, R; Koskela, P; Paavonen, J; Peto, R; Schiller, JT; Hakama, M. Serologically diagnosed infection with human papillomavirus type 16 and risk for subsequent development of cervical carcinoma: nested case- control study. BMJ 1996;312(7030):537-9.

138. Liaw, KL; Glass, AG; Manos, MM; Greer, CE; Scott, DR; Sherman, M; Burk, RD; Kurman, RJ; Wacholder, S; Rush, BB; Cadell, DM; Lawler, P; Tabor, D; Schiffman, M. Detection of human papillomavirus DNA in cytologically normal women and subsequent cervical squamous intraepithelial lesions. J Natl Cancer Inst 1999;91(11):954-60.

139. Liaw, KL; Hildesheim, A; Burk, RD; Gravitt, P; Wacholder, S; Manos, MM; Scott, DR; Sherman, ME; Kurman, RJ; Glass, AG; Anderson, SM; Schiffman, M. A prospective study of human papillomavirus (HPV) type 16 DNA detection by polymerase chain reaction and its association with acquisition and persistence of other HPV types. J Infect Dis 2001;183(1):8-15.

140. Liaw, KL; Hsing, AW; Chen, CJ; Schiffman, MH; Zhang, TY; Hsieh, CY; Greer, CE; You, SL; Huang, TW; Wu, TC; et, al. Human papillomavirus and cervical neoplasia: a case-control study in Taiwan. Int J Cancer 1995;62(5):565-71.

141. Liu, T; Soong, SJ; Alvarez, RD; Butterworth, CE Jr. A longitudinal analysis of human papillomavirus 16 infection, nutritional status, and cervical dysplasia progression. Cancer Epidemiol Biomarkers Prev 1995;4(4):373-80.

142. Londesborough, P; Ho, L; Terry, G; Cuzick, J; Wheeler, C; Singer, A. Human papillomavirus genotype as a predictor of persistence and development of high-grade lesions in women with minor cervical abnormalities. Int J Cancer 1996;69(5):364-8.

143. Lorincz, A. Detection of human papillomavirus DNA without amplification: prospects for clinical utility. IARC Sci Publ 1992;119:135-45.

144. Luft, F; Klaes, R; Nees, M; Durst, M; Heilmann, V; Melsheimer, P; von Knebel Doeberitz, M. Detection of integrated papillomavirus sequences by ligation-mediated PCR (DIPS-PCR) and molecular characterization in cervical cancer cells. Int J Cancer 2001;92(1):9-17.

145. Maciag, PC; Schlecht, NF; Souza, PS; Franco, EL; Villa, LL; Petzl-Erler, ML. Major histocompatibility complex class II polymorphisms and risk of cervical cancer and human papillomavirus infection in Brazilian women. Cancer Epidemiol Biomarkers Prev 2000;9(11):1183-91.

146. Maciag, PC; Villa, LL. Genetic susceptibility to HPV infection and cervical cancer. Braz J Med Biol Res 1999;32(7):915-22.

147. Makni, H; Franco, EL; Kaiano, J; Villa, LL; Labrecque, S; Dudley, R; Storey, A; Matlashewski, G. P53 polymorphism in codon 72 and risk of human papillomavirus-induced cervical cancer: effect of inter-laboratory variation. Int J Cancer

2000;87(4):528-33.

148. Maldonado, G; Greenland, S. Simulation study of confounder-selection strategies. Am J Epidemiol 1993;138(11):923-36.

149. Manos, MM; Ting, Y; Wright, DK; Lewis, AJ; Broker, TR; Wolinski, SM. Use of polymerase chain reaction amplification for the detection of genital human papillomaviruses. Cancer Cells 1989;7:209-14.

150. McCrory, DC; Matchar, DB; Bastian, L; et al. Evaluation of cervical cytology. Evidence report/technology assessment No. 5. AHCPR Publication No. 99-E010. Rockville, MD: Agency for Health Care Policy and Research, U.S. Department of Health and Human Services; 1999.

151. McDougall, JK; Nelson, JA; Myerson, D; Beckmann, AM; Galloway, DA. HSV, CMV, and HPV in human neoplasia. J Invest Dermatol 1984;83(1 Suppl):72s-6s.

152. McIndoe, WA; McLean, MR; Jones, RW; Mullins, PR. The invasive potential of carcinoma in situ of the cervix. Obstet Gynecol 1984;64(4):451-8.

153. Meanwell, CA; Cox, MF; Blackledge, G; Maitland, NJ. HPV 16 DNA in normal and malignant cervical epithelium: implications for the aetiology and behaviour of cervical neoplasia. Lancet 1987;1(8535):703-7.

154. Meijer, CJ; van den Brule, AJ; Snijders, PJ; Helmerhorst, T; Kenemans, P; Walboomers, JM. Detection of human papillomavirus in cervical scrapes by the polymerase chain reaction in relation to cytology: possible implications for cervical cancer screening. IARC Sci Publ 1992;119:271-81.

155. Melnikow, J; Nuovo, J; Willan, AR; Chan, BK; Howell, LP. Natural history of cervical squamous intraepithelial lesions: a meta- analysis. Obstet Gynecol 1998;92(4 Pt 2):727-35.

156. Menczer, J; Fintsi, Y; Arbel-Alon, S; Tell, L; Friedman, E; Jackman, A; Sherman, L. The presence of HPV 16, 18 and p53 immunohistochemical staining in tumor tissue of Israeli Jewish women with cervical and vulvar neoplasia. Eur J Gynaecol Oncol 2000;21(1):30-4.

157. Miller, AB. The role of early diagnosis and screening in Oncology. Diagn Oncol 1991;1:19-27.

158. Miller, AB; Knight, J; Narod, S. The natural history of cacner of the cervix, and the implications for screening policy. In: Miller, AB; Chamberlain, J; Day, NE; Hakama, M; Prorok, PC, (eds.). Cancer Screening. Cambridge, UK: Cambridge Press; 1991.

159. Miller, AB; Nazeer, S; Fonn, S; Brandup-Lukanow, A; Rehman, R; Cronje, H; Sankaranarayanan, R; Koroltchouk, V; Syrjanen, K; Singer, A; Onsrud, M. Report on consensus conference on cervical cancer screening and management. Int J Cancer 2000;86(3):440-7.

160. Mirra, A. P., (coordenador). Cancer Incidence in Sao Paulo County, Brazil: 1983-1988-1993: Trend in the period 1969-1993. Brazilia, DF: Ministerio da Saude; 1999.

161. Mitchell, HS. Longitudinal analysis of histologic high-grade disease after negative cervical cytology according to endocervical status. Cancer 2001;93(4):237-40.

162. Mitchell, MF; Hittelman, WN; Hong, WK; Lotan, R; Schottenfeld, D. The natural history of cervical intraepithelial neoplasia: an argument for intermediate endpoint biomarkers. Cancer Epidemiol Biomarkers Prev 1994;3(7):619-26.

163. Mitchell, MF; Schottenfeld, D; Tortolero-Luna, G; Cantor, SB; Richards-Kortum, R. Colposcopy for the diagnosis of squamous intraepithelial lesions: a meta-analysis. Obstet Gynecol 1998;91(4):626-31.

164. Mitchell, MF; Tortolero-Luna, G; Wright, T; Sarkar, A; Richards-Kortum, R; Hong, WK; Schottenfeld, D. Cervical human papillomavirus infection and intraepithelial neoplasia: a review. J Natl Cancer Inst Monogr 1996;(21):17-25.

165. Mitrani-Rosenbaum, S; Tsvieli, R; Tur-Kaspa, R. Oestrogen stimulates differential transcription of human papillomavirus type 16 in SiHa cervical carcinoma cells. J Gen Virol 1989;70 (Pt 8):2227-32.

166. Montz, FJ. Management of high-grade cervical intraepithelial neoplasia and low- grade squamous intraepithelial lesion and potential complications. Clin Obstet Gynecol 2000;43(2):394-409.

167. Moreno, V; Bosch, FX; Munoz, N; Meijer, CJ; Shah, KV; Walboomers, JM; Herrero, R; Franceschi, S. Effect of oral contraceptives on risk of cervical cancer in women with human papillomavirus infection: the IARC multicentric case-control study. Lancet 2002;359(9312):1085-92.

168. Morris, M; Tortolero-Luna, G; Malpica, A; Baker, VV; Cook, E; Johnson, E; Follen Mitchell, M. Cervical intraepithelial neoplasia and cervical cancer. Obstet Gynecol Clin North Am 1996;23(2):347-410.

169. Morrison, BJ; Coldman, AJ; Boyes, DA; Anderson, GH. Forty years of repeated screening: the significance of carcinoma in situ. Br J Cancer 1996;74(5):814-9.

170. Moscicki, AB; Ellenberg, JH; Vermund, SH; Holland, CA; Darragh, T; Crowley-Nowick, PA; Levin, L; Wilson, CM. Prevalence of and risks for cervical human papillomavirus infection and squamous intraepithelial lesions in adolescent girls: impact of infection with human immunodeficiency virus. Arch Pediatr Adolesc Med 2000;154(2):127-34.

171. Moscicki, AB; Hills, N; Shiboski, S; Powell, K; Jay, N; Hanson, E; Miller, S; Clayton, L; Farhat, S; Broering, J; Darragh, T; Palefsky, J. Risks for incident human papillomavirus infection and low-grade squamous intraepithelial lesion development in young females. JAMA 2001;285(23):2995-3002.

172. Moscicki, AB; Palefsky, J; Gonzales, J; Schoolnik, GK. Human papillomavirus infection in sexually active adolescent females: prevalence and risk factors. Pediatr Res 1990;28(5):507-13.

173. Moscicki, AB; Palefsky, JM; Gonzales, J; Smith, G; Schoolnik, GK. Colposcopic and histologic findings and human papillomavirus (HPV) DNA test variability in young women positive for HPV DNA. J Infect Dis 1992;166:951-7.

174. Moscicki, AB; Shiboski, S; Broering, J; Powell, K; Clayton, L; Jay, N; Darragh, TM; Brescia, R; Kanowitz, S; Miller, SB; Stone, J; Hanson, E; Palefsky, J. The natural history of human papillomavirus infection as measured by repeated DNA testing in adolescent and young women. J Pediatr 1998;132(2):277-84.

175. Moscicki, AB; Winkler, B; Irwin, CE Jr; Schachter, J. Differences in biologic maturation, sexual behavior, and sexually transmitted disease between adolescents with and without cervical intraepithelial neoplasia. J Pediatr 1989;115(3):487-93.

176. Muderspach, L; Wilczynski, S; Roman, L; Bade, L; Felix, J; Small, LA; Kast, WM; Fascio, G; Marty, V; Weber, J. A phase I trial of a human papillomavirus (HPV) peptide vaccine for women with high-grade cervical and vulvar intraepithelial neoplasia who are HPV 16 positive. Clin Cancer Res 2000;6(9):3406-16.

177. Munoz, N; Bosch, FX. The causal link between HPV and cervical cancer and its implications for prevention of cervical cancer. Bull Pan Am Health Organ 1996;30(4):362-77.

178. Munoz, N; Bosch, FX; Desanjose, S; Tafur, L; Izarzugaza, I; Gili, M; Viladiu, P; Navarro, C; Martos, C; Ascunce, N; Gonzalez, LC; Kaldor, JM; Guerrero, E; Lorincz, A; Santamaria, M; Deruiz, PA; Aristizabal, N; Shah, K. The causal link between human papillomavirus and invasive cervical cancer: a population-based case-control study in Colombia and Spain. Int J Cancer 1992;52(5):743-9.

179. Munoz, N; Franceschi, S; Bosetti, C; Moreno, V; Herrero, R; Smith, JS; Shah, KV; Meijer, CJ; Bosch, FX. Role of parity and human papillomavirus in cervical cancer: the IARC multicentric case-control study. Lancet 2002;359(9312):1093-101.

180. Myers, ER; McCrory, DC; Nanda, K; Bastian, L; Matchar, DB. Mathematical model for the natural history of human papillomavirus infection and cervical carcinogenesis. Am J Epidemiol 2000;151(12):1158-71.

181. Nagao, S; Yoshinouchi, M; Miyagi, Y; Hongo, A; Kodama, J; Itoh, S; Kudo, T. Rapid and sensitive detection of physical status of human papillomavirus type 16 DNA by quantitative real-time PCR. J Clin Microbiol 2002;40(3):863-7.

182. Nanda, K; McCrory, DC; Myers, ER; Bastian, LA; Hasselblad, V; Hickey, JD; Matchar, DB. Accuracy of the Papanicolaou test in screening for and follow-up of cervical cytologic abnormalities: a systematic review. Ann Intern Med 2000;132(10):810-9.

183. Nawa, A; Nishiyama, Y; Kobayashi, T; Wakahara, Y; Okamoto, T; Kikkawa, F; Suganuma, N; Goto, S; Kuzuya, K; Tomoda, Y. Association of human leukocyte antigen-B1*03 with cervical cancer in Japanese women aged 35 years and younger. Cancer 1995;75(2):518-21.

184. NCI Workshop. The Bethesda System for reporting cervical/vaginal cytologic diagnoses: revised after the second National Cancer Institute Workshop, Bethesda, Maryland, USA. April 29-30, 1991. Acta Cytol 1993;37:115-24.

185. Nobbenhuis, MA; Walboomers, JM; Helmerhorst, TJ; Rozendaal, L; Remmink, AJ; Risse, EK; van der Linden, HC; Voorhorst, FJ; Kenemans, P; Meijer, CJ. Relation of human papillomavirus status to cervical lesions and consequences for cervical-cancer screening: a prospective study. Lancet 1999;354(9172):20-5.

186. Olsen, AO; Gjoen, K; Sauer, T; Orstavik, I; Naess, O; Kierulf, K; Sponland, G; Magnus, P. Human papillomavirus and cervical intraepithelial neoplasia grade II- III: a population-based case-control study. Int J Cancer 1995;61(3):312-5.

187. Ostor, AG. Natural history of cervical intraepithelial neoplasia: a critical review. Int J Gynecol Pathol 1993;12(2):186-92.

188. Palefsky, JM; Holly, EA. Molecular virology and epidemiology of human papillomavirus and cervical cancer. Cancer Epidemiol Biomarkers Prev 1995;4(4):415-28.

189. Palefsky, JM; Minkoff, H; Kalish, LA; Levine, A; Sacks, HS; Garcia, P; Young, M; Melnick, S; Miotti, P; Burk, R. Cervicovaginal human papillomavirus infection in human immunodeficiency virus-1 (HIV)-positive and high-risk HIV-negative women. J Natl Cancer Inst 1999;91(3):226-36.

190. Papanicolaou, GN. Atlas of exfoliative cytology. Cambridge, MA: Harvard University Press; 1954.

191. Papanicolaou, GN; Traut, HF. Diagnosis of Uterine Cancer by Vaginal Smears. New York: The Commonwealth Fund; 1943.

192. Parazzini, F; La Vecchia, C; Negri, E; Fedele, L; Franceschi, S; Gallotta, L. Risk factors for cervical intraepithelial neoplasia. Cancer 1992;69(9):2276-82.

193. Parkin, DM; Bray, FI; Devesa, SS. Cancer burden in the year 2000. The global picture. Eur J Cancer 2001;37 Suppl 8:S4-66.

194. Parkin, DM; Hodgson, P; Clayden, AD. Incidence and prevalence of preclinical carcinoma of cervix in a British population. Br J Obstet Gynaecol 1982;89(7):564-70.

195. Pater, MM; Mittal, R; Pater, A. Role of steroid hormones in potentiating transformation of cervical cells by human papillomaviruses. Trends Microbiol 1994;2(7):229-34.

196. Peitsaro, P; Johansson, B; Syrjanen, S. Integrated human papillomavirus type 16 is frequently found in cervical cancer precursors as demonstrated by a novel quantitative real-time PCR technique. J Clin Microbiol 2002;40(3):886-91.

197. Peto, R; Pike, MC; Armitage, P; Breslow, NE; Cox, DR; Howard, SV; Mantel, N; McPherson, K; Peto, J; Smith, PG. Design and analysis of randomized clinical trials requiring prolonged observation of each patient. I. Introduction and design. Br J Cancer 1976;34(6):585-612.

198. Pfister, H. Papillomaviruses: particles genome organization and proteins. Syrjanen, K; Gissman, L; Koss, L. Papillomaviruses and Human Disease. Berlin: Springer-Verlag; 1987. pp. 1-18.

199. Pisani, P; Parkin, DM; Munoz, N; Ferlay, J. Cancer and infection: estimates of the attributable fraction in 1990. Cancer Epidemiol Biomarkers Prev 1997;6(6):387-400.

200. Prentice, RL; Williams, BJ; Peterson, AV. On the Regression Analysis of Multivariate Failure Time Data. Biometrika 1981;68(2):373-9.

201. Qu, W; Jiang, G; Cruz, Y; Chang, CJ; Ho, GY; Klein, RS; Burk, RD. PCR detection of human papillomavirus: comparison between MY09/MY11 and GP5+/GP6+ primer systems. J Clin Microbiol 1997;35(6):1304-10.

202. Ratnam, S; Franco, EL; Ferenczy, A. Human papillomavirus testing for primary screening of cervical cancer precursors. Cancer Epidemiol Biomarkers Prev 2000;9(9):945-51.

203. Reagan, JW; Seidemann, IL; Saracusa, Y. The cellular morphology of carcinoma in situ

and dysplasia or atypical hyperplasia of the uterine cervix. Cancer 1953;6:224-35.

204. Rho, J; Roy-Burman, A; Kim, H; de Villiers, EM; Matsukura, T; Choe, J. Nucleotide sequence and phylogenetic classification of human papillomavirus type 59. Virology 1994;203(1):158-61.

205. Richardson, H; Franco, E; Pintos, J; Bergeron, J; Arella, M; Tellier, P. Determinants of low-risk and high-risk cervical human papillomavirus infections in Montreal University students. Sex Transm Dis 2000;27(2):79-86.

206. Richart, RM. Influence of diagnostic and therapeutic procedures on the distribution of cervical intraepithelial neoplasia. Cancer 1966;19(11):1635-8.

207. Richart, RM. A modified terminology for cervical intraepithelial neoplasia. Obstet Gynecol 1990;75(1):131-3.

208. Richart, RM. Natural history of cervical intraepithelial neoplasia. Clin Obstet Gynecol 1968;5:748-84.

209. Rothman, KJ; Greenland, S. Measures of Disease Frequency. In: Rothman, KJ; Greenland, S, (eds.). Modern Epidemiology. second ed. Philadelphia: Lippincott-Raven; 1998. pp. 29-46.

210. Rousseau, MC; Pereira, JS; Prado, JC; Villa, LL; Rohan, TE; Franco, EL. Cervical coinfection with human papillomavirus (HPV) types as a predictor of acquisition and persistence of HPV infection. J Infect Dis 2001;184(12):1508-17.

211. Rousseau, MC; Trevisan, A; Villa, L; Rohan, T; Franco, EViral load as a predictor of HPV infection persistence among women in the Ludwig-McGill cohort study in Sao Paulo Brazil. In: Proceedings of the: 19th International Papillomavirus Conference; Castão do Santinho, Florianopolis, Brazil. 2001. p. 130.

212. Sanjeevi, CB; Hjelmstrom, P; Hallmans, G; Wiklund, F; Lenner, P; Angstrom, T; Dillner, J; Lernmark, A. Different HLA-DR-DQ haplotypes are associated with cervical intraepithelial neoplasia among human papillomavirus type-16 seropositive and seronegative Swedish women. Int J Cancer 1996;68(4):409-14.

213. Sasagawa, T; Basha, W; Yamazaki, H; Inoue, M. High-risk and multiple human papillomavirus infections associated with cervical abnormalities in Japanese women. Cancer Epidemiol Biomarkers Prev 2001;10(1):45-52.

214. Sastre-Garau, X; Loste, MN; Vincent-Salomon, A; Favre, M; Mouret, E; de la Rochefordiere, A; Durand, JC; Tartour, E; Lepage, V; Charron, D. Decreased frequency of HLA-DRB1 13 alleles in Frenchwomen with HPV- positive carcinoma of the cervix. Int J Cancer 1996;69(3):159-64.

215. Sawaya, GF; Grady, D; Kerlikowske, K; Valleur, JL; Barnabei, VM; Bass, K; Snyder, TE; Pickar, JH; Agarwal, SK; Mandelblatt, J. The positive predictive value of cervical smears in previously screened postmenopausal women: the Heart and Estrogen/progestin Replacement Study (HERS). Ann Intern Med 2000;133(12):942-50.

216. Sawaya, GF; Kerlikowske, K; Lee, NC; Gildengorin, G; Washington, AE. Frequency of cervical smear abnormalities within 3 years of normal cytology. Obstet Gynecol 2000;96(2):219-23.

217. Schiff, M; Becker, TM; Masuk, M; van Asselt-King, L; Wheeler, CM; Altobelli, KK; North, CQ; Nahmias, AJ. Risk factors for cervical intraepithelial neoplasia in southwestern American Indian women. Am J Epidemiol 2000;152(8):716-26.

218. Schiffman, M; Adrianza, ME. ASCUS-LSIL Triage Study. Design, methods and characteristics of trial participants. Acta Cytol 2000;44(5):726-42.

219. Schiffman, MH. Validation of hybridization assays: correlation of filter in situ, dot blot and PCR with Southern blot. IARC Sci Publ 1992;(119):169-79.

220. Schiffman, MH; Bauer, HM; Hoover, RN; Glass, AG; Cadell, DM; Rush, BB; Scott, DR; Sherman, ME; Kurman, RJ; Wacholder, S; et, al. Epidemiologic evidence showing that human papillomavirus infection causes most cervical intraepithelial neoplasia. J Natl Cancer Inst 1993;85(12):958-64.

221. Schiffman, MH; Haley, NJ; Felton, JS; Andrews, AW; Kaslow, RA; Lancaster, WD; Kurman, RJ; Brinton, LA; Lannom, LB; Hoffmann, D. Biochemical epidemiology of cervical neoplasia: measuring cigarette smoke constituents in the cervix. Cancer Res 1987;47:3886-8.

222. Schiller, JT. Papillomavirus-like particle vaccines for cervical cancer. Mol Med Today 1999;5(5):209-15.

223. Schlecht, NF; Franco, EL; Rohan, TE; Kjaer, SK; Schiffman, MH; Moscicki, AB; Duffy, SW. Repeatability of sexual history in longitudinal studies on HPV infection and cervical neoplasia: determinants of reporting error at follow-up interviews. J Epidemiol Biostat 2001;6(5):393-407.

224. Schlecht, NF; Kulaga, S; Robitaille, J; Ferreira, S; Santos, M; Miyamura, RA; Duarte-Franco, E; Rohan, TE; Ferenczy, A; Villa, LL; Franco, EL. Persistent human papillomavirus infection as a predictor of cervical intraepithelial neoplasia. JAMA 2001;286(24):3106-14.

225. Schneider, A; Hotz, M; Gissmann, L. Increased prevalence of human papillomaviruses in the lower genital tract of pregnant women. Int J Cancer 1987;40(2):198-201.

226. Schneider, A; Zahm, DM; Greinke, C; Kirchmayr, R; Nindl, I. Different detectability of high-risk HPV in smears from incident and prevalent high-grade squamous intraepithelial lesions of the cervix. Gynecol Oncol 1997;65(3):399-404.

227. Schneider, DL; Herrero, R; Bratti, C; Greenberg, MD; Hildesheim, A; Sherman, ME; Morales, J; Hutchinson, ML; Sedlacek, TV; Lorincz, A; Mango, L; Wacholder, S; Alfaro, M; Schiffman, M. Cervicography screening for cervical cancer among 8460 women in a high- risk population. Am J Obstet Gynecol 1999;180(2 Pt 1):290-8.

228. Sherman, ME; Schiffman, M; Cox, JT. Effects of age and human papilloma viral load on colposcopy triage: data from the randomized Atypical Squamous Cells of Undetermined Significance/Low-Grade Squamous Intraepithelial Lesion Triage Study (ALTS). J Natl Cancer Inst 2002;94(2):102-7.

229. Sherman, ME; Solomon, D; Schiffman, M. Qualification of ASCUS. A comparison of equivocal LSIL and equivocal HSIL cervical cytology in the ASCUS LSIL Triage Study. Am J Clin Pathol 2001;116(3):386-94.

230. Slattery, ML; Abbott, TM; Overall, JC Jr; Robison, LM; French, TK; Jolles, C; Gardner, JW;

West, DW. Dietary vitamins A, C, and E and selenium as risk factors for cervical cancer. Epidemiology 1990;1(1):8-15.

231. Solomon, D. The 1988 Bethesda system for reporting cervical/vaginal cytologic diagnoses. Developed and approved at the National Cancer Institute Workshop, Bethesda, Maryland, USA. December, 12-13, 1988. Acta Cytol 1989;33:567-74.

232. Solomon, D; Davey, D; Kurman, R; Moriarty, A; O'Connor, D; Prey, M; Raab, S; Sherman, M; Wilbur, D; Wright, T Jr; Young, N. The 2001 Bethesda System: terminology for reporting results of cervical cytology. JAMA 2002;287(16):2114-9.

233. Solomon, D; Schiffman, M; Tarone, R. Comparison of three management strategies for patients with atypical squamous cells of undetermined significance: baseline results from a randomized trial. J Natl Cancer Inst 2001;93(4):293-9.

234. Spitzer, M; Krumholz, BA; Chernys, AE; Seltzer, V; Lightman, AR. Comparative utility of repeat Papanicolaou smears, cervicography, and colposcopy in the evaluation of atypical Papanicolaou smears. Obstet Gynecol 1987;69(5):731-5.

235. Stafl, A. Cervicography: a new method for cervical cancer detection. Am J Obstet Gynecol 1981;139(7):815-25.

236. Stern, E; Neely, PM. Dysplasia of the uterine cervix: incidence of regression, recurrence, and cancer. Cancer 1964;17:508-12.

237. Sun, CA; Liu, JF; Wu, DM; Nieh, S; Yu, CP; Chu, TY. Viral load of high-risk human papillomavirus in cervical squamous intraepithelial lesions. Int J Gynaecol Obstet 2002;76(1):41-7.

238. Sun, XW; Kuhn, L; Ellerbrock, TV; Chiasson, MA; Bush, TJ; Wright, TC Jr. Human papillomavirus infection in women infected with the human immunodeficiency virus. N Engl J Med 1997;337(19):1343-9.

239. Sun, Y; Eluf-Neto, J; Bosch, FX; Munoz, N; Walboomers, JM; Meijer, CJ; Shah, KV; Clayman, B; Viscidi, RP. Serum antibodies to human papillomavirus 16 proteins in women from Brazil with invasive cervical carcinoma. Cancer Epidemiol Biomarkers Prev 1999;8(10):935-40.

240. Swan, DC; Tucker, RA; Tortolero-Luna, G; Mitchell, MF; Wideroff, L; Unger, ER; Nisenbaum, RA; Reeves, WC; Icenogle, JP. Human papillomavirus (HPV) DNA copy number is dependent on grade of cervical disease and HPV type. J Clin Microbiol 1999;37(4):1030-4.

241. Terry, G; Ho, L; Jenkins, D; Hills, M; Singer, A; Mansell, B; Beverley, E. Definition of human papillomavirus type 16 DNA levels in low and high grade cervical lesions by a simple polymerase chain reaction technique. Arch Virol 1993;128(1-2):123-33.

242. The New Zealand Contraception and Health Study Group. An attempt to estimate the incidence of cervical dysplasia in a group of New Zealand women using contraception. Epidemiology 1995;6(2):121-6.

243. Therneau, TM; Grambsch, PM. Modelling Survival Data: Extending the Cox Model. New York: Springer-Verlag; 2000.

244. Thomas, KK; Hughes, JP; Kuypers, JM; Kiviat, NB; Lee, SK; Adam, DE; Koutsky, LA.

Concurrent and sequential acquisition of different genital human papillomavirus types. J Infect Dis 2000;182(4):1097-102.

245. Trevisan, A; Schlecht, N; Rousseau, MC; Sobrinho, J; Prado, J; Rohan, T; Franco, E; Villa, LCorrelates of viral load among HPV-infected women enrolled in a cohort study in Brazil. Proceeding of the: 19th International Papillomavirus Conference; Castão do Santinho, Florianopolis, Brazil. 2001. p. 130.

246. University of Zimbabwe/JHPIEGO Cervical Cancer Project. Visual inspection with acetic acid for cervical-cancer screening: test qualities in a primary-care setting. University of Zimbabwe/JHPIEGO Cervical Cancer Project. Lancet 1999;353(9156):869-73.

247. van den Brule, AJ; Snijders, PJ; Gordijn, RL; Bleker, OP; Meijer, CJ; Walboomers, JM. General primer-mediated polymerase chain reaction permits the detection of sequenced and still unsequenced human papillomavirus genotypes in cervical scrapes and carcinomas. Int J Cancer 1990;45(4):644-9.

248. van Duin, M; Snijders, PJ; Schrijnemakers, HF; Voorhorst, FJ; Rozendaal, L; Nobbenhuis, MA; van den Brule, AJ; Verheijen, RH; Helmerhorst, TJ; Meijer, CJ. Human papillomavirus 16 load in normal and abnormal cervical scrapes: an indicator of CIN II/III and viral clearance. Int J Cancer 2002;98(4):590-5.

249. Vermund, SH; Schiffman, MH; Goldberg, GL; Ritter, DB; Weltman, A; Burk, RD. Molecular diagnosis of genital human papillomavirus infection: comparison of two methods used to collect exfoliated cervical cells. Am J Obstet Gynecol 1989;160(2):304-8.

250. Vermund, SH; Wilson, CM; Rogers, AS; Partlow, C; Moscicki, AB. Sexually transmitted infections among HIV infected and HIV uninfected high-risk youth in the REACH study. Reaching for Excellence in Adolescent Care and Health. J Adolesc Health 2001;29(3 Suppl):49-56.

251. Villa, LL; Franco, EL. Epidemiologic correlates of cervical neoplasia and risk of human papillomavirus infection in asymptomatic women in Brazil. J Natl Cancer Inst 1989;81(5):332-40.

252. Villa, LL; Sichero, L; Rahal, P; Caballero, O; Ferenczy, A; Rohan, T; Franco, EL. Molecular variants of human papillomavirus types 16 and 18 preferentially associated with cervical neoplasia. J Gen Virol 2000;81(Pt 12):2959-68.

253. Vizcaino, AP; Moreno, V; Bosch, FX; Munoz, N; Barros-Dios, XM; Parkin, DM. International trends in the incidence of cervical cancer: I. Adenocarcinoma and adenosquamous cell carcinomas. Int J Cancer 1998;75(4):536-45.

254. Walboomers, JM; Jacobs, MV; Manos, MM; Bosch, FX; Kummer, JA; Shah, KV; Snijders, PJ; Peto, J; Meijer, CJ; Munoz, N. Human papillomavirus is a necessary cause of invasive cervical cancer worldwide. J Pathol 1999;189(1):12-9.

255. Wallin, KL; Wiklund, F; Angstrom, T; Bergman, F; Stendahl, U; Wadell, G; Hallmans, G; Dillner, J. Type-specific persistence of human papillomavirus DNA before the development of invasive cervical cancer. N Engl J Med 1999;341(22):1633-8.

256. Wang, ZH; Kjellberg, L; Abdalla, H; Wiklund, F; Eklund, C; Knekt, P; Lehtinen, M; Kallings, I; Lenner, P; Hallmans, G; Mahlck, CG; Wadell, G; Schiller, J; Dillner, J. Type specificity and significance of different isotypes of serum antibodies to human

papillomavirus capsids. J Infect Dis 2000;181(2):456-62.

257. Wank, R; Meulen, JT; Luande, J; Eberhardt, HC; Pawlita, M. Cervical intraepithelial neoplasia, cervical carcinoma, and risk for patients with HLA-DQB1*0602,*301,*0303 alleles. Lancet 1993;341(8854):1215.

258. Wank, R; Thomssen, C. High risk of squamous cell carcinoma of the cervix for women with HLA- DQw3. Nature 1991;352(6337):723-5.

259. Wei, LJ; Lin, DY; Weissfeld, L. Regression Analysis of Multivariate Incomplete Failure Time Data by Modeling Marginal Distributions. JASA 1989;84(408):1065-73.

260. Wentzensen, N; Ridder, R; Klaes, R; Vinokurova, S; Schaefer, U; Doeberitz, MK. Characterization of viral-cellular fusion transcripts in a large series of HPV16 and 18 positive anogenital lesions. Oncogene 2002;21(3):419-26.

261. Womack, SD; Chirenje, ZM; Gaffikin, L; Blumenthal, PD; McGrath, JA; Chipato, T; Ngwalle, S; Munjoma, M; Shah, KV. HPV-based cervical cancer screening in a population at high risk for HIV infection. Int J Cancer 2000;85(2):206-10.

262. Woodman, CB; Collins, S; Winter, H; Bailey, A; Ellis, J; Prior, P; Yates, M; Rollason, TP; Young, LS. Natural history of cervical human papillomavirus infection in young women: a longitudinal cohort study. Lancet 2001;357(9271):1831-6.

263. Wynder, E; Cornfield, J; Shroff, P; Doraishwami, KR. A study of environmental factors in carcinoma of the cervix. Am J Obstet Gynecol 1954;68:1016-52.

264. Xi, LF; Demers, GW; Koutsky, LA; Kiviat, NB; Kuypers, J; Watts, DH; Holmes, KK; Galloway, DA. Analysis of human papillomavirus type 16 variants indicates establishment of persistent infection. J Infect Dis 1995;172(3):747-55.

265. Ylitalo, N. Human papillomavirus and cervical carcinoma in situ: implications for future screening . Stockholm, Sweden: Karolinska Institutet ; 2000. Doctoral Dissertation.

266. Ylitalo, N; Josefsson, A; Melbye, M; Sorensen, P; Frisch, M; Andersen, PK; Sparen, P; Gustafsson, M; Magnusson, P; Ponten, J; Gyllensten, U; Adami, HO. A prospective study showing long-term infection with human papillomavirus 16 before the development of cervical carcinoma in situ. Cancer Res 2000;60(21):6027-32.

267. Ylitalo, N; Sorensen, P; Josefsson, A; Frisch, M; Sparen, P; Ponten, J; Gyllensten, U; Melbye, M; Adami, HO. Smoking and oral contraceptives as risk factors for cervical carcinoma in situ. Int J Cancer 1999;81(3):357-65.

268. Ylitalo, N; Sorensen, P; Josefsson, AM; Magnusson, PK; Andersen, PK; Ponten, J; Adami, HO; Gyllensten, UB; Melbye, M. Consistent high viral load of human papillomavirus 16 and risk of cervical carcinoma in situ: a nested case-control study. Lancet 2000;355(9222):2194-8.

269. Youkeles, L; Forsythe, AB; Stern, E. Evaluation of Papanicolaou smear and effect of sample biopsy in follow- up of cervical dysplasia. Cancer Res 1976;36(6):2080-4.

270. Zahm, DM; Nindl, I; Greinke, C; Hoyer, H; Schneider, A. Colposcopic appearance of cervical intraepithelial neoplasia is age dependent. Am J Obstet Gynecol 1998;179(5):1298-304.

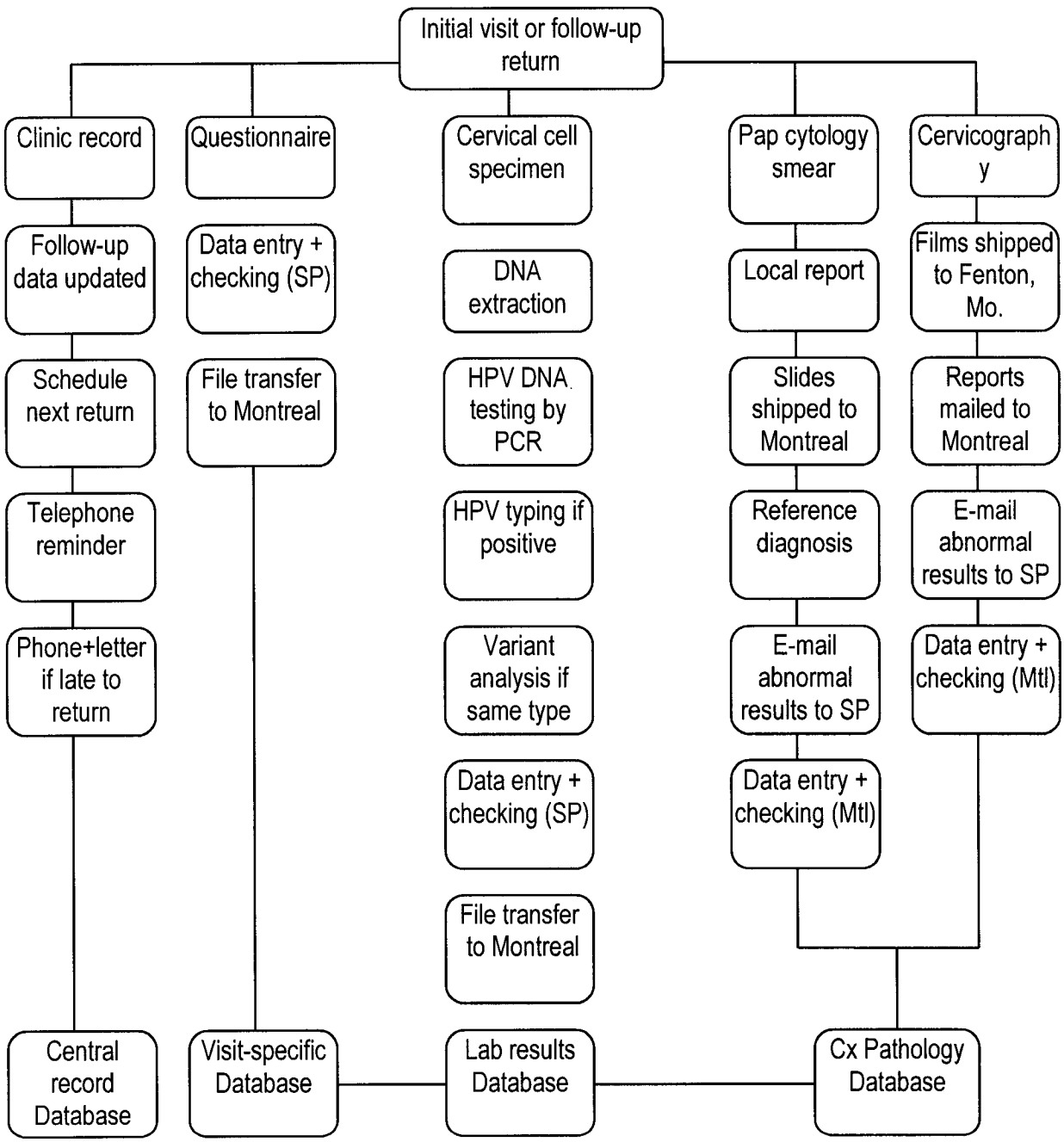
271. Zeger, SL; Liang. K. Y. Longitudinal data analysis for discrete and continuous outcomes. Biometrics 1986;42(1):121-30.

272. Ziegler, RG; Brinton, LA; Hamman, RF; Lehman, HF; Levine, RS; Mallin, K; Norman, SA; Rosenthal, JF; Trumble, AC; Hoover, RN. Diet and the risk of invasive cervical cancer among white women in the United States. Am J Epidemiol 1990;132:432-45.

273. zur Hausen, H; de Villiers, EM. Human papillomaviruses. Annu Rev Microbiol 1994;48:427-47.

10. APPENDICES

10.1. Appendix A: Ludwig-McGill study follow-up chart



10.2. Appendix B: Summary table of questionnaire data

Questionnaire and color code	Question Numbers	Information sought
1	1-18	age, ethnicity, marital status, job titles last 10 yrs, schooling, religion, income, household goods, neighborhoods of residence and where subject lived for the longest time.
(white)	19-30	tobacco consumption: frequency, type, duration, cessation.
	31-37	alcohol consumption variables: specifically type of beverage, frequency, and duration.
	38-53	age at menarche, last menstrual period, type of menstrual absorbent, gynecologic symptoms and treatments, hygiene habits, sexually-transmitted diseases, previous Pap smears.
	54-62	age at first intercourse, number of pregnancies and outcomes, sexual practices during pregnancy and after delivery.
	63-87	numbers and types of sexual partners, frequency of sexual activity during various periods in the subject's past.
	88-99	contraceptive methods: history of oral contraceptive use, frequency of condom use and other barrier methods.
	100-107	practice of anal and oral intercourse.
2	1-6	age, ethnicity, marital status.
(blue)	7-8	smoking and drinking since last visit.
	9-16	last menstrual period, type of menstrual absorbent, recent gynecologic symptoms and treatments, hygiene habits.
	17-24	age at first intercourse, lifetime and recent numbers of partners, frequency of sexual activity, contraceptive methods, practice of anal and oral intercourse since last interview.
	25-26	frequency of consumption of selected food items and vitamin supplements during the last 5 years.

* Questionnaires for enrollment and first follow-up visits used in the Ludwig-McGill Cohort Study appended at end of thesis

10.3. Appendix C: Description of statisical formulas used in analyses

10.3.1. Cox proportional hazards model

The hazard rates were calculated using the following formula for Cox proportional hazards for time-independent variables [*38*]:

$$h(t,X) = h_o(t) \times e^{\sum_{i=1}^{p} \beta i Xi}$$

where $h(t,X)$ = the hazard function at time t for an individual with the vector of explanatory variables X,

$h_0(t)$ = an arbitrary and unspecified baseline hazard function,

βi = the regression parameter associated with the i^{th} explanatory variable, and

Xi =the i^{th} explanatory variable, i=1,...,p.

The ratio of two hazards, or HR, can then be derived to compare two units of observation with a constant hazard function that depends on their covariate values. Estimation of these HR is based on a partial likelihood function that cancels out the unspecified baseline hazard function, $h_o(t)$, and accounts for right censoring of survival time [*37*]. The point estimate, βi, similarly to the logistic regression estimate, is interpreted as the hazards for subjects with a covariate Xi indicating a particular exposure status compared to subjects with the covariate Xi at its non-exposed, referent category, assuming all other covariates are held fixed. The exponentiated coefficients give the instantaneous RR for an increase of one unit for the covariate in question. By virtue of how covariates are constructed in regression models as described above, one can derive RRs from the ratio of the hazards using the following formula where the nuisance parameter $h_0(t)$ for the baseline hazard of both units of observation cancels out: The result is the exponent of an additive regression equation similar to the binomial logistic model [*128*]:

$$\frac{h(t,X^*)}{h(t,X)} = \frac{h_o(t) \times e^{\sum_{i=1}^{p} \beta i Xi^*}}{h_o(t) \times e^{\sum_{i=1}^{p} \beta i Xi}} = e^{\sum_{i=1}^{p} \beta i (Xi^* - Xi)}$$

where $\beta i(Xi)$ = the regression parameter associated with the referent dummy category, Xi = 1 for the i^{th} explanatory variable,

$\beta i(Xi^*)$ = the regression parameter associated with the dummy comparison category, Xi^* = 2, 3,...n, of the i^{th} explanatory variable, and

$h_0(t)$ = the equal arbitrary and unspecified baseline hazard functions in the numerator and denominator cancel out.

All incident cases of SIL were compared with subjects with no detected lesions or ASCUS during the entire period of follow-up.

10.3.2. *Cox proportional hazards model for time-dependent variables*

To allow for the inclusion of time-dependent covariates, the Cox proportional hazards model can be expanded using the following formula [*128*]:

$$h(t, X(t)) = h_o(t) \times e^{\sum_{i=1}^{p} \beta i Xi + \sum_{j=1}^{p} \delta j Xj(t)}$$

where $h(t,X(t))$ = the hazard function at time t for an individual with the vector of time-dependent or independent explanatory variables $X(t)$,

$h_0(t)$ = an arbitrary and unspecified non-negative baseline hazard function,

βi = regression parameter associated with the i^{th} explanatory variable,

Xi = the i^{th} explanatory variable, i=1,...,p,

δj = regression parameter associated with the j^{th} time-dependent explanatory variable, and

$Xj(t)$ =the j^{th} time-dependent explanatory variable, j=1,...,p.

The resulting HR, represented by, δj, is interpreted as the hazards for subjects with a time-dependent covariate Xj compared to subjects without the covariate Xi, at given time t, assuming all other covariates at the given time are held constant.

10.3.3. *Generalized Estimating Equation*

The computation process used in GEEs is based on the marginal or population average regression model approach, which involves a two-step process [*52*]. That is the estimates are determined first by fitting the model ignoring the correlation and then estimating the variance in a separate function. Outcomes can be continuous, binary or rates and are modeled using a generalized linear model method. This involves modeling the marginal expectation or average response, $E(Y_{ij})$, over the sub-population sharing a same exposure ($\boldsymbol{x}_{ij}$), as a function of the explanatory variables (p_{ij}). In the binary outcome circumstance $\log[E(Y_{ij})/(1- E(Y_{ij}))]$ is translated into a linear function through the use of the logit link function:

$$\text{Logit}(\mu_{ij}) = \beta. + \boldsymbol{x}_{ij}'\beta$$

The second step, the marginal variance is generalized as a function of the marginal means:

$$\text{Var}(Y_{ij}) = v(\mu ij)\phi$$

Where individuals are indexed by i, i=1,...,n, ,

observations on individuals are indexed by j, j=1,...,n_i ,

Y_{ij} is the outcome on individual i at time j ,

$Y_i = (Y_{i1}, \ldots, Y_{ini})$ is a vector of responses for subject i,

$\text{Var}(Y_{ij})$= is a matrix V_i ,

v is a known variance function, and

ϕ is a scale parameter which may need to be estimated.

Estimation is then done solving for the equation:

$$\Sigma\left(\frac{\partial \mu_i}{\partial \beta}\right) Var(Y_i)^{-1}(Y_i - \mu_i) = 0$$

The correlation between outcomes COV(Y_{ii} | Y_{ij}) is a function of the marginal means and can also be modified by additional parameters. This correlation matrix is then collapsed into a single measure p(t). Kendall's tau is used to calculate the correlation estimates in the data. GEEs model this dependence structure and the marginal distributions separately.

When carrying out GEE analysis, a link function (logit for binary outcome data, log for counts, or identity for continuous responses), a correlation structure (independent, unstructured, exchangeable or autoregressive), and a clustering variable are defined. In the case of repeated response data, clustering is defined by the subject id.

10.3.4. Marginal hazards model for multivariate failure time data

To further take into account the repeated occurrences of SIL over time, the Cox proportional hazards model for time-dependent variables can be modified to deal with the non-independent nature of repeated events within subjects. These marginal hazards models are described in detail by Therneau and Grambsch [*243*]. Based on the published models for marginal hazards the Cox models are stratified to allow the baseline hazard functions to differ for the covariate groups identified. The resulting hazard function [*h(t,X)*] at time *t* for an individual with the vector of explanatory variables *X* was described by the following formula:

$$h(t,X) = h_{oj}(t) \times e^{\sum_{i=1}^{P} \beta i Xi}$$

where $h_{0j}(t)$ = an arbitrary and unspecified baseline hazard function for each stratum (*j*) (e.g. period of observation),

βi = the regression parameter associated with the i^{th} explanatory variable (i.e. HPV status at start of period of observation), and

Xi = the i^{th} explanatory variable, i=1,...,p.

The nuisance or baseline hazard function and the regression parameters may vary with *j*. The joint distribution of all the β_j's can be assumed to be normal with a covariance matrix that can be estimated. Evaluation of exposure differences is obtained by linearly combining the β_j's. In the case of the analyses for this project (see Paper III) stratification was performed to allow the baseline hazard to vary with each period of observation defined by the index visit (t_0) where an HPV test was performed. Computationally, this is accomplished by a counting process style with each interval period being assigned to a separate stratum.

10.3.5. Cumulative survival function

The cumulative survival probability if the probability of suriving from 0 up to time t_i.[*121*] The method is based on a non-parametric approach for the calculation of the overall probability of survival in a given population. The cumulative survival function or probability, S(t), is derived for the exact times at which each failure event is measured by cytology by calculating the product of 1 – (minus) the number of k events at time *t*, over the number of subjects at risk just prior to the k^{th} failure. This function, S(*t*), is illustrated by the following formula:

$$S(t) = \prod (1 - ki / ni)$$

Censored observations at time t_i reduce the number of subjects at risk (n_i) by ki, the number of failure events, but do not change the cumulative survival probability at time t_i. The plot of the cumulative survival probabilities against time is known as the Kaplan-Meier survival curve. The cumulative distribution of outcome events over time are shown by plotting 1-S(t) over time since initiation of follow-up in the study. The median time was taken when 50% of the cases at baseline had transitioned to the outcome event.

Statistical comparison of the survival functions between exposure groups, i.e., HPV positive versus HPV negative subjects, can be performed using the log rank test. Under the null hypothesis of no difference between groups, the number of expected events (e_i) over time in each group, is defined as the sum of expected events at time t_i where $e_i=(k_i X n_i) / n_i$. Based on the Mantel-Haenszel equation, the log rank test statistic can be obtained from chi square distribution function for 1 degree of freedom, assuming a pairwise comparison of two exposure groups:

10.3.6. Mean duration

In addition to the latter actuarial estimates of mean duration of preinvasive lesions an alternate formula was also used based on the epidemiologic tenet that, within a stationary population and in the absence of migration. The average duration can be estimated with the general formula:

$$mean\ duration\ (D) = \frac{P}{[I * (1 - P)]}$$

where P = the prevalence proportion at each visit,

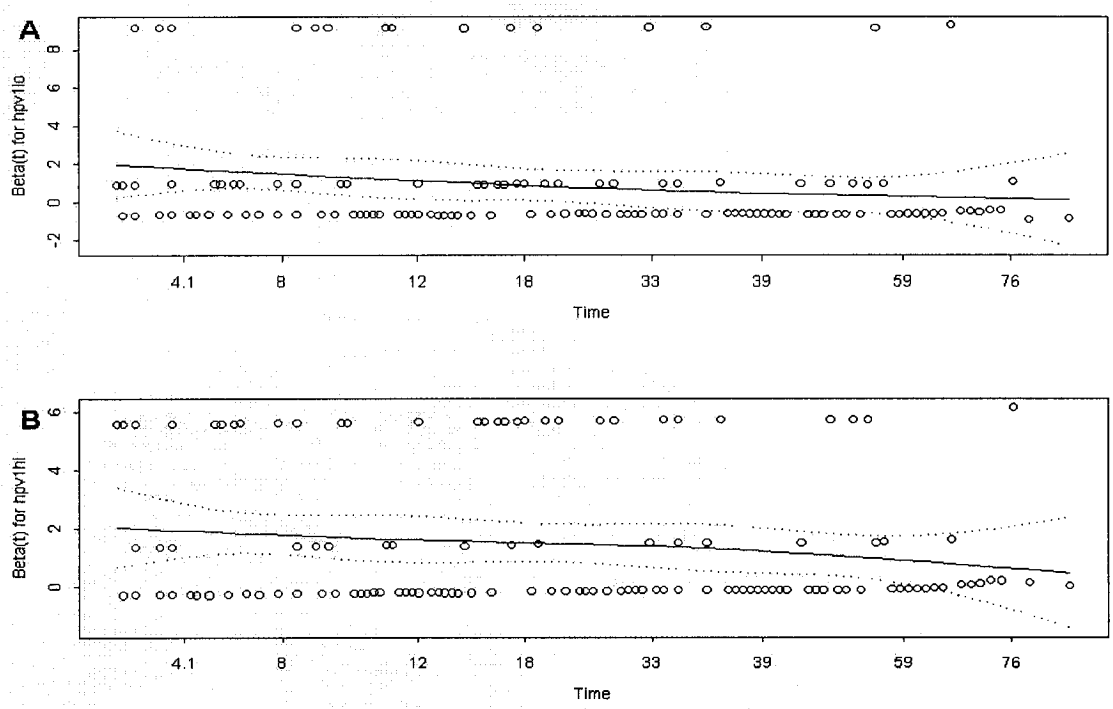
I = the incidence rate over the period of follow-up, and

D = the mean duration of the preinvasive lesion.

This formula holds provided that the point prevalence in each strata is less than 0.1, a condition that was met for all precursor lesion analyzed in the study [*76*].

10.4. Appendix D: Additional Statistical Analyses

Figure A-1: Beta estimates* and Schoenfeld residuals over time in months since enrollment to first SIL event for non-oncogenic and oncogenic HPV infection at enrollment



* Beta estimates over time in months by Cox regression analysis adjusted for age, excluding prevalent cases of SIL at enrollment. Figure A: no-oncogenic HPV at enrollment, Figure B: oncogenic HPV. Beta estimates (solid line), 95% confidence intervals (dotted line), Schoenfeld residuals (circles).

Figure A-2. Consistency of point estimates (hazard ratios) over time derived by marginal hazards regression for oncogenic HPV (top graph) and non-oncogenic HPV (bottom graph) using different exclusion criteria for prevalent lesions

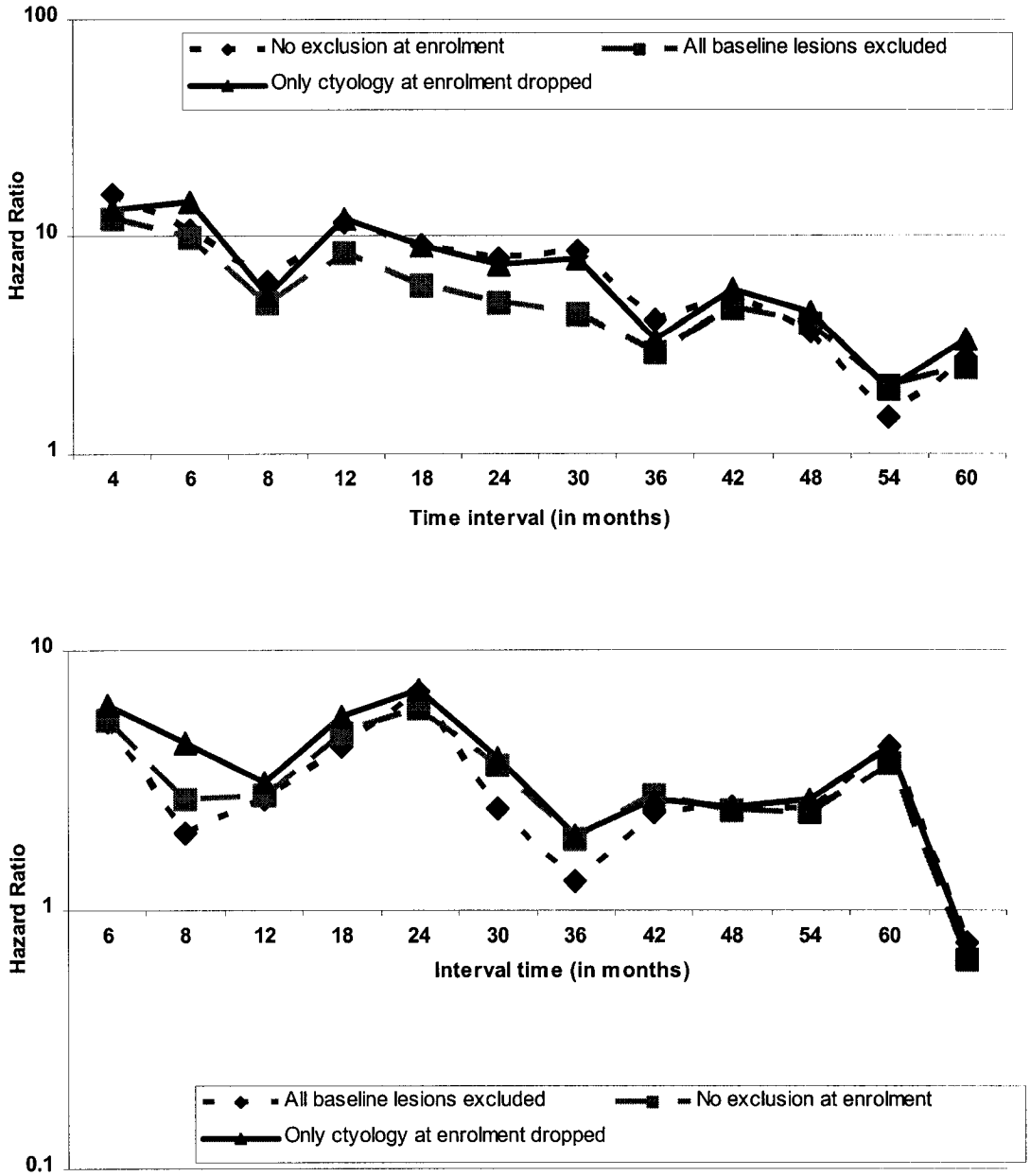


Figure A-3. Consistency of point estimates (odds ratios) over time derived by GEE regression for oncogenic HPV (top graph) and non-oncogenic HPV (bottom graph) using different exclusion criteria for prevalent lesions

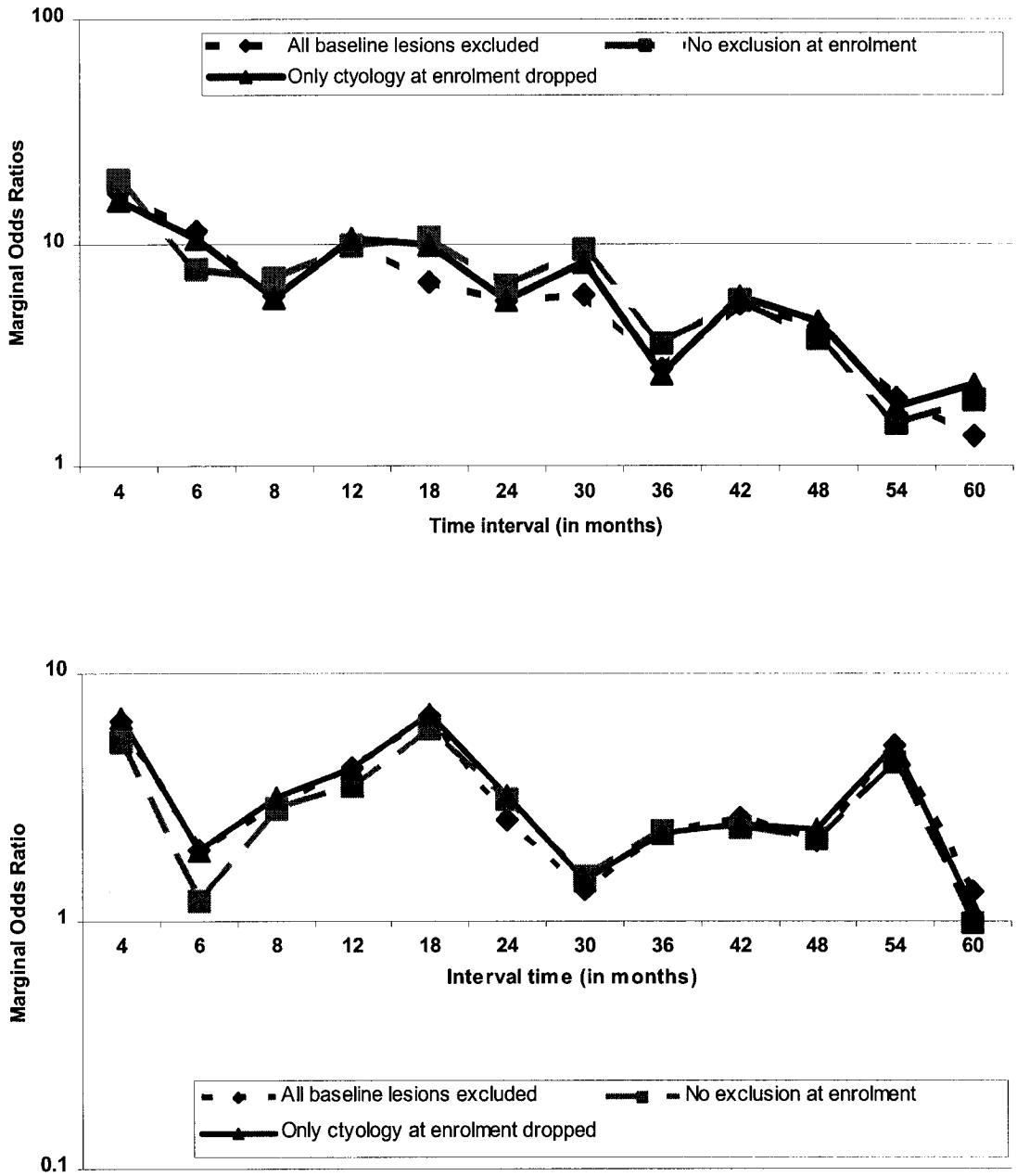


Table A-1.1: Socio-demographic and risk behavior characteristics collected at enrollment and first return visit for study participants according to compliance with scheduled visits over time

Socio-demographic characteristic	Enrollment visit only (%) (N=276)	Second or third visits only (%) (N=303)	All first year visits (%) (N=112)	First two years of visits (%) (N=115)	Two or more years of visits (%) (N=1651)	Total (%) (N=2457)*
Age at enrollment						
18-24	65 (23.6)	72 (23.8)	28 (25.0)	27 (23.5)	318 (19.3)	510 (20.8)
25-34	111 (40.2)	127 (41.9)	36 (32.1)	44 (38.3)	657 (39.8)	975 (39.7)
35-44	65 (23.6)	77 (25.4)	39 (34.5)	33 (28.7)	488 (29.6)	702 (28.6)
45-70	35 (12.70	27 (8.9)	9 (8.0)	11 (9.6)	188 (11.4)	270 (11.0)
Ethnicity						
White	186 (67.4)	185 (61.1)	79 (70.5)	65 (56.5)	1067 (64.6)	1582 (64.4)
Nonwhite	88 (31.9)	118 (38.9)	33 (29.5)	50 (43.5)	584 (35.4)	873 (35.5)
Missing	2 (0.7)					2 (0.1)
Level of education						
<Elementry	68 (24.6)	84 (27.7)	26 (23.2)	31 (27.0)	345 (20.9)	554 (22.5)
Elementry	144 (52.2)	172 (56.8)	65 (58.0)	64 (55.7)	990 (60.0)	1435 (58.4)
HighSchool	49 (17.8)	40 (13.2)	14 (12.5)	18 (15.70	275 (16.7)	396 (16.1)
College	15 (5.4)	7 (2.3)	6 (5.4)	2 (1.7)	40 (2.4)	70 (2.8)
Smoking						
Never	115 (41.7)	133 (43.9)	48 (42.9)	55 (47.8)	817 (49.5)	1168 (47.5)
Current	108 (39.1)	125 (41.3)	37 (33.0)	39 (33.9)	552 (33.4)	861 (35.0)
Former	53 (19.2)	45 (14.9)	27 (24.1)	21 (18.3)	282 (17.1)	428 (17.4)
Frequency of smoking since enrollment						
None	NA	173 (57.3)	64 (59.3)	69 (60.5)	1064 (64.5)	1370 (63.0)
<=10/day	NA	65 (21.5)	21 (19.4)	30 (26.3)	349 (21.2)	465 (21.4)
>10/day	NA	64 (21.2)	22 (20.4)	15 (13.2)	237 (14.4)	338 (15.5)
Total	276 (100.0)	303 (100)	112 (100)	115 (100)	1651 (100)	2457 (100)

* The total number of participants includes only those who met the eligibility criteria (N=2462) minus five subjects did not complete the interviews.

Table A-1.2: Reproductive health and hygiene characteristics collected at enrollment for study participants according to compliance with scheduled visits over time

Reproductive health and hygiene characteristic	Enrollment visit only (%) (N=276)	Second or third visits only (%) (N=303)	All first year visits (%) (N=112)	First two years of visits (%) (N=115)	Two or more years of visits (%) (N=1651)	Total (%) (N=2457)*
Age at menarche						
0-11	59 (21.4)	70 (23.1)	24 (21.4)	26 (22.6)	369 (22.4)	548 (22.3)
12-19	217 (78.6)	233 (76.9)	88 (78.6)	89 (77.4)	1275 (77.2)	1902 (77.4)
Use of hygienic tampons						
No	238 (86.2)	270 (89.1)	97 (86.6)	103 (89.6)	1445 (87.5)	2153 (87.6)
Yes	38 (13.8)	33 (10.9)	15 (13.4)	12 (10.4)	206 (12.5)	304 (12.4)
Use of non-commercial feminine hygiene absorbents						
No	183 (66.3)	195 (64.4)	66 (58.9)	73 (63.5)	1050 (63.6)	1567 (63.8)
Yes	93 (33.7)	108 (35.6)	46 (41.1)	42 (36.5)	601 (36.4)	890 (36.2)
Presence of genital sores						
No	241(87.3)	260 (85.8)	92 (82.1)	99 (86.1)	1403 (85.0)	2095 (85.3)
Yes	35 (12.7)	43 (14.2)	20 (17.9)	16 (13.9)	248 (15.0)	362 (14.7)
Vaginal discomfort						
No	180 (65.2)	195 (64.4)	75 (67.0)	71 (61.7)	1087 (65.8)	1608 (65.4)
Yes	96 (34.8)	108 (35.6)	37 (33.0)	44 (38.3)	564 (34.2)	849 (34.6)
Genital discomfort in pervious 2 days						
No	180 (65.2)	195 (64.4)	75 (67.0)	71 (61.7)	1087 (65.8)	1608 (65.4)
Yes	96 (34.8)	108 (35.6)	37 (33.0)	44 (38.3)	564 (34.2)	849 (34.6)
Vaginal douching						
Never/ Occasionally	250 (90.6)	269 (88.8)	102 (91.1)	105 (91.3)	1483 (89.8)	2209 (89.9)
Frequent	26 (9.4)	34 (11.2)	10 (8.9)	10 (8.7)	168 (10.2)	248 (10.1)
Oral contraceptive use						
Never	61 (22.1)	55 (18.2)	21 (18.8)	20 (17.4)	243 (14.7)	400 (16.3)
<6yrs	131 (47.5)	172 (56.8)	63 (56.3)	73 (63.5)	908 (55.0)	1347 (54.8)
>6yrs	84 (30.4)	76 (25.1)	28 (25.0)	22 (19.1)	500 (30.3)	710 (28.9)
Total number of pregnancies						
0-1	61 (22.1)	47 (15.5)	22 (19.6)	16 (13.9)	270 (16.4)	416 (16.9)
2-3	115 (41.7)	111 (36.6)	45 (40.2)	45 (39.1)	723 (43.8)	1039 (42.3)
4-6	69 (25.0)	107 (35.3)	34 (30.4)	36 (31.3)	491 (29.7)	737 (30.0)
7 or more	27 (9.8)	33 (10.9)	9 (8.0)	17 (14.8)	161 (9.8)	247 (10.1)
Missing	4 (1.4)	5 (1.7)	2 (1.8)	1 (0.9)	6 (0.4)	18 (0.7)
Total	276 (100.0)	303 (100)	112 (100)	115 (100)	1651 (100)	2457 (100)

* The total number of participants includes only those who met the eligibility criteria (N=2462) minus five subjects did not complete the interviews.

Table A-1.3: Sexual behavior characteristics collected at enrollment and first return visit for study participants according to compliance with scheduled visits over time

Sexual behavior characteristics	Enrollment visit only (%) (N=276)	Second or third visits only (%) (N=303)	All first year visits (%) (N=112)	First two years of visits (%) (N=115)	Two or more years of visits (%) (N=1651)	Total (%) (N=2457)*
Age of first intercourse						
<=15	78 (28.3)	89 (29.4)	33 (29.5)	36 (31.3)	439 (26.6)	675 (27.5)
16-17	65 (23.6)	90 (29.7)	29 (25.9)	30 (26.1)	416 (25.2)	630 (25.6)
18-19	54 (19.6)	69 (22.8)	21 (18.8)	24 (20.9)	349 (21.1)	517 (21.0)
20-50	79 (28.6)	55 (18.2)	29 (25.9)	25 (21.7)	447 (27.1)	635 (25.8)
Number of new sexual partners since enrollment						
0	NA	281 (93.0)	90 (83.3)	99 (86.8)	1527 (92.5)	1997 (91.9)
1 or more	NA	14 (4.6)	7 (6.5)	8 (7.0)	54 (3.3)	83 (3.8)
Unknown	NA	7 (2.3)	10 (11.2)	7 (6.1)	69 (4.2)	94 (4.3)
Number of recent sexual partners since enrollment						
0	NA	26 (8.6)	10 (9.3)	11 (9.6)	139 (8.4)	186 (8.6)
1	NA	269 (89.1)	94 (87.0)	95 (83.3)	1467 (88.9)	1925 (88.5)
2 or more	NA	7 (2.3)	4 (3.7)	7 (6.1)	41 (2.5)	59 (2.7)
Lifetime number of sexual partners at enrollment						
0-1	125 (45.3)	120 (39.6)	41 (36.6)	51 (44.3)	750 (45.4)	1087 (44.2)
2-3	92 (33.3)	118 (38.9)	44 (39.3)	39 (33.9)	563 (34.1)	856 (34.8)
4 or more	59 (21.4)	65 (21.5)	27 (24.1)	25 (21.7)	337 (20.4)	513 (20.9)
Number of sexual partners in the 5 years prior to enrollment						
0-1	207 (75.0)	222 (73.3)	75 (67.0)	89 (77.4)	1307 (79.2)	1900 (77.3)
2 or more	67 (24.3)	81 (26.7)	37 (33.0)	26 (22.6)	342 (20.7)	553 (22.5)
Number of sexual partners in last year prior to enrollment						
0-1	260 (94.2)	285 (94.1)	104 (92.9)	110 (95.7)	1569 (95.0)	2328 (94.7)
2 of more	9 (3.3)	12 (4.0)	6 (5.4)	3 (2.6)	71 (4.3)	101 (4.1)
Missing	7 (2.5)	6 (2.0)	2 (1.8)	2 (1.7)	11 (0.7)	28 (1.1)
Oral sex						
Never	141 (51.1)	144 (47.5)	49 (43.8)	63 (54.8)	741 (44.9)	1138 (46.3)
Ever	135 (48.9)	159 (52.5)	63 (56.3)	52 (45.2)	910 (55.1)	1319 (53.7)
Condom use						
Never/occasionally	266 (96.4)	289 (95.4)	105 (93.8)	111 (96.5)	1594 (96.5)	2365 (96.3)
Always	10 (3.6)	14 (4.6)	7 (6.3)	4 (3.5)	57 (3.5)	92 (3.7)
Anal intercourse						
Never	189 (68.5)	191 (63.0)	67 (59.8)	65 (56.5)	1030 (62.4)	1542 (62.8)
Ever	87 (31.5)	112 (37.0)	45 (40.2)	50 (43.5)	621 (37.6)	915 (37.2)
Report of sexually transmitted disease (STD) at enrollment						
No STD	219 (79.3)	234 (77.2)	78 (69.6)	80 (69.6)	1268 (76.8)	1879 (76.5)
HPV-related	14 (5.1)	12 (4.0)	9 (8.0)	7 (6.1)	64 (3.9)	106 (4.3)
Other-STD	42 (15.2)	56 (18.5)	23 (20.5)	28 (24.3)	314 (19.0)	463 (18.8)
Total	276 (100.0)	303 (100)	112 (100)	115 (100)	1651 (100)	2457 (100)

* The total number of participants includes only those who met the eligibility criteria (N=2462) minus five subjects did not complete the interviews.

10.5. Appendix E: Ethics certificates and questionnaires

McGill

Faculty of Medicine | Faculté de médecine

June 18, 2002

Dr. Eduardo Franco, Director
Division of Epidemiology
McGill University

Dear Dr. Franco:

We are writing in response to your request for continuing review for the study A07-M29-95 entitled *"Molecular Epidemiology of Persistent Cervical HPV Infection (Natural History of HPV Infection and CIN in a High Risk Area), [formerly "Molecular Epidemiology of Persistent Cervical HPV Infection"].*

The progress report was reviewed and we are pleased to inform you that full board re-approval for the study was provided on *June 17, 2002*, valid until *July 2003.* The certification of annual review has been enclosed. We note that the current consent form is dated *July 16, 1996*

We ask you to take note of the investigator's responsibility to assure that the current protocol and consent document are deposited on an annual basis with the Research Ethics Board of each hospital where patient enrollment or data collection is conducted.

Should any modification or unanticipated development occur prior to the next review, please advise the IRB promptly.

Yours sincerely,

J. Lawrence Hutchison, M.D.
Chair
Institutional Review Board

cc: A07-M29-95

McGill Faculty of Medicine
Institutional Review Board
-Continuing Review-

Principal Investigator: Dr. Eduardo L. Franco Department/Institution: Oncology / McGill University

IRB Review Number: AO7-M29-86 Study Number (if any):_____ Review Interval Annual

Title of Research Study: Molecular Epidemiology of Persistent Cervical HPV Infection
(Natural History of HPV Infection and CIN in a High Risk Area)

Date of initial IRB approval: May 6, 1996 Date of previous continuing review (if applicable) May 28,2001

INTERIM REPORT (PLEASE CHECK OR SPECIFY)

Current Status of Study:

Active Study: √ On Hold:_____ Closed to Enrolment: √

Interim Analysis: √ Final Analysis:_____ Study Not Activated*: _____

*If the study has not become active at McGill, please provide correspondence to explain; enclosed:_____

McGill hospital(s) where study has received approval of local Research Ethics Board(s) (if applicable):

JGH: ☐ MCH: ☐ MGH: ☐ MNH/MNI: ☐

RVH: ☐ SMH: ☐ Other: ☐ _____

McGill hospital(s) where study has not received approval of local Research Ethics Board(s) (if applicable): N /A

If study sponsorship or financial support has changed, please provide correspondence to explain; enclosed: N /A

Number of subjects to be enrolled at McGill: N/A Number of McGill subjects enrolled to date: N /A

Number of McGill subjects enrolled since last review: N /A Have McGill subjects withdrawn from the study?: N /A

Has the study been revised since the last review?: N /A Have the study revisions been approved by the IRB?: Yes

Has the consent form been revised since the last review?: N /A Date of the current consent form: July 16, 1996

Are there new data since the last review that could influence a subject's willingness to provide continuing consent?: NO

Have there been any serious adverse experiences (SAEs)?: NO

Have all serious adverse experiences (SAEs) and safety reports relevant to the study been reported to the IRB?: N /A

SIGNATURES:

Principal Investigato _____ Date: 23/4/02

IRB Chair:_____ Date: 18 JUN 2002

/ V

LUDWIG INSTITUTE FOR CANCER RESEARCH

SÃO PAULO

BRUSSELS

LAUSANNE

LONDON

MELBOURNE

NEW YORK

SAN DIEGO

STOCKHOLM

UPPSALA

ZURICH

JUL 10 1998

July 8, 1998

Dr. Eduardo L. Franco
Professor and Director
Division of Cancer Epidemiology

Re: Annual Renewal of IRB approval for project RO1 CA 70269 (SPA# S-13400-01 and related identifiers)

Dear Dr. Franco:

I refer to your letter of June 1 with a progress report on your study entitled "Molecular Epidemiology of Persistent HPV infection." I am pleased to inform you that your report was judged satisfactory by this institution's Ethical and Scientific Review Board. As the IRB president, I hereby extend the approval granted to your project for another year, until July 1, 1999.

Please note that this is the only notice of renewal concerning the continued ethical approval for your project. Since our IRB serves also your other study site it applies to both local Single Project Assurances granted to your project.

Sincerely yours,

Humberto Torloni, MD
President,
Ethical and Scientific Review Board

cc: Dr. LLVilla
Dr. RR Brentani
Dr. MV Renaud

INSTITUTO LUDWIG DE PESQUISA SOBRE O CÂNCER

RUA PROF. ANTONIO PRUDENTE
109-4º ANDAR
01509.010, SÃO PAULO, SP. BRASIL

University of Toronto

OFFICE OF RESEARCH SERVICES

PROTOCOL REFERENCE #3078

December 2, 1997

Dr. T. Rohan
Dept. of Public Health Sciences

Dear Dr. Rohan:

Re: Research protocol entitled, "Molecular Epidemiology of Persistent HPV Infection" (RO1 CA70269, Assurance ID # M-5001-34)

We are writing to advise you that a Review Committee composed of Drs. L. Marrett, J. McLaughlin and Professor B. Dickens has granted approval to the amendment to the above-named research study.

The approved revised addendum to the consent form is attached. Subjects should receive a copy of their consent form.

During the course of the research, any significant deviations from the approved protocol **(that is, any deviation which would lead to an increase in risk or a decrease in benefit to human subjects)** and/or any unanticipated developments within the research should be brought to the attention of the Office of Research Services.

Best wishes for the successful completion of your project.

Yours sincerely,

Susan Pilon
Executive Officer
Human Subjects Review Committee

SP/mr Enclosure
cc: Professor H. Skinner

Data da entrevista: D: _____ M: _____ A: _____ Hora do início: _____

1. Nº no estudo: _____

2. Registro M.E.V.N.C: _____

3. Qual o seu nome? _____

4. Em que dia, mês e ano a Sra. nasceu? D: _____ M: _____ A: _____

   Portanto, quantos anos a Sra. tem? _____ anos

5. Grupo étnico (*Interpretação da entrevistadora*):
   ¦1¦ Branca ¦2¦ Mulata ¦3¦ Negra ¦4¦ Amarela ¦5¦ Índio ou descendente

6. Qual é o seu estado civil?
   ¦1¦ Solteira ¦2¦ Casada ¦3¦ Viúva
   ¦4¦ Separada ¦5¦ Vive maritalmente ("juntada" ou amigada)

7. Quais foram as suas ocupações/empregos nos últimos dez anos?

   _____

   _____

   _____

   _____

8. Até que grau escolar a Sra. estudou?
   ¦1¦ Analfabeta ¦2¦ Primário incompleto ¦3¦ Primário completo
   ¦4¦ Secundário incompleto ¦5¦ Secundário completo
   ¦6¦ Técnico-profissionalizante ¦7¦ Superior

9. Qual a sua religião?
   ¦1¦ Católica ¦2¦ Crente ¦3¦ Protestante ¦4¦ Judia (Israelita)

   ¦5¦ Espírita ¦6¦ Umbandista ¦7¦ outra (qual?) _____ ¦8¦ não tem

10. Incluindo a Sra., quantas pessoas vivem na sua casa? _____ pessoas

11. Qual é a sua renda familiar, ou seja, a da Sra. mais a dos que vivem em sua casa?

    CR$ _____ cruzeiros reais

12. Quais dos seguintes itens a Sra. tem em casa?

a) geladeira	¦1¦ Sim	¦2¦ Não
b) T.V. a cores	¦1¦ Sim	¦2¦ Não
c) telefone	¦1¦ Sim	¦2¦ Não
d) video-cassete	¦1¦ Sim	¦2¦ Não
e) carro	¦1¦ Sim	¦2¦ Não
f) carros adicionais	¦1¦ Sim	¦2¦ Não

13. Em que bairro a Sra. mora: _____

14. Há quantos anos a Sra. mora nesse local? _____ anos

15. Onde a Sra. nasceu?: Cidade:_____ Estado _____

16. Essa cidade era: ¦1¦ área rural ¦2¦ área urbana ¦3¦ subúrbio ¦8¦ não sabe

17. Onde a Sra. morou a maior parte de sua vida?
(*após os 12 anos de idade*) Cidade: _____ Estado _____

18. Esse local era: ¦1¦ área rural ¦2¦ área urbana ¦3¦ subúrbio ¦8¦ não sabe

19. A Sra. fuma ou já fumou?
¦1¦ Sim ¦2¦ Nunca (*se nunca, vá direto à questão 31*)

Se sim, fazer as perguntas referentes a cada tipo de tabaco:

Cigarros de papel industrializados

20. Quantos cigarros a Sra. fuma/fumava em média **por dia**, aproximadamente?
¦1¦ no máximo 1 ¦2¦ de 2 a 5 ¦3¦ de 6 a 10 ¦4¦ 11 a 20
¦5¦ mais que 20 ¦6¦ mais que 40 (2 maços)

21. Que tipos de cigarro a Sra. fuma/fumava?
¦1¦ somente com filtro ¦2¦ principalmente com filtro, as vezes sem filtro
¦3¦ principalmente sem filtro, as vezes com filtro ¦4¦ somente sem filtro

22. Com que idade a Sra. começou a fumar regularmente ? _____ anos

23. (*Se ainda fuma, perguntar*) Há quantos anos a Sra. fuma? _____ anos

24. (*Se parou, perguntar*) Durante quantos anos a Sra. fumou? _____ anos

Cigarros de fumo de corda, de palha ou papel

25. Quantos cigarros de fumo de corda ou palha a Sra. fuma/fumava em média **por dia**, aproximadamente?
¦1¦ no máximo 1 ¦2¦ de 2 a 5 ¦3¦ de 6 a 10 ¦4¦ 11 a 20
¦5¦ mais que 20 ¦6¦ mais que 40 (2 maços)

26. Com que idade a Sra. começou a fumar regularmente? _____ anos

27. (*Se ainda fuma, perguntar*) Há quantos anos a Sra. fuma? _____ anos

28. (*Se parou, perguntar*) Durante quantos anos a Sra. fumou? _____ anos

29. (*Se parou, perguntar*) Há quantos anos a Sra. parou? _____ anos

30. A Sra. fuma/fumou charuto ou cachimbo? ¦1¦ Sim ¦2¦ Não

Consumo de bebidas alcoólicas

31. A Sra. costuma/costumava consumir bebidas alcoólicas, mesmo que ocasionalmente?
¦1¦ Sim ¦2¦ Nunca (*se nunca, vá direto à q. 38*)

32. A Sra. costuma/costumava beber cerveja?
¦1¦ Não/ocasionalmente ¦2¦ no máximo um copo por semana
¦3¦ de 2 a 5 por semana ¦4¦ de 6 a 10 ¦5¦ 11 a 30 ¦6¦ mais que 30

33. A Sra. costuma/costumava beber vinho?
¦1¦ Não/ocasionalmente ¦2¦ no máximo um copo por semana
¦3¦ de 2 a 5 por semana ¦4¦ de 6 a 10 ¦5¦ 11 a 30 ¦6¦ mais que 30

34. A Sra. costuma/costumava beber pinga ou cachaça?
¦1¦ Não/ocasionalmente ¦2¦ no máximo um copo por semana
¦3¦ de 2 a 5 por semana ¦4¦ de 6 a 10 ¦5¦ 11 a 30 ¦6¦ mais que 30

35. A Sra. costuma/costumava uisque, gim, vodca ou outra bebida forte?
¦1¦ Não/ocasionalmente ¦2¦ no máximo um copo por semana
¦3¦ de 2 a 5 por semana ¦4¦ de 6 a 10 ¦5¦ 11 a 30 ¦6¦ mais que 30

36. Há quantos anos a Sra. bebe essas quantidades? _____ anos
(*as referidas acima?*)

37. Durante quantos anos a Sra. bebeu? _____ anos
(*Se parou de beber, perguntar*)

Eu gostaria agora de lhe fazer algumas perguntas sobre sua vida intima. Eu entendo que este é um assunto pessoal, mas conhecer estas informações será de grande auxilio na nossa pesquisa. Eu volto a lembrar a Sra. que todas as respostas serão mantidas em total segredo. Nunca estes dados serão revelados a alguém.

38. Que idade a Sra. tinha quando menstruou pela primeira vez? _____ anos
(*se menopausada e passar a q. 11*)

39. Quando a Sra. teve a sua última menstruação? D: ___ M: ___ A: ___

¦_¦ Puérpera ¦_¦ Lactante ¦_¦ Menopausada

40. Quando está/estava menstruada, o que a Sra. usa/usava como absorvente íntimo?

a) absorvente tipo "MODESS" comercial	¦1¦ Sim	¦2¦ Não
b) absorvente interno tipo OB/Tampax	¦1¦ Sim	¦2¦ Não
c) toalhinha de pano	¦1¦ Sim	¦2¦ Não
d) outro (*qual?*_____)	¦1¦ Sim	¦2¦ Não

41. Nos últimos **cinco anos**, quantas vezes a Sra. sentiu coceira na região genital?
¦1¦ Nenhuma vez ¦2¦ Algumas vezes (1-9) ¦3¦ Muitas vezes (10+)

42. Nos últimos **cinco anos**, quantas vezes a Sra. sentiu dor/ardor na região genital?
¦1¦ Nenhuma vez ¦2¦ Algumas vezes (1-9) ¦3¦ Muitas vezes (10+)

43. Nos últimos **cinco anos**, quantas vezes a Sra. teve corrimento vaginal?
¦1¦ Nenhuma vez ¦2¦ Algumas vezes (1-9) ¦3¦ Muitas vezes (10+)

44. Já fez ou faz uso de algum produto para tratamento ginecológico?
(mostrar ou ler a lista de nomes de medicamentos impressa no verso da página anterior)

_____

_____

45. A Sra. já utilizou algum produto que não seja de farmácia para tratamento ginecológico?
¦1¦ Sim (*qual:* _____) ¦2¦ Não

46. Nos últimos **dois dias**, a Sra. teve corrimento, coceira ou ardor na região genital? ¦1¦ Sim ¦2¦ Não

47a. A Sra. usa/já usou algum sistema que force água/líquidos para o interior da vagina tais como duchas, bidês, etc?
¦1¦ sim, sempre ¦2¦ sim, frequentemente
¦3¦ de vez em quando ¦4¦ nunca

47b. (*Se sim*) Com que produto? _____

48. Durante os períodos menstruais a Sra. costuma/costumava lavar seus órgãos genitais? (*Além do banho diário*)
¦1¦ Não ¦2¦ Sim, uma vez por dia ¦3¦ Sim, mais de uma vez por dia

49. A Sra. já teve alguma vez feridas na vagina ou vulva? ¦1¦ Sim ¦2¦ Não

50. Alguma vez a Sra. soube por um(a) médico(a) que tinha uma doença venérea ou sexualmente transmissivel? (*se sim, qual?*)
¦1¦ Gonorreia ¦2¦ Cancro ¦3¦ Crista de galo ou condiloma ¦4¦ Sifilis
¦5¦ Herpes ¦6¦ Tricomoniase ¦7¦ Candidiase 8 Nunca teve

51. A Sra. já fez um exame de prevenção do câncer de colo de útero, também chamado Papanicolaou ou citológico ou citologia oncótica? ¦1¦ Sim ¦2¦ Não

52. (*Se sim*) Quantas vezes? _____ vezes

53. (*Se sim*) Quando foi a última vez que fez este exame?
¦1¦ No último ano ¦2¦ Há mais de um ano, mas menos que cinco
¦3¦ Há mais que cinco anos ¦8¦ Não lembra

Lembrete: Foro intimo e confidencialidade

54. Com que idade a Sra. teve a sua primeira relação sexual? _____ anos
(Se virgem, vá direto a q. 100, depois de certificar-se que ela nunca engravidou)

55. Quantas vezes a Sra. já engravidou? _____ vezes
(Se nunca, vá direto a q. 63)

56. Quantas destas gestações resultaram em partos normais? _____

57. Quantas foram por operação cesariana? _____

58. Quantas resultaram em aborto? _____

59. Em que ano foi a sua última gravidez? _____

60. Foi uma gestação completa? ¦1¦ Sim ¦2¦ Não

61. Enquanto grávida a Sra. continuava tendo relações sexuais com seu marido/parceiro? ¦1¦ Sim ¦2¦ Não

62. A Sra. costuma/costumava resguardar-se de relações sexuais após cada parto? ¦1¦ Sim ¦2¦ Não

63. Com que idade a Sra. começou a ter relações sexuais pelo menos uma vez por mês

Aos _____ anos -> ¦_¦ se nunca foi constante

64. **Durante a sua vida inteira**, com quantos homens a Sra. manteve relações sexuais?

(*Insista para que ela dê uma resposta mesmo q. aproximada*) _____

65. Quantos destes parceiros foram regulares, isto é, com os quais a Sra. teve relações sexuais regulares durante um periodo minimo de 6 meses, independentemente de morar na mesma casa?

_____

66. Que a Sra. saiba, quantos destes parceiros sexuais não foram fiéis, isto é, tiveram contato sexual com outras mulheres?

_____

67. No total, quantos parceiros sexuais a Sra. teve **antes de 20 anos**? _____

68. Quantos destes parceiros (*antes dos 20*) tinham menos que 20 anos? _____

69. Quantos destes parceiros (*antes dos 20*) tinham mais que 30 anos? _____

Se a paciente tiver menos que 20 anos, passar para a q. 73

70. No total, quantos parceiros sexuais a Sra. teve **depois dos 20 anos**? _____

71. Quantos destes parceiros (*depois dos 20*) tinham menos que 20 anos? _____

72. Quantos destes parceiros (*depois dos 20*) tinham mais que 30 anos? _____

73. Desde o inicio de sua vida sexual houve periodos em que a Sra. não teve relações por mais que um ano? Se sim, quantos periodos *(total em anos)*?

_____ anos

74. Em geral, considerando a maior parte de sua vida sexual, com que frequência a Sra. tem mantido/manteve relações sexuais? *(Descrever a frequência e duração para cada período lembrado pela paciente)*

_____

_____

_____

75. Enquanto menstruada, a Sra. evita/evitava ter relações sexuais com seu marido/parceiro? (*Também considerar a hipótese do marido evitar)*
¦1¦ Sempre evitei ou marido evitou ¦2¦ De vez em quando
¦3¦ Só nos primeiros dias ¦4¦ Nunca

76. A Sra. costuma/costumava lavar seus genitais antes de ter relações sexuais?
¦1¦ Sempre ¦2¦ de vez em quando ¦3¦ Nunca

77. A Sra. costuma/costumava lavar seus genitais depois das relações sexuais?
¦1¦ Sempre ¦2¦ de vez em quando ¦3¦ Nunca

78. Durante os **últimos cinco anos**, com quantos homens a Sra. manteve relações sexuais? (*Insista para que ela dê uma resposta mesmo que aproximada)*

_____

79. Que a Sra. saiba, quantos destes parceiros sexuais não foram fiéis, isto é, tiveram contato sexual com outras mulheres?

_____

80. Quantos destes parceiros tinham menos que 20 anos?

_____

81. Quantos destes parceiros tinham mais que 30 anos?

_____

82. Durante este periodo dos **últimos 5 anos**, com que frequência a Sra. teve relações sexuais com seu marido ou parceiro(s)?

_____/semana _____/mes _____/ano

Dizer à paciente para se lembrar apenas dos últimos 12 meses

83. Durante os **últimos 12 meses**, com quantos homens a Sra. teve relações sexuais? (*Insista para que ela dê uma resposta mesmo que aproximada*)

_____

84. Que a Sra. saiba, quantos destes parceiros sexuais não foram fiéis, isto é, tiveram contato sexual com outras mulheres?

_____

85. Quantos destes parceiros tinham menos que 20 anos?

_____

86. Quantos destes parceiros tinham mais que 30 anos?

_____

87. Durante este periodo dos **últimos 12 meses**, com que frequência a Sra. teve relações sexuais com seu marido ou parceiro(s)?

_____/semana _____/mes _____/ano

Fim das perguntas ciclicas

Detalhes sobre os métodos anticoncepcionais usados

88. Quais são os métodos que a Sra. ou o seu marido/parceiro(s) tem usado/usaram para evitar filhos? (*assinalar todos os mencionados)*
¦1¦ pílula anticoncepcional ¦2¦ laqueadura ¦3¦ vasectomia
¦4¦ D.I.U. ¦5¦ condom ¦6¦ diafragma
¦7¦ Geléia espermicida ¦8¦ Coito interrompido/tabelinha/muco cervical
¦9¦ Outro: _____ ¦10¦ Não sabe

89. *(Se P.A., perguntar)* Com que idade a Sra. começou a usar P.A.? _____ anos

90. *(Se P.A.)* Durante quantos anos a Sra. tem tomado/tomou P.A.? _____ anos

91. *(Se P.A.)* Durante este(s) período(s) a Sra. obedeceu os intervalos regulares de descanso recomendados pelo médico? ¦1¦ Sim ¦2¦ Não

92. *(Se parou, perguntar)* Há quanto tempo parou de tomar P.A.? _____ anos

93. *(Se laqueadura, perguntar)*
Há quanto tempo foi a laqueadura que a Sra. fez? _____ anos

94. *(Se vasectomia do parceiro mais frequente, perguntar)*
Há quanto tempo foi a vasectomia do seu marido/parceiro? _____ anos

95. *(Se D.I.U., perguntar)*
Com que idade a Sra. usou D.I.U. pela primeira vez? _____ anos

96. *(Se D.I.U.)* A Sra. ainda usa D.I.U.? ¦1¦ Sim ¦2¦ Não

97. *(Se condom, perguntar)* Com que frequência seu marido/parceiro(s) usa(m) camisinha? ¦1¦ Muito raramente ¦2¦ As vezes ¦3¦ Sempre

98. *(Se diafragma, perguntar)* com que frequência a Sra. tem utilizado/utilizou diafragma? ¦1¦ Muito raramente ¦2¦ As vezes ¦3¦ Sempre

99. *(Se geléia, perguntar)* A Sra. tem usado a geléia espermicida de que maneira?
¦1¦ Principalmente como método único ¦2¦ Principal/ associado ao diafragma
¦3¦ Principalmente associado à camisinha

Lembrete: Foro intimo e confidencialidade

100. A Sra. já praticou/pratica coito anal, isto é, relação com penetração pelo anus?
¦1¦ Sim, frequentemente ¦2¦ Sim, raramente ¦3¦ Não

Se não, vá direto à questão 105

101. (*Se sim, perguntar*)
Com quantos parceiros a Sra. já praticou/pratica coito anal? _____

Se mais de 1 na resposta anterior, iniciar as próximas perguntas enfatizando que a entrevistada deve se referir ao parceiro com quem ela mais frequentemente praticou coito anal.

102. O seu marido/parceiro realizava/realiza penetração vaginal em seguida ao coito anal? ¦1¦ Sim ¦2¦ Não ¦3¦ As vezes

103. (*Se sim, perguntar*) Antes da penetração vaginal o seu marido/parceiro fazia/faz a higiene do pênis? ¦1¦ Sim ¦2¦ Não ¦3¦ As vezes

104. (*Alternativamente*) Se o seu marido/parceiro usava/usa camisinha para o coito anal, antes da penetração vaginal ele a retirava ou trocava/retira ou troca?
¦1¦ Sim ¦2¦ Não ¦3¦ As vezes
¦4¦ Não usava/usa camisinha

105. O seu marido/parceiro tinha/tem o hábito de praticar sexo oral na Sra., ou seja, contato da boca ou lingua dele nos seus genitais?
¦1¦ Sim, frequentemente ¦2¦ Sim, raramente ¦3¦ Não

106. (*Se sim, perguntar*) Com quantos parceiros a Sra. já praticou/pratica sexo oral desta maneira?

_____

Se não, vá direto ao final

Se mais de 1 na resposta anterior, iniciar as próximas perguntas enfatizando que a entrevistada deve se referir ao parceiro com quem ela mais frequentemente praticou/pratica sexo oral.

107. (*Se sim, perguntar*) O seu marido/parceiro realizava/realiza penetração vaginal em seguida ao sexo oral?

¦1¦ Sim ¦2¦ Não ¦3¦ As vezes

Eu agradeco muito a sua colaboração com a nossa pesquisa. Se a Sra. tiver alguma pergunta, sinta-se a vontade em fazê-la. Caso queira comunicar-se comigo depois a Sra pode me procurar aqui durante a semana.

Horário de término da entrevista: _____:_____

COMENTARIOS DA ENTREVISTADORA:

_____

_____

_____

_____

_____

Enfermeira: _____

Questionário codificado em _____/_____/_____ *por* _____

Dados digitados em _____/_____/_____ *por* _____

Dados conferidos em _____/_____/_____ *por* _____

Data da entrevista: D: ____ M: ____ A: ____ Hora do início: _____

1. Nº no estudo: _____
2. Registro M.E.V.N.C: _____

Esta é nossa segunda entrevista, mas levaremos menos tempo desta vez, porque muitas das informações já são conhecidas. Como antes, não há respostas corretas ou incorretas, portanto a Sra. não precisa se preocupar com o que diz, o mais importante é a sua sinceridade e um esforço de memória. Lembro à senhora que todos os dados fornecidos são mantidos em total segredo. Eles jamais serão revelados à ninguém. Algumas questões podem parecer que são repeticões da nossa entrevista anterior. Elas são importantes para se conhecer mudanças que podem ter ocorrido nos últimos quatro meses, isto é, desde a nossa última conversa. Nós agradecemos a sua compreensão.

3. Só para confirmar, qual o seu nome? _____
4. Em que dia, mês e ano a Sra. nasceu? D: _____ M: _____ A: _____
5. Grupo étnico (Interpretacão da entrevistadora):
   ¦1¦ Branca ¦2¦ Mulata ¦3¦ Negra ¦4¦ Amarela ¦5¦ Indio ou descendente
6. Qual é o seu estado civil atual?
   ¦1¦ Solteira ¦2¦ Casada ¦3¦ Viuva ¦4¦ Separada ¦5¦ Vive maritalmente

Perguntas sobre fumo e álcool

7. Desde sua última visita, a Sra. fumou cigarros de papel ou outras formas de tabaco comercializado? Se sim, com que frequência?
   ¦1¦ no máximo 1 por dia ¦2¦ de 2 a 5 ¦3¦ de 6 a 10
   ¦4¦ 11 a 20 ¦5¦ mais que 20 ¦6¦ mais que 40 (2 macos)
   ¦7¦ não fumou no período
8. Também desde sua última visita, a Sra. consumiu bebidas alcoólicas? Se sim, com que frequência?
   ¦1¦ Não ¦2¦ no máximo um copo/dose por semana ¦3¦ de 2 a 5 por semana
   ¦4¦ de 6 a 10 ¦4¦ 11 a 30 ¦5¦ mais que 30

Eu gostaria agora de lhe fazer algumas perguntas sobre sua vida íntima. Eu entendo que este é um assunto pessoal, mas conhecer estas informacões será de grande auxílio na nossa pesquisa. Como da última vez, eu lembro a Sra. que todas as respostas serão mantidas em total segredo. Nunca estes dados serão revelados a alguém. Desta vez, esta parte da entrevista não será tão longa!

Usar a forma apropriada, dependendo do caso, por ex.: marido, namorado, os respectivos plurais)

9. Quando foi a sua última menstruacão? D: _____ M: _____ A: _____

¦_¦ Lactante ¦_¦ Menopausada *(se menopausada, passar à q. 11)*

10. Desde sua última visita, o que a Sra usou quando menstruada, como absorvente íntimo?

a) absorvente tipo "MODESS" comercial	¦1¦ Sim	¦2¦ Não
b) absorvente interno tipo OB/Tampax	¦1¦ Sim	¦2¦ Não
c) toalhinha de pano	¦1¦ Sim	¦2¦ Não
d) outro (qual? _____		

11. Desde sua última visita, com que frequencia a Sra. sentiu coceira na região genital?
¦1¦ Nenhuma vez ¦2¦ Algumas vezes (1-9) ¦3¦ Muitas vezes (10+)

12. Desde sua última visita, com que frequencia a Sra. sentiu dor/ardor na região genital?
¦1¦ Nenhuma vez ¦2¦ Algumas vezes (1-9) ¦3¦ Muitas vezes (10+)

13. Desde sua última visita, com que frequência a Sra. teve corrimento vaginal?
¦1¦ Nenhuma vez ¦2¦ Algumas vezes (1-9) ¦3¦ Muitas vezes (10+)

14. Desde sua última visita, a Sra. usou algum produto para tratamento ginecológico? Gostaríamos de saber também se alguns desses produtos não foram adquiridos em farmácia *(mostrar ou ler a lista de nomes de medicamentos).*

_____

15. Nos últimos dois dias, a Sra. teve corrimento, coceira ou ardor na região genital?
¦1¦ Sim ¦2¦ Não

16. Desde sua última visita, a Sra. usou algum sistema que force água/líquidos para o interior da vagina tais como duchas, bides, etc?
¦1¦ sim, sempre ¦2¦ sim, frequentemente
¦3¦ de vez em quando ¦4¦ raramente ¦5¦ nunca

Lembrete: Foro íntimo e confidencialidade

17. Com que idade a Sra. teve a sua primeira relação sexual? _____ anos
(Se virgem, vá direto à q. 21)

18. Durante a sua vida inteira, com quantos homens a Sra. manteve relações sexuais?
(Insista para que ela dê uma resposta mesmo q. aproximada)

_____

19. Desde sua última visita, com quantos homens a Sra. teve relações sexuais?
(Insista numa resposta mesmo que aproximada)

_____

20. Nestes últimos 4 meses, ou seja, desde a sua última visita, a Sra. teve relações sexuais com algum parceiro novo?
Se sim, quantos? _____

21. Desde sua última visita, com que frequência, em média, a Sra. teve relações sexuais com seu marido ou parceiro(s)?

_____/semana ou _____/mês ou _____/no período

22. Quais são os métodos que a Sra. ou o seu marido/parceiro(s) usaram nestes 4 meses para evitar filhos? (assinalar todos os mencionados)

¦1¦ pílula anticoncepcional ¦2¦ laqueadura ¦3¦ vasectomia
¦4¦ D.I.U. ¦5¦ condom ¦6¦ diafragma
¦7¦ Geléia espermicida ¦8¦ Coito interrompido/tabelinha/muco cervical
¦9¦ Outro:_____ ¦10¦ Não sabe

23. Desde a última visita, a Sra. praticou coito anal, isto é, relação com penetração pelo ânus?

¦1¦ Sim, algumas vezes ¦2¦ Sim, pelo menos uma vez ¦3¦ Não

24. Desde a última visita, o seu marido/parceiro praticou sexo oral na Sra., ou seja, contato da boca ou língua dele nos seus genitais?

¦1¦ Sim ¦2¦ Não

Agora falaremos mais do passado e não mais dos hábitos recentes. Eu gostaria agora de lhe fazer algumas perguntas sobre seus hábitos alimentares, particularmente nos últimos 5 anos. Eu compreendo que lembrar-se de fatos como consumo de alimentos no passado é difícil. Procure fazer um esforço de memória, escolhendo as alternativas apresentadas que são mais próximas da realidade. Uma resposta aproximada é sempre melhor do que nenhuma. Talvez se a Sra. se lembrar de algum fato de sua vida acontecido há mais ou menos 5 anos, será mais fácil precisar na memória as outras coisas.

25. Em média, durante os últimos 5 anos, com que frequência semanal ou mensal a Sra. tem consumido os alimentos que eu vou citar a seguir:

A.Laranjas: a própria fruta ou em forma de sucos:

¦1¦ Nunca ou menos que 1 vez/ano
¦2¦ Menos que 1 vez/mês, mais que 1/ano
¦3¦ 1 vez/mês, em media
¦4¦ 2-3 vezes/mes
¦5¦ 1-3 vezes/semana
¦6¦ 4-6 vezes/semana
¦7¦ Diariamente ou mais frequentemente

B. Limão, em forma de sucos ou limonada:

¦1¦ Nunca ou menos que 1 vez/ano
¦2¦ Menos que 1 vez/mês, mais que 1/ano
¦3¦ 1 vez/mês, em media
¦4¦ 2-3 vezes/mes
¦5¦ 1-3 vezes/semana
¦6¦ 4-6 vezes/semana
¦7¦ Diariamente ou mais frequentemente

C. Cenoura, qualquer que seja a forma de preparo:

¦1¦ Nunca ou menos que 1 vez/ano
¦2¦ Menos que 1 vez/mês, mais que 1/ano
¦3¦ 1 vez/mês, em media
¦4¦ 2-3 vezes/mes
¦5¦ 1-3 vezes/semana
¦6¦ 4-6 vezes/semana
¦7¦ Diariamente ou mais frequentemente

D. Abóbora, qualquer que seja a forma de preparo:

¦1¦ Nunca ou menos que 1 vez/ano
¦2¦ Menos que 1 vez/mês, mais que 1/ano
¦3¦ 1 vez/mês, em media
¦4¦ 2-3 vezes/mes
¦5¦ 1-3 vezes/semana
¦6¦ 4-6 vezes/semana
¦7¦ Diariamente ou mais frequentemente

E. Mamão:
¦1¦ Nunca ou menos que 1 vez/ano
¦2¦ Menos que 1 vez/mês, mais que 1/ano
¦3¦ 1 vez/mês, em media
¦4¦ 2-3 vezes/mes
¦5¦ 1-3 vezes/semana
¦6¦ 4-6 vezes/semana
¦7¦ Diariamente ou mais frequentemente

F. Couve, qualquer que seja a forma de preparo:
¦1¦ Nunca ou menos que 1 vez/ano
¦2¦ Menos que 1 vez/mês, mais que 1/ano
¦3¦ 1 vez/mês, em media
¦4¦ 2-3 vezes/mes
¦5¦ 1-3 vezes/semana
¦6¦ 4-6 vezes/semana
¦7¦ Diariamente ou mais frequentemente

G. Espinafre, qualquer que seja a forma de preparo:
¦1¦ Nunca ou menos que 1 vez/ano
¦2¦ Menos que 1 vez/mês, mais que 1/ano
¦3¦ 1 vez/mês, em media
¦4¦ 2-3 vezes/mes
¦5¦ 1-3 vezes/semana
¦6¦ 4-6 vezes/semana
¦7¦ Diariamente ou mais frequentemente

H. Brócoli, qualquer que seja a forma de preparo:
¦1¦ Nunca ou menos que 1 vez/ano
¦2¦ Menos que 1 vez/mês, mais que 1/ano
¦3¦ 1 vez/mês, em media
¦4¦ 2-3 vezes/mes
¦5¦ 1-3 vezes/semana
¦6¦ 4-6 vezes/semana
¦7¦ Diariamente ou mais frequentemente

I. Alface:
¦1¦ Nunca ou menos que 1 vez/ano
¦2¦ Menos que 1 vez/mês, mais que 1/ano
¦3¦ 1 vez/mês, em media
¦4¦ 2-3 vezes/mes
¦5¦ 1-3 vezes/semana
¦6¦ 4-6 vezes/semana
¦7¦ Diariamente ou mais frequentemente

O. Outras verduras (acelga,agrião,rucula,escarola)

¦1¦ Nunca ou menos que 1 vez/ano
¦2¦ Menos que 1 vez/mês, mais que 1/ano
¦3¦ 1 vez/mês, em media
¦4¦ 2-3 vezes/mes
¦5¦ 1-3 vezes/semana
¦6¦ 4-6 vezes/semana
¦7¦ Diariamente ou mais frequentemente

J. Ovos, incl. alimentos c/o bolos e massas:
¦1¦ Nunca ou menos que 1 vez/ano
¦2¦ Menos que 1 vez/mês, mais que 1/ano
¦3¦ 1 vez/mês, em media
¦4¦ 2-3 vezes/mes
¦5¦ 1-3 vezes/semana
¦6¦ 4-6 vezes/semana
¦7¦ Diariamente ou mais frequentemente

K. Leite e iogurtes:
¦1¦ Nunca ou menos que 1 vez/ano
¦2¦ Menos que 1 vez/mês, mais que 1/ano
¦3¦ 1 vez/mês, em media
¦4¦ 2-3 vezes/mes
¦5¦ 1-3 vezes/semana
¦6¦ 4-6 vezes/semana
¦7¦ Diariamente ou mais frequentemente

L. Queijo:

¦1¦ Nunca ou menos que 1 vez/ano
¦2¦ Menos que 1 vez/mês, mais que 1/ano
¦3¦ 1 vez/mês, em media
¦4¦ 2-3 vezes/mes
¦5¦ 1-3 vezes/semana
¦6¦ 4-6 vezes/semana
¦7¦ Diariamente ou mais frequentemente

M. Manteiga:
¦1¦ Nunca ou menos que 1 vez/ano
¦2¦ Menos que 1 vez/mês, mais que 1/ano
¦3¦ 1 vez/mês, em media
¦4¦ 2-3 vezes/mes
¦5¦ 1-3 vezes/semana
¦6¦ 4-6 vezes/semana
¦7¦ Diariamente ou mais frequentemente

N. Figado:
¦1¦ Nunca ou menos que 1 vez/ano
¦2¦ Menos que 1 vez/mês, mais que 1/ano
¦3¦ 1 vez/mês, em media
¦4¦ 2-3 vezes/mes
¦5¦ 1-3 vezes/semana
¦6¦ 4-6 vezes/semana
¦7¦ Diariamente ou mais frequentemente

26. Ainda nos últimos 5 anos, em média, a Sra. tomou complexos vitamínicos ou vitaminas isoladas sob forma de comprimidos, capsulas, drágeas, ou comprimidos efervescentes (por ex.: durante períodos de gestação)? Se sim, a Sra conseguiria precisar o tipo e o número de dias no total em que os consumiu? (somando mentalmente todos os dias em que fez uso do medicamento nos últimos 5 anos).

A. Vitamina C, isoladamente:
¦1¦ Nenhuma vez ¦2¦ Algumas (< 30 dias) ¦3¦ Frequentemente (> 30 dias)

B. Vitamina E, isoladamente:
¦1¦ Nenhuma vez ¦2¦ Algumas (< 30 dias) ¦3¦ Frequentemente (> 30 dias)

C. Complexo B, isoladamente:
¦1¦ Nenhuma vez ¦2¦ Algumas (< 30 dias) ¦3¦ Frequentemente (> 30 dias)

D. Complexos multi-vitaminicos (geralmente durante gravidez):
¦1¦ Nenhuma vez ¦2¦ Algumas (< 30 dias) ¦3¦ Frequentemente (> 30 dias)

E. Outras vitaminas/indeterminado:
¦1¦ Nenhuma vez ¦2¦ Algumas (< 30 dias) ¦3¦ Frequentemente (> 30 dias)

Eu agradeço muito a sua colaboração com a nossa pesquisa. Se a Sra. tiver alguma pergunta, sinta-se à vontade em fazê-la. Conto em re-encontrá-la daqui a 4 meses, novamente.

Horário de término da entrevista: _____:_____

COMENTARIOS DA ENTREVISTADORA:

_____

_____

_____

_____

Enfermeira: _____

Questionário codificado em ____/____/____ *por* _____

Dados digitados em ____/____/____ *por* _____

Dados conferidos em ____/____/____ *por* _____

ALL THESIS STUDENTS:

IF YOU INTEND TO SUBMIT A MANUSCRIPT-BASED THESIS (where manuscripts are published, in press, accepted, submitted, or in preparation), the information on this page must be provided and signed by you, your thesis supervisor and all co-authors, and submitted along with your abstract and nominee worksheet.

LIST **ALL AUTHORS** of **ALL MANUSCRIPTS** in your thesis:

MANUSCRIPT # 1 Title: Persistent human papillomavirus infection as a predictor of cervical neoplasia
Authors: Schlecht NF, Kulaga S, Robitaille J, Ferreira S, Santos M, Miyamura RA, Duarte-Franco E, Rohan TE, Ferenczy A, Villa LL, Franco EL.

MANUSCRIPT # 2 Title: Viral load as a predictor of risk of cervical intraepithelial neoplasia
Authors: Schlecht NF, Trevisan A, Duarte-Franco E, Rohan TE, Ferenczy A, Villa LL, Franco EL.

MANUSCRIPT # 3 Title: Time dependence of the association between human papillomavirus infection and cervical cancer precursor lesions
Authors: Schlecht NF, Platt R, Negassa A, Duarte-Franco E, Rohan TE, Ferenczy A, Villa LL, Franco EL.

MANUSCRIPT # 4 Title: Role of Human Papillomavirus infection on the Rate of Progression and Regression of Precursor Lesions in Cervical Cancer
Authors: Schlecht NF, Platt R, Duarte-Franco E, Rohan TE, Ferenczy A, Villa LL, Franco EL.

This represents an accurate count of **all** authors and **all** manuscripts included in the thesis.

Signature of Student: _____ Signature of Thesis Supervisor:

Date: 16-07-2002 Signature of Thesis Co-supervisor: _____

RELEASE FORM

I give permission to be included as a co-author of manuscript(s) that are within the thesis:

1) _____ Sophie Kulaga MSc
Signature of co-author Typed name of co-author

2) _____ Juliette Robitaille BSc
Signature of co-author Typed name of co-author

3) _____ Silvaneide Ferreira BSc
Signature of co-author Typed name of co-author

4) _____ Monica Santos BSc
Signature of co-author Typed name of co-author

5) _____ Romulo A. Miyamura BSc
Signature of co-author Typed name of co-author

6) _____ Eliane Duarte-Franco MD, MPH

Signature of co-author	Typed name of co-author
7) _____ Signature of co-author	Thomas E. Rohan MD, PhD Typed name of co-author
8) _____ Signature of co-author	Alex Ferenczy MD Typed name of co-author
9) _____ Signature of co-author	Luisa L. Villa PhD Typed name of co-author
10) _____ Signature of co-author	Eduardo L. Franco PhD Typed name of co-author
11) _____ Signature of co-author	Andrea Trevisan BSc Typed name of co-author
12) _____ Signature of co-author	Robert Platt MSc, PhD Typed name of co-author
14) _____ Signature of co-author	Abdissa Negassa, PhD Typed name of co-author
15) _____ Signature of co-author	_____ Typed name of co-author
16) _____ Signature of co-author	_____ Typed name of co-author
17) _____ Signature of co-author	_____ Typed name of co-author
18) _____ Signature of co-author	_____ Typed name of co-author
19) _____ Signature of co-author	_____ Typed name of co-author
20) _____ Signature of co-author	_____ Typed name of co-author
21) _____ Signature of co-author	_____ Typed name of co-author
22) _____ Signature of co-author	_____ Typed name of co-author

7) _____ Signature of co-author	Thomas E. Rohan MD, PhD Typed name of co-author
8) _____ Signature of co-author	Alex Ferenczy MD Typed name of co-author
9) _____ Signature of co-author	Luisa L. Villa PhD Typed name of co-author
10) _____ Signature of co-author	Eduardo L. Franco PhD Typed name of co-author
11) _____ Signature of co-author	Andrea Trevisan BSc Typed name of co-author
12) _____ Signature of co-author	Robert Platt MSc, PhD Typed name of co-author
14) _____ Signature of co-author	Abdissa Negassa, PhD Typed name of co-author
15) _____ Signature of co-author	Typed name of co-author
16) _____ Signature of co-author	Typed name of co-author
17) _____ Signature of co-author	Typed name of co-author
18) _____ Signature of co-author	Typed name of co-author
19) _____ Signature of co-author	Typed name of co-author
20) _____ Signature of co-author	Typed name of co-author
21) _____ Signature of co-author	Typed name of co-author
22) _____ Signature of co-author	Typed name of co-author

7)_____ _____ Thomas E. Rohan MD, PhD _____
Signature of co-author Typed name of co-author

8)_____ _____ Alex Ferenczy MD _____
Signature of co-author Typed name of co-author

9)_____ _____ Luisa L. Villa PhD _____
Signature of co-author Typed name of co-author

10)_____ _____ Eduardo L. Franco PhD _____
Signature of co-author Typed name of co-author

11)_____ _____ Andrea Trevisan BSc _____
Signature of co-author Typed name of co-author

_____ _____ Robert Platt MSc, PhD _____
Typed name of co-author

14)_____ _____ Abdissa Negassa, PhD _____
Signature of co-author Typed name of co-author

15)_____ _____
Signature of co-author Typed name of co-author

16)_____ _____
Signature of co-author Typed name of co-author

17)_____ _____
Signature of co-author Typed name of co-author

18)_____ _____
Signature of co-author Typed name of co-author

19)_____ _____
Signature of co-author Typed name of co-author

20)_____ _____
Signature of co-author Typed name of co-author

21)_____ _____
Signature of co-author Typed name of co-author

22)_____ _____
Signature of co-author Typed name of co-author

7) _____ Thomas E. Rohan MD, PhD
Signature of co-author — Typed name of co-author

8) _____ Alex Ferenczy MD
Signature of co-author — Typed name of co-author

9) _____ Luisa L. Villa PhD
Signature of co-author — Typed name of co-author

10) _____ Eduardo L. Franco PhD
Signature of co-author — Typed name of co-author

11) _____ Andrea Trevisan BSc
Signature of co-author — Typed name of co-author

12) _____ Robert Platt MSc, PhD
Signature of co-author — Typed name of co-author

_____ Abdissa Negassa, PhD
Signature of co-author — Typed name of co-author

15) _____
Signature of co-author — Typed name of co-author

16) _____
Signature of co-author — Typed name of co-author

17) _____
Signature of co-author — Typed name of co-author

18) _____
Signature of co-author — Typed name of co-author

19) _____
Signature of co-author — Typed name of co-author

20) _____
Signature of co-author — Typed name of co-author

21) _____
Signature of co-author — Typed name of co-author

22) _____
Signature of co-author — Typed name of co-author

ALL THESIS STUDENTS:

IF YOU INTEND TO SUBMIT A MANUSCRIPT-BASED THESIS (where manuscripts are published, in press, accepted, submitted, or in preparation), the information on this page must be provided and signed by you, your thesis supervisor and all co-authors, and submitted along with your abstract and nominee worksheet.

LIST **ALL AUTHORS** of **ALL MANUSCRIPTS** in your thesis:

MANUSCRIPT # 1 Title: Persistent human papillomavirus infection as a predictor of cervical neoplasia
Authors: Schlecht NF, Kulaga S, Robitaille J, Ferreira S, Santos M, Miyamura RA, Duarte-Franco E, Rohan TE, Ferenczy A, Villa LL, Franco EL.

MANUSCRIPT # 2 Title: Viral load as a predictor of risk of cervical intraepithelial neoplasia
Authors: Schlecht NF, Trevisan A, Duarte-Franco E, Rohan TE, Ferenczy A, Villa LL, Franco EL.

MANUSCRIPT # 3 Title: Time dependence of the association between human papillomavirus infection and cervical cancer precursor lesions
Authors: Schlecht NF, Platt R, Negassa A, Duarte-Franco E, Rohan TE, Ferenczy A, Villa LL, Franco EL.

MANUSCRIPT # 4 Title: Role of Human papillomavirus infection on the rate of progression and regression of precursor lesions in cervical cancer
Authors: Schlecht NF, Platt R, Duarte-Franco E, Rohan TE, Ferenczy A, Villa LL, Franco EL.

This represents an accurate count of **all** authors and **all** manuscripts included in the thesis.

Signature of Student:_____ Signature of Thesis Supervisor:_____

Date:_____ Signature of Thesis Co-supervisor:_____

RELEASE FORM

I give permission to be included as a co-author of manuscript(s) that are within the thesis:

1)_____ _____ Sophie Kulaga MSc _____
Signature of co-author Typed name of co-author

2 _____ Juliette Robitaille BSc RT (CSLT) CT (ASCP)
Typed name of co-author

3)_____ _____ Silvaneide Ferreira BSc _____
Signature of co-author Typed name of co-author

4)_____ _____ Monica Santos BSc _____
Signature of co-author Typed name of co-author

5)_____ _____ Romulo A. Miyamura BSc _____
Signature of co-author Typed name of co-author

6)_____ _____ Eliane Duarte-Franco MD, MPH ___
Signature of co-author Typed name of co-author

7)_____ Signature o	Thomas E. Rohan MD, PhD Typed name of co-author
8)_____ Signature of co-author	Alex Ferenczy MD Typed name of co-author
9)_____ Signature of co-author	Luisa L. Villa PhD Typed name of co-author
10)_____ Signature of co-author	Eduardo L. Franco PhD Typed name of co-author
11)_____ Signature of co-author	Andrea Trevisan BSc Typed name of co-author
12)_____ Signature of co-author	Robert Platt MSc, PhD Typed name of co-author
14)_____ Signature of co-author	Abdissa Negassa, PhD Typed name of co-author
15)_____ Signature of co-author	_____ Typed name of co-author
16)_____ Signature of co-author	_____ Typed name of co-author
17)_____ Signature of co-author	_____ Typed name of co-author
18)_____ Signature of co-author	_____ Typed name of co-author
19)_____ Signature of co-author	_____ Typed name of co-author
20)_____ Signature of co-author	_____ Typed name of co-author
21)_____ Signature of co-author	_____ Typed name of co-author
22)_____ Signature of co-author	_____ Typed name of co-author

ALL THESIS STUDENTS:

IF YOU INTEND TO SUBMIT A MANUSCRIPT-BASED THESIS (where manuscripts are published, in press, accepted, submitted, or in preparation), the information on this page must be provided and signed by you, your thesis supervisor and all co-authors, and submitted along with your abstract and nominee worksheet.

LIST **ALL AUTHORS** of **ALL MANUSCRIPTS** in your thesis:

MANUSCRIPT # 1 Title: Persistent human papillomavirus infection as a predictor of cervical neoplasia
Authors: Schlecht NF, Kulaga S, Robitaille J, Ferreira S, Santos M, Miyamura RA, Duarte-Franco E, Rohan TE, Ferenczy A, Villa LL, Franco EL.

MANUSCRIPT # 2 Title: Viral load as a predictor of risk of cervical intraepithelial neoplasia
Authors: Schlecht NF, Trevisan A, Duarte-Franco E, Rohan TE, Ferenczy A, Villa LL, Franco EL.

MANUSCRIPT # 3 Title: Time dependence of the association between human papillomavirus infection and cervical cancer precursor lesions
Authors: Schlecht NF, Platt R, Negassa A, Duarte-Franco E, Rohan TE, Ferenczy A, Villa LL, Franco EL.

MANUSCRIPT # 4 Title: Role of Human papillomavirus infection on the rate of progression and regression of precursor lesions in cervical cancer
Authors: Schlecht NF, Platt R, Duarte-Franco E, Rohan TE, Ferenczy A, Villa LL, Franco EL.

This represents an accurate count of **all** authors and **all** manuscripts included in the thesis.

Signature of Student:_____ Signature of Thesis Supervisor:_____

Date:_____ Signature of Thesis Co-supervisor:_____

RELEASE FORM

I give permission to be included as a co-author of manuscript(s) that are within the thesis:

_____ Signature of co-author	_____ Sophie Kulaga MSc _____ Typed name of co-author
2)_____ Signature of co-author	_____ Juliette Robitaille BSc _____ Typed name of co-author
3)_____ Signature of co-author	_____ Silvaneide Ferreira BSc _____ Typed name of co-author
4)_____ Signature of co-author	_____ Monica Santos BSc _____ Typed name of co-author
5)_____ Signature of co-author	_____ Romulo A. Miyamura BSc _____ Typed name of co-author
6)_____ Signature of co-author	_____ Eliane Duarte-Franco MD, MPH ___ Typed name of co-author

ALL THESIS STUDENTS:

IF YOU INTEND TO SUBMIT A MANUSCRIPT-BASED THESIS (where manuscripts are published, in press, accepted, submitted, or in preparation), the information on this page must be provided and signed by you, your thesis supervisor and all co-authors, and submitted along with your abstract and nominee worksheet.

LIST **ALL AUTHORS** of **ALL MANUSCRIPTS** in your thesis:

MANUSCRIPT # 1 Title: Persistent human papillomavirus infection as a predictor of cervical neoplasia
Authors: Schlecht NF, Kulaga S, Robitaille J, Ferreira S, Santos M, Miyamura RA, Duarte-Franco E, Rohan TE, Ferenczy A, Villa LL, Franco EL.

MANUSCRIPT # 2 Title: Viral load as a predictor of risk of cervical intraepithelial neoplasia
Authors: Schlecht NF, Trevisan A, Duarte-Franco E, Rohan TE, Ferenczy A, Villa LL, Franco EL.

MANUSCRIPT # 3 Title: Time dependence of the association between human papillomavirus infection and cervical cancer precursor lesions
Authors: Schlecht NF, Platt R, Negassa A, Duarte-Franco E, Rohan TE, Ferenczy A, Villa LL, Franco EL.

MANUSCRIPT # 4 Title: Role of Human papillomavirus infection on the rate of progression and regression of precursor lesions in cervical cancer
Authors: Schlecht NF, Platt R, Duarte-Franco E, Rohan TE, Ferenczy A, Villa LL, Franco EL.

This represents an accurate count of **all** authors and **all** manuscripts included in the thesis.

Signature of Student:_____ Signature of Thesis Supervisor:_____

Date:_____ Signature of Thesis Co-supervisor:_____

RELEASE FORM

I give permission to be included as a co-author of manuscript(s) that are within the thesis:

1)_____ _____ Sophie Kulaga MSc _____
Signature of co-author Typed name of co-author

2)_____ _____ Juliette Robitaille BSc _____
Signature of co-author Typed name of co-author

_____ _____ Silvaneide Ferreira BSc _____
Typed name of co-author

_____ _____ Monica Santos BSc _____
Typed name of co-author

_____ Romulo A. Miyamura BSc _____
Signature of co-author Typed name of co-author

6)_____ _____ Eliane Duarte-Franco MD, MPH ___
Signature of co-author Typed name of co-author

7) _____ Signature of co-author	Thomas E. Rohan MD, PhD Typed name of co-author
8) _____ Signature of co-author	Alex Ferenczy MD Typed name of co-author
9) _____ Signature of co-author	Luisa L. Villa PhD Typed name of co-author
10) _____ Signature of co-author	Eduardo L. Franco PhD Typed name of co-author
11) _____ Signature of co-author	Andrea Trevisan BSc Typed name of co-author
12) _____ Signature of co-author	Robert Platt MSc, PhD Typed name of co-author
14) _____ Signature of co-author	Abdissa Negassa, PhD Typed name of co-author
15) _____ Signature of co-author	_____ Typed name of co-author
16) _____ Signature of co-author	_____ Typed name of co-author
17) _____ Signature of co-author	_____ Typed name of co-author
18) _____ Signature of co-author	_____ Typed name of co-author
19) _____ Signature of co-author	_____ Typed name of co-author
20) _____ Signature of co-author	_____ Typed name of co-author
21) _____ Signature of co-author	_____ Typed name of co-author
22) _____ Signature of co-author	_____ Typed name of co-author