THE ROLE OF INSULIN-LIKE GROWTH FACTOR-I AND INSULIN-LIKE GROWTH FACTOR BINDING PROTEIN-3 ON THE DEVELOPMENT OF CERVICAL SQUAMOUS INTRAEPITHELIAL LESIONS

Rhea Ferguson

Department of Epidemiology, Biostatistics, and Occupational Health

McGill University, Montreal

June 2010

A thesis submitted to McGill University in partial fulfilment of the requirements for the degree of Masters in Science

Abstract

Objectives: Higher levels of circulating insulin-like growth factor-I (IGF-I) and lower levels of its binding protein (IGFBP-3) have been linked to an increased risk of certain epithelial cancers. It is unclear whether IGF-1 plays a similar role in the development of cervical squamous intraepithelial lesions (SIL). I investigated the association between circulating levels IGF-1 and IGFBP-3 and development of SIL.

Methods: Blood serum samples from a nested case-control study were analyzed. Two controls were age and risk-set matched to each case. Conditional logistic regression was used for the statistical analysis.

Results: While the odds ratios of higher quartiles of circulating IGF-1 showed a higher risk of developing SIL, as compared to baseline, none of the associations were significant. The same was found for both IGFBP-3 and the molar ratio IGF-1:IGFBP-3.

Conclusions: IGF-1 and IGFBP-3 may play at most a minor role in the development of cervical SIL.

Résumé

Objectifs: Des niveaux élevés de circulation du facteur de croissance analogue à l'insuline (IGF-I) et des niveaux inférieurs de sa protéine de liaison (IGFBP-3) sont associés à un risque accru de certains cancers épithéliaux, mais leur rôle dans le développement lésions squameuses intraépithéliales cervicales (SIL) demeure incertain. L'association entre les taux circulatoires d'IGF-1 et d'IGFBP-3 et le développement de SIL a été évaluée.

Méthodes: Des échantillons de sérum sanguin d'une étude cas-témoins nichée dans une cohorte ont été analysés. Deux sujets du groupe contrôle ont été pairés quand à l'âge et certains facteurs de risque à chaque cas. L'analyse statistique a été effectuée par régression logistique conditionnelle.

Résultats: Bien que les rapports de cotes des quartiles supérieurs d'IGF-1, d'IGFBP-3 et le rapport molaire IGF-1: l'IGFBP-3 suggèrent un risque accru de développer des SIL, par rapport aux valeurs initiales, aucune des associations ne sont statistiquement significatives.

Conclusions: IGF-1 et IGFBP-3 pourraient jouer tout au plus un rôle mineur dans le développement de SIL du col de l'utérus.

Statement of Support

The Ludwig-McGill Cohort Study was funded by the Ludwig Institute for Cancer Research, the Medical Research Council (MA-13647), the National Institute of Health (R01 CA70269-01), and the Canadian Institutes of Health Research (MOP-77650).

During this study, I was funded by a Canadian Institutes of Health Research's Frederick Banting and Charles Best Canada Graduate Scholarships Master's Award. Funding for my research work was provided by the Cancer Research Society Division of Epidemiology, McGill University.

Acknowledgements

My master's thesis was completed under the supervision of Dr. Eduardo Franco, and I could not have asked for a better supervisor. I would like to thank him for all of the help and guidance he has offered me, in both my academic and personal life. He goes out of his way to support his students, and to provide them with the best learning opportunities he can, and I have learned so much during my time here. I will miss working for him.

I am also grateful to Dr. Ramana-kumar Agnihotram, without whom I would have spent a great deal of time being frustrated, staring at my computer screen with no idea what to do next. He provided invaluable support when it came to statistical analysis, and helped me learn STATA in record time. His cheerful disposition and willingness to drop everything in order to answer your question did not go unappreciated, and it was a pleasure to work with him. I will miss him very much!

I would like to sincerely thank Dr. Pollak, Dr. Coutlée, and their laboratory staff for their hard work in providing the data that was fundamental to the completion of thesis.

The members of my thesis committee were exceptional individuals and must be thanked. Dr. Salah Mahmud, Dr. Robert Platt, and Dr. Michael Pollak provided me with invaluable feedback and insight, and helped me get through the thesis writing process with only minimal mental scarring.

I was fortunate to meet and make many new friends during my time at McGill, and I was lucky enough to share my office with one of them; Candace Sirjoosingh. She is a great friend, and made the many hours at work fly by. She always provided helpful and insightful feedback to any problem, both academic and personal, and helped me laugh through all the difficult times. I wish her all the best in her future as an epidemiologist!

I must also thank my parents. Throughout my entire life they have supported me unconditionally in every way they could. I'll never be able to articulately express my gratitude, but I want them to know that I appreciate every second I have with them and love them so much. This year has been particularly hard, as my dad was diagnosed with cancer. Despite having his own life to worry about, he is always concerned about me and how the diagnosis will affect my studies, and how it is hard for me to be away at school. I have worked so hard because I know he would want nothing else, and he had better be at my graduation!

Thanks also to Candida Pizzolongo, the one that keeps this office running so efficiently. She is so kind and thoughtful, and genuinely cares about every one of us. I loved coming in every day and seeing her at her desk, she and I always had such fun conversations. I will keep in touch.

Many people helped me edit my thesis in its various stages of completion, and to them I am very grateful. Thanks to my sister, Janna, and Chris for always offering to read my work, even when I know you must have thought it was boring.

Table of Contents

List of Tables

1. Introduction

1.1 Descriptive epidemiology of invasive cervical cancer

In 2008 an estimated 529 000 women worldwide developed cervical cancer and approximately 274 000 women died from it, making cervical cancer the third most common cancer in women, and the 7th most common overall (1). More than 85% of the global disease burden occurs in developing countries, with Eastern and Western Africa having the greatest burdens (age standardized rate of greater than 30 per 100 000) (1). Other areas of high risk include South Africa, South-Central Asia, South America, and Middle Africa (1-3). The lowest rates of cervical cancer can be found in Western Asia, North America, Australia/New Zealand and Japan (1, 2). While a correlation between incidence and mortality is seen across all countries, some regions, such as Africa, have a disproportionally higher rate of cervical cancer death as compared to other countries (4). Women in developing countries, such as those in Africa, have a 50% chance of surviving more than 5 years, whereas this number is 66% in developed countries (5). As this disease affects relatively young women, it is the single biggest cause of life years lost due to cancer in the developing world, and in Latin America, the Caribbean, and Eastern Europe it is responsible for more years lost than tuberculosis, maternal mortality, or AIDS (2).

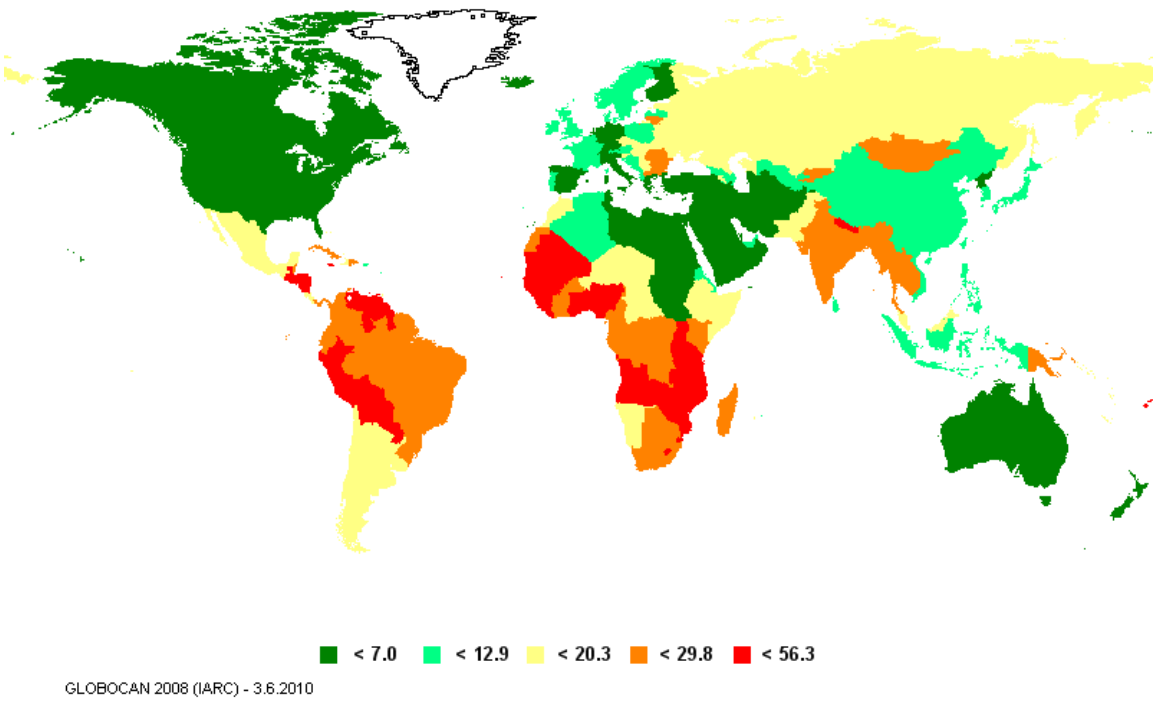

Figure 1: Estimated age-standardized incidence rate per 100,000 of cervical cancer, all ages. Figure adapted from Globocan 2008 (1).

1.2 Classification and staging of cervical precursor lesions

The main histological types of invasive cervical cancer include squamous cell carcinomas, which account for about 80% of cancers, with adenocarcinomas and mixed adenosquamous tumours making up the majority of the remaining 20% (6). Cervical cancer results when pre-invasive lesions, which have not invaded the connective tissue adjacent to the epithelial layer of the cervix, progress and breach this layer. There are two separate schemes used to classify cervical dysplasia: that of the World Health Organization (WHO), which is based on histopathology and uses the term cervical intraepithelial neoplasias (CIN), and the newer Bethesda System, which is based on cytology, and classifies lesions as squamous intraepithelial lesions (SIL). Table 1 outlines the correspondence between the different terminologies.

The system used by the Ludwig-McGill Cohort Study, the focus of this thesis, was that of the Bethesda System, first outlined by the National Cancer Institute in 1988 (7). The Bethesda System was developed to be a uniform reporting system for cervical/vaginal cytology to: 1) improve the communication between cytopathologists and the referring physicians, 2) facilitate cytologic-histopathologic correlation, 3) facilitate research into the epidemiology, biology, and pathology of cervical disease, and 4) provide reliable data for national and international statistic analyses and comparisons (7). The Bethesda System has two main terms, "low-grade squamous intraepithelial lesions" (LG-SIL) and "high-grade squamous intraepithelial lesions" (HG-SIL). The system limits the use of "atypical cells" to cases in which the findings are of undetermined significance.

Table 1: Correspondence between different terminologies

Cytomorphological Changes	WHO	Bethesda System	Proportion of Atypical Basal Layer
Normal	Normal	Within normal limits (WNL)	--
Inflammatory/Atypia (multiple qualifiers)	Normal	Benign cellular changes	--
Inflammatory/Atypia (epithelial cell abnormalities)	Normal	ASCUS/AGCUS	--
Mild dysplasia	CIN1	LG-SIL	1/3 of layer
Moderate dysplasia	CIN 2	HG-SIL	1/3-2/3 of layer
Severe dysplasia	CIN 3	HG-SIL	2/3 – whole layer
Carcinoma in situ (CIS)	CIN 3	HG-SIL	2/3 – whole layer
Invasive Cervical Cancer (ICC)	ICC	ICC	

Abbreviations: CIN, cervical intraepithelial neoplasia; ASCUS, atypical squamous cells of undetermined significance; AGCUS, atypical glandular cells of undetermined significance; LG-SIL, low grade squamous intraepithelial lesions; HG-SIL, high grade squamous intraepithelial lesion.

1.3 Risk Factors for HPV infection

Although human papillomavirus (HPV) infections are common and vital to cervical cancer development, most infected women will clear their infections within a few months to a few years after acquisition (6), never developing the disease. The absolute risk of a LG-SIL progressing to a HG-SIL is only 15-25% over 2-4 years (6). These facts indicate that while HPV infection is a necessary cause of cervical cancer, it is not sufficient. Other risk factors must play a role in the aetiology of cervical cancer, including variables related to the virus, host, and environment.

Examples of viral cofactors include the type of HPV, the viral load, and whether there is viral integration into the host genome. The host influences infections through endogenous hormones, genetic factors, and factors related to immune response. The environment can play a role as well; hormonal contraceptives, tobacco smoking, parity, and co-infections with other sexually transmitted agents can influence HPV infection (8).

The best-known risk factors for the acquisition of an HPV infection are the number of sexual partners (life-time and recent), age at first intercourse, smoking, other STIs, chronic inflammation, and immunosuppressive conditions including HIV infection (9). Markers of sexual activity are strongly associated with all types of HPV infections (9), however, the most consistent determinant of infection is age, with most studies showing a sharp decrease in risk of infection after the age of 25 or 30 (10-12). This decrease in risk is independent of the changes in sexual behaviour, suggesting this could be due to the protective effect of a specific immune response to the virus.

1.4 HPV infection and cervical cancer

Clinical, subclinical or latent human papillomavirus (HPV) infections are the most common sexually transmitted infection worldwide, in both women and men (5), with studies estimating the prevalence of HPV in asymptomatic women in the general population to be in the range of 2-44% (9). This wide range in estimated prevalence can be attributed to the age differences amongst the groups tested, as HPV prevalence is highest amongst sexually active young adults (13), and to the differences in sensitivity of various HPV DNA assays. The International Agency for Research on Cancer (IARC) has classified HPV types 16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, 59, and 66 as being carcinogenic to human beings (14), and a persistent infection with one of these high-risk types is known to be a necessary causal factor in the development of cervical cancer (15, 16). So necessary is HPV to the development of cervical cancer that there is an HPV prevalence of greater than 99.7% among women with the disease (16). Approximately 5-15% of HPV negative women in the general population will contract a high-risk HPV type each year (17), and there are an estimated 20 million Americans infected with HPV at any one time (9).

Low-risk (LR) HPV types, or types unlikely to cause cancer, are responsible for anogenital warts (condylomata acuminata) and low-grade cervical lesions, and thus it is hard to measure the occurrence of LR types in a population directly; the only method available being through hospital series or physician consultation statistics. Despite this handicap, it is thought that lesions caused by LR-HPV types will affect 1% of all sexually active adults in the United States (9).

2.0 Insulin-like growth factor and insulin-like growth factor binding protein-3

2.1 The insulin-like growth factor family of proteins

The insulin-like growth factor (IGF) family of proteins consists of several ligand hormones, receptors and binding proteins, and it has received recent attention with respect to cancer risk. The two main IGF peptides, IGF-I and IGF-II, are structurally similar to insulin (50% homology to pro-insulin), and the IGF-1 receptor (IGFR-I) and the human insulin receptor show 84% homology in the tyrosine kinase domains (18). Despite this homology, IGFs have insulin-like activity that is not inhibited by anti-insulin antibodies. The prime physiological role of the IGF family of proteins is to regulate cellular proliferation and apoptosis in relation to diet, but it also plays a role in energy metabolism, body size, longevity, and other organ-specific functions (18, 19).

IGF-1 is a peptide that can act as both a circulating hormone and as a local growth factor, depending on where it is produced. Most circulating IGF-I is produced in the liver so it must travel, via the blood stream, to the tissue or organ upon which it exerts its effects. Circulating IGFs are carried by serum binding proteins that are also synthesized mainly in the liver, the primary one of which is IGF-binding protein-3 (IGFBP-3). These binding proteins, to which more than 90% of serum IGF-I is bound (20), increase the half-life of circulating IGF-I from approximately 11 minutes to 12-15 hours (18). Some tissues can produce their own IGF-I, allowing it to act through autocrine or paracrine mechanisms of action. This might affect how well systematic concentrations of IGF-I can predict IGF signalling. Hepatic production of IGF-I is regulated by growth hormone (GH) that is produced in the pituitary gland, which can decrease IGF-I levels during periods of starvation or malnutrition.

IGF-I is a ligand to IGFR-I, a cell-surface tyrosine kinase receptor that binds IGF-I with a very high affinity. Once activated, the receptor initiates signalling pathways that favour cell proliferation and survival. IGFBP-3 influences the bioavailability of IGF-I by competing with IGFR-I for the ligand, allowing it to play a mediating role in IGFR-I signalling.

2.2 Determinants of circulating IGF-1 serum concentrations

Circulating serum concentrations of IGF-I and IGFBP-3 vary widely between individuals and high levels have been implicated in the pathogenesis of several cancers. For these reasons it is of interest to determine which anthropometric and lifestyle factors influence IGF-I and IGFBP-3 expression.

A cross-sectional study by Kaklamani *et al.* investigated age, height, BMI, cigarette smoking (cigarettes/d), alcohol drinking (glasses/mo), and coffee drinking (cups/mo) as possible determinants of IGF-I/IGFBP-3 serum concentration and found that only age and sex are significant predictors of circulating IGF-1 levels, and that smoking and sex are the significant predictors of circulating IGFBP-3 (21). Another cross-sectional study, by Goodman-Gruen *et al.* found that IGF-1 levels were associated with age, sex, and alcohol use (22). This second study, however, used a study population of elderly men and women, and is thus not directly comparable to the first. A Swedish study also reported decreasing IGF-1 levels with increasing age, and that the protein levels differed significantly between men and women, with men having higher mean concentrations. The study also identified smoking as a determinant of IGF-1 concentration in men, and coffee consumption as a determinant in women (23). Overall, the three studies agreed that age and sex were the most important determinants of IGF-I

serum concentration, but they did not reach consensus regarding the role that cigarettes, coffee, or alcohol plays, if any.

Serum IGF-1concentration increases slowly during childhood, peaking at approximately 14.5 years of age for girls, at a mean concentration of 524 μgs/L (24). For women, the mean decline in IGF-I is 37μgs/L per 10-year increase in age (23). There does not appear to be any seasonal or diurnal variation in IGF-I or IGFBP-3 concentrations (22).

GH, which is sensitive to nutritional status, regulates IGF-I produced by the liver; adequate intake of calories and protein is required to maintain an individual's normal serum concentration. Several studies have shown that long-term caloric restriction (CR) without malnutrition is a robust intervention in increasing the lifespan and health of rodents in a lab environment. CR decreases serum IGF-1 levels by up to 40% in rodents, potentially making this the key mechanism through which the rodents experience CR-dependent lifespan extension and cancer protection (25). IGF-1 and IGFBP-3 are relatively insensitive to short term changes in caloric intake, but CR, over a longer period of time, causes both IGF-1 and IGFBP-3 to decrease in adults (26). This reduction in protein concentrations could be a result of cells becoming GH-resistant in an environment of elevated GH levels. GH increases both IGF-I and IGFBP-3 in individuals that are well fed. Fontana *et al.* conducted three studies in which they determined that long term CR alone wasn't sufficient to reduce total and free IGF-I concentrations in humans if protein intake was high. They suggest that chronic protein intake is a more powerful determinant of circulating IGF-I than caloric intake, and that in order for IGF-I and IGFBP-3 serum concentrations to decrease, protein intake must be restricted as well (25).

2.3 IGF and neoplasia

Normal physiological IGF signalling plays a role in controlling the rate of cellular renewal, proliferation and apoptosis, making its signalling of interest to cancer researchers. It is important to note that there is no physiological level of IGF-I that is considered 'normal'; there is considerable variation in the levels of IGF-I and IGFBP-3 between individuals. However, it has been indicated that the risk of common cancers is higher in those with concentrations of circulating IGF-I levels in the high end of the physiological range, compared to those who have levels at the lower end of the scale. Many studies have reported a positive association between high levels of IGF-I and the development of various cancers, including prostate (27-29), breast (30-32), lung (33), colorectal (34) and ovarian (35) cancers. A meta-analysis of case-control studies estimating the association between concentrations of IGF-I and IGFBP-3 and various epithelial cancers found that higher concentrations of IGF-I were associated with an increased risk of prostate, colorectal, and premenopausal breast cancer, whereas higher concentrations of IGFBP-3 were associated with increased risk of premenopausal breast cancer (36). One way IGF-I signalling may impact cancer development is by favouring neoplastic progression of small lesions that would otherwise have remained clinically unapparent (37, 38). In support of this theory, model systems have shown that proliferation and metastasis of cancer cells may be increased by increasing IGFR-1 activation (39), and that the growth of tumour cells has been increased by both serum IGF-I and IGF-I produced by the tumour stroma itself. Tumour cells with IGF-I receptors may promote their own growth by synthesizing endogenous IGF-I, which could contribute to the partial autonomy and rapid growth characteristic of malignant cells (18).

IGFR-1 signalling promotes cancer cell proliferation and invasion (18, 40) and survival signalling required for anchorage-independent growth (38), an ability necessary for tumour cells to become metastatic. Further evidence for a link between IGF-1 signalling and cancer includes research showing cells in culture without the IGFR-1 receptor being refractory to transformation by viral oncogenes (41, 42), transforming at a rate that is 1000 times less than that of the cells with IGFR-1 (41). The down-regulation of the receptor in cell culture has been found to cause apoptosis and growth inhibition of cancer cells (41, 42), with one group reporting the reversal of transformed phenotypes in human cervical cancer cell lines in both the presence and absence of HPV (42). Shen *et al.* investigated the regulatory mechanism of IGFR-1 signalling and its importance in cervical cancer development and found that IGFR-1 proteins were abundant in cervical cancer cell lines but not in normal cervical epithelial cells (40). They also found that by blocking IGF-I stimulating effects with an antagonistic antibody, in a mouse model, cervical tumour growth was inhibited, and the tumours began to regress. This finding was corroborated in mice by Nakamura *et al.* (42).

It is generally hypothesized in the literature that high levels of IGFBP-3 should be protective against the development of cancer because of its ability to bind and sequester IGF-1, with some studies having found IGFBP-3 to be responsible for cellular senescence and growth inhibition (43, 44), and IGFBP-3 expression being lost in prostate and lung cancers (33, 45). Conversely, however, studies have also shown a correlation between high serum levels of IGFBP-3 with an increased risk of prostate, colorectal and breast cancers (46). One theory explaining the dual role IGFBP-3 seeming plays in either suppressing or promoting neoplastic development is that it has functions independent of its IGF-1 binding capacity which may differ depending on cell types or culture

conditions. Baege *et al.* demonstrated that E6/E7-immortalized cervical cells secreted 500-fold more IGFBP-3 than non-infected cells and displayed increased mitogenic sensitivity to IGF-1 after chronic pre-exposure to the binding protein. They concluded that *in vivo* expression of IGFBP-3 in cervical dysplasia may contribute to a selective growth advantage for HPV-immortalized cells (46).

2.4 IGF and HPV

Some research has shown that the E7 protein encoded by HPV-16 can target IGFBP-3, repressing its tumour-suppressor activity (ability to induce apoptosis and halt proliferation), and also signal its proteasome-dependent degradation in cervical cancer cells (47, 48). This reduction in IGFBP-3 levels could mean less bound IGF-1 with a resultant increase in IGFR-1 signalling. Another study found that E6/E7 immortalized human ectocervical epithelial cells secrete IGFBP-3, and yet are more sensitive to the mitogenic activity of IGF-1 (46).

Observational studies have also found a links between serum IGF concentrations and the risk of developing SILs. One study found a significant inverse association between IGFBP-3 and risk of incident detection of HR-HPV, and with HR-HPV positive SIL. The same study also found a significant decreased clearance of HR-HPV associated with a high IGF-1/IGFBP-3 molar ratio (49). Harris *et al.* found that having a high IGF-I:IGFBP-3 molar ratio was associated with an increased persistence of oncogenic HPV infection (a lower rate of clearance), whereas IGFBP-3 was inversely associated with both incident detection of HR-HPV types and the incidence of HR-HPV positive SIL (50).

3.0 Limitations of Prior Studies

Only a few observational or case control studies published have investigated the role of IGF-1 and IGFBP-3 on the risk of developing SIL, and their findings were contradictory. Schaffer *et al.* found that increasing levels of IGF-I were associated with a reduced risk of HG-SIL, and that higher levels of the peptide were associated with a reduced risk of being positive for HPV-16 or HVP-18 among controls (51). These findings are supported by the work of Serrano *et al.* (52) who also found significantly lower levels of IGF-I and IGF-1:IGFBP-3 molar ratio in cases as compared to the reference category. The findings of Wu *et al.*, however, contradict this. Their study reported IGF-I levels in the highest quartile were associated with a significant increase in risk of SIL compared to those in the lowest quartile (53). Another study investigated the association between levels of serum IGF-1 and cervical cancer (54), and found that high over-expression of the IGFR-1 receptor was an independent predictor of cervical cancer death and recurrence, that pre-operative total serum IGF-I or IGFBP-3 levels failed to predict cervical cancer death and recurrence. They also reported that there was a lack of correlation between circulating IGF-I or IGFBP-3 with IGFR-1 over-expression in the cervical cancer cells, suggesting a likely autocrine or paracrine IGF-I stimulation of IGFR-1 signalling.

Studies investigating the role of circulating concentrations of IGF-1 and IGFBP-3 in the development of SIL and cervical cancer have reported differing results. To determine if there is a true association between levels of IGF-I, IGFBP-3 and cervical cancer, a nested case-control study is necessary. Circulating levels of IGF-I are modulated by GH, which in turn is modulated by nutritional status. GH can decrease its

stimulation of IGF-1 production in times of poor nutrition, a state common to cancer patients. Despite the fact that intervention occurs in this study before patients develop cancer, this study would prevent any bias that would occur should IGF-I level be associated with the disease status. Circulating levels of IGF-1 and IGFBP-3 should, in a best-case scenario, be determined in a large number of healthy individuals before subsequent long-term observation, which we were able to do. After follow-up, women who have developed SILs were identified and assays were performed on their stored blood samples. This method minimizes the possibility that conclusions would be biased by the effect of the disease on the IGF-I and IGFBP-3 levels. Prior to our study, there has been no such prospective investigation of the role of circulating IGF-1 and IGFBP-3 levels on the development of SIL.

It is important to keep in mind that cervical cancer differs from other epithelial neoplasias because cervical cancer cells are immortalized via a viral infection; IGF-I and IGFBP-3 may not play the same role in its development as it does in other epithelial cancers.

4.0 Design and Methods

4.1 Statement of Objectives

This study was designed to test the hypothesis that circulating levels of IGF-I and IGFBP-3 may be associated with cervical lesion development. This study is the first nested case-control study investigating the baseline circulating levels of IGF-1 and IGFBP-3 as a predictor for SIL development, as well as levels measured at time of diagnosis. All previous studies have only measured IGF-1 and IGFBP-3 at time of case ascertainment. The primary specific aim of this project is thus as follows:

To estimate the association between circulating IGF-1, IGFBP-3, and IGF-1:IGFBP-3 molar ratio quartiles and risk of developing any SIL or HG-SIL.

The following were ancillary aims:

To identify factors influencing circulating IGF-1 and IGFBP-3 levels in the Brazilian population (eg: age, education, salary, year of measurement);

To estimate the association between circulating IGF-1, IGFBP-3, and IGF-1:IGFBP-3 molar ratio quartiles and risk of any HPV infection, LR-HPV infection, or HR-HPV infection;

To estimate whether circulating IGF-1 level influences time to SIL development.

If enrolment levels of IGF-1/IGFBP-3 are predictive of the subsequent risk of cervical lesions (LG-SIL and HG-SIL) independently of HPV variables, the net effect of IGF/IGFBP levels would be demonstrable only in the strata of patients who are infected by HPV because of the very low risk of incident lesions among women who have remained free of HPV infection during the beginning of follow-up.

Unlike case-control studies which do not permit temporal inference, or cohort studies based on a single IGF/HPV measurement at entry that do not permit the assessment of infection incidence and clearance, this nested case control study is ideal for testing the etiologic role of IGFs in the development of cervical lesions because of the availability of two measurements at different time points.

4.2 The Ludwig-McGill Cohort Study

4.2.1 Overview

The Ludwig-McGill Cohort study was designed to investigate the molecular epidemiology and natural history of cervical HPV infection and neoplasia. Participants were recruited between 1993 and 1997, and in its entirety the study is comprised of over 2500 women. Follow-up was for over 5 years, and extended into 2002. The participants were seen every 4 months in the first year, and once a year thereafter (4, 55, 56). The nested case control study that is the subject of this thesis includes 603 women, which includes incident/prevalent cases with 2:1 matching of controls to cases.

4.2.2 Study Setting

The study was carried out in the city of São Paulo, in the most populous and industrialized state in Brazil, which had a population of approximately 12 million at the beginning of this study. São Paulo is a city with one of the highest risks worldwide of cervical cancer (incidence age-standardized of 21.1 per 100 000) (57). The study was conducted at the Maternidade Escola Dr. Mario de Moraes Altenfelder Silva Municipal Hospital (MEVNC). The clinic is part of a network of primary, secondary, and tertiary health care institutions maintained by the municipal health department.

4.2.3 Subject Recruitment

Selected participants for the study were women attending a comprehensive maternal- and child-health program, catering to low-income families. Two trained nurses approached women selected at random from daily visit lists of outpatients from family medicine, gynaecology, and family planning clinics at MEVNC, and then determined

their eligibility via interview. The nurses also explained the nature of the study and its general purpose.

Women were eligible to participate in the study if they were between the ages of 18 and 60. A previous study (58) in a similar population has shown that prevalence of HPV is at its highest in women in their early twenties, that incidence of carcinoma *in situ* peaks in women in their mid thirties, and that incidence of invasive cervical cancer is greatest in women about 60 years old. Women were also eligible if they were permanent residents of the city São Paulo, so as to minimize loss of follow-up, and if they were not currently pregnant and had no intention of becoming so in the following 12 months. A pregnancy would interfere with cervical cell sample collection and cervicography, and by excluding pregnant women we would minimize a bias resulting from missing information. Women also had to have an intact uterus and no current referral for a hysterectomy, which would restrict the study population to those women at risk for cervical cancer. They must not have reported use of vaginal medication in the previous 2 days, as recent use of vaginal medication could hamper the ability to collect good quality cervical cell specimens. Lastly, women must not have had any treatment for cervical disease by electrocoagulation, cryotherapy, or conization in the previous 6 months. This criterion eliminated women who would have already been diagnosed with the outcome of interest (pre-neoplastic or neoplastic lesions).

Women were also considered ineligible if they were not interested in complying with all scheduled return visits, for at least 2 years.

Those that were potentially eligible were given an in-depth description of participation, and informed that an incentive to compliance consisting of a meal ticket would be given at each completed visit. Meal tickets had inflation adjusted cash value,

and came in various denominations, beginning at US $5 at the enrolment visit, and increasing $5 per subsequent visit to a maximum of $20. This strategy resulted in excellent rates of follow-up compliance, despite the complexity of the procedures used in the study and the requirement for blood specimens. All of the study procedures and the informed consent were approved by the institutional review boards and ethical committees at the participating institutions: McGill University, Montreal, Quebec, Canada; the University of Toronto, Ontario, Canada; and the Ludwig Institute for Cancer Research and the MEVNC clinic, both in São Paulo, Brazil.

4.3 Data Collection

4.3.1 QUESTIONNAIRES

In the first four visits, and in the annual returns, subjects completed an interviewer-administered structured questionnaire specific to the current visit, given by nurses that had been extensively trained in interview strategies. The information collected in the interviews covered all classes of risk factors for HPV infection and cervical neoplasia, such as sociodemographics, reproductive health, sexual practices, smoking, and diet. A sample questionnaire has been included in the appendix.

4.3.2 CERVICAL SPECIMENS, HPV DNA, AND IGF/IFBP-3

At each clinical visit, an Accelon biosampler (Medscand, Inc., Hollywood, FL, USA) was used to collect ectocervical and endocervical cells, which were then used for cytological screening and HPV testing. A cervical cell smear was fixed onto a slide in 95% ethanol, and then the sampler containing the exfoliated cells was immersed into a tube of Tris-EDTA buffer pH 7.4. The tube underwent agitation to release the cells from the sampler, and samples were subsequently kept at 4° at the clinic for 5 days at most

before being brought to the Ludwig Institute, were they were kept at -20°C until testing. The slides were stained and read at the Ludwig Institute's cytopathology laboratory for an initial diagnosis, and then sent to the laboratory of Dr. Alex Ferenczy at the Jewish General Hospital in Montreal for reading. The Canadian cytopathology reports were based on the Bethesda system for cytological diagnoses (7). To prevent concerns regarding false negative results, women also underwent a cervicography during their first year of participation, and then again at 24 and 48 months. This test was used to detect clinically relevant lesions that are visually identifiable, thus providing a safety net to supplement the information obtained from the cytological readings. Cervicography has been proposed as a useful tool in large-scale studies in high-risk populations, where well-trained colposcopists are hard to recruit (59). The international rights to cervicography and its trademarks are held by National Testing Laboratories Worldwide (NTL), Fenton, Missouri, USA. The NTL technical representative visited São Paulo and trained two of the study nurses in the use of the cerviscope and on the cervicography procedure. Log sheets and rolls of film were prepared according to the NTL protocol instructions and then shipped monthly to Missouri, for development and evaluation by NTL's expert colposcopists. Results were then mailed to the project manager in Montreal, for computer data entry.

If a woman was found to have lesions of moderate dysplasia or worse in the initial screening, they were referred for colposcopy at MEVNC. This referral also happened if the cytopathology review performed in Montreal revealed a diagnosis of a HG-SIL or worse, or if the cervigram indicated HG- lesions or worse. At colposcopy, if any lesion tissue was present, a biopsy was taken for histological assessment and, if indicated, the woman was treated according to the local prevailing protocol. Those women with

positive biopsies were removed from the study and did not contribute any more person time.

Table 2: Study procedures and instruments applied at enrolment and at different prescheduled follow-up returns in the Ludwig-McGill cohort study, São Paulo, Brazil, accrual period 1993-1997

Procedures	0	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	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Host susceptibility markers												
IGF-1	✓	← ✓ (at first HPV diagnosis + at time of SIL diagnosis) →										
IGFBP-3	✓	← ✓ (at first HPV diagnosis + at time of SIL diagnosis) →										
Cervical Pathology												
Local Pap cytology	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Cytology Review	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Cervicography		(once 1st year)				✓				✓		
Colposcopy + biopsy			← ✓ (Whenever needed if HG-SIL) →									
Questionnaire information												
Sociodemographics	✓											
Diet		✓								✓		
Reproductive health	✓		✓									
Sexual behaviour, smoking	✓	✓	✓	✓		✓		✓		✓		
Health Attitudes and beliefs										✓		
Compliance incentive												
Meal tickets (USD)	5	10	15	20	20	20	20	20	20	20	20	20

4.3.3. RISK-SET SAMPLING

There were 2462 women in the Ludwig-McGill Cohort Study who were eligible for selection into the nested case control study. At the beginning of follow-up, 51 prevalent SILs were identified in the Ludwig-McGill cohort, and 150 subsequent SILs were diagnosed over the course of the study. Figure 2 describes the selection of subjects from the Ludwig-McGill Cohort Study into the nested case control study. The criteria for sampling women included risk-set sampling (matching based on follow-up time), enrolment date (controls were matched on the same month of enrolment into the study as the case), and woman's age (within 5 years if possible).

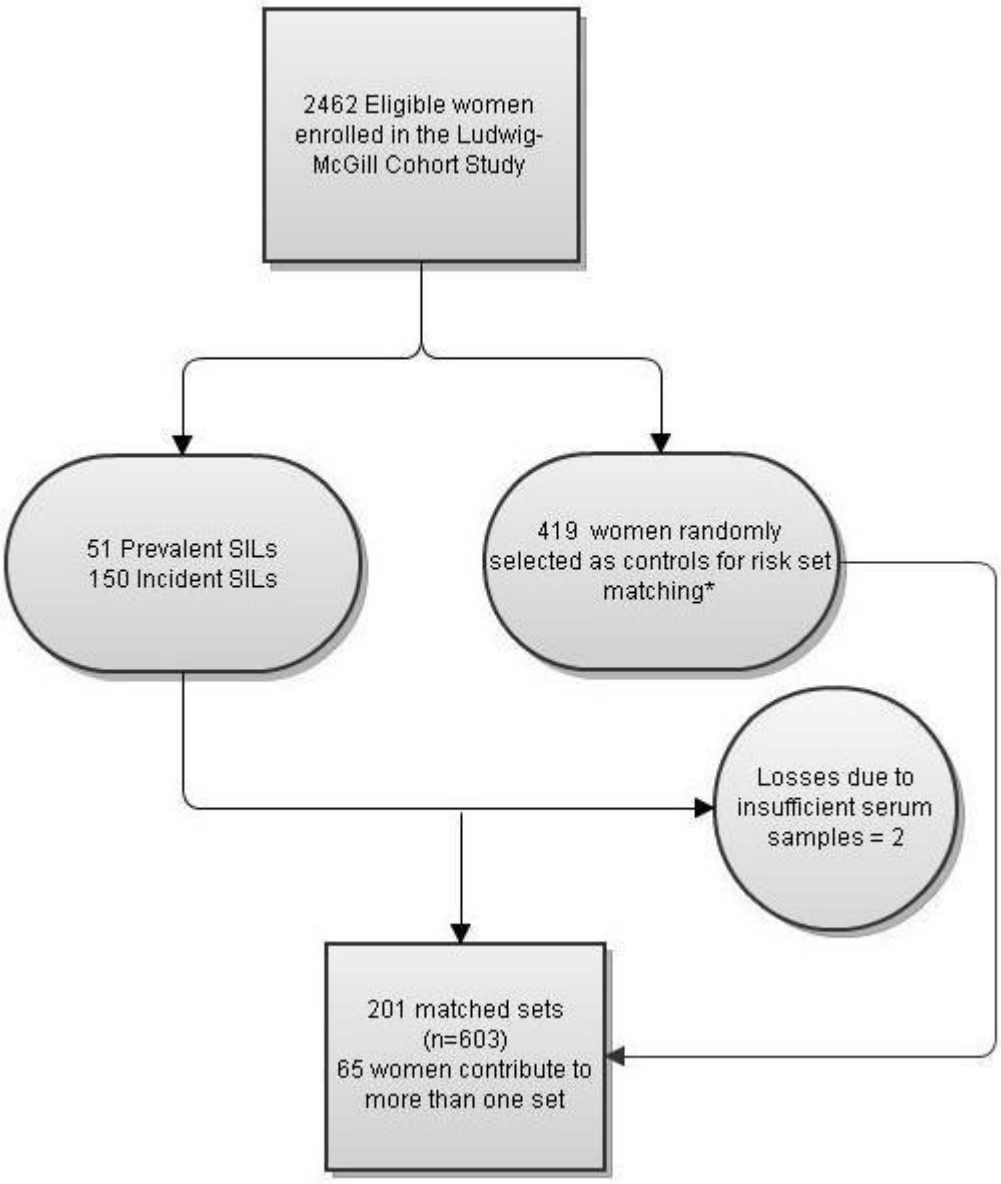

Figure 2: Flowchart describing the selection of subjects from the Ludwig-McGill Cohort Study into the nested case control study. * Incident cases could potentially act as controls prior to lesion development. Criteria for risk set sampling: controls had to have the same minimum follow-up time; an ASCUS event for the candidate control before the index SIL of its respective case invalidates eligibility of control; controls were matched on same month of enrolment into the study as the case and on age (within 5 years if possible); all cases were of first instance of any SIL and two controls were selected per case; and case control sets were chosen based on the highest lesion grade attained.

4.3.4 IGF MEASUREMENT

After the risk-set sampling was complete, the required samples were sent from the Ludwig Institute in Brazil to the laboratory of Dr. Pollak at the Jewish General Hospital in Montreal, Canada. For budgetary reasons, the samples remained in the freezer (-80°C) for about 3 years before assays could be completed.

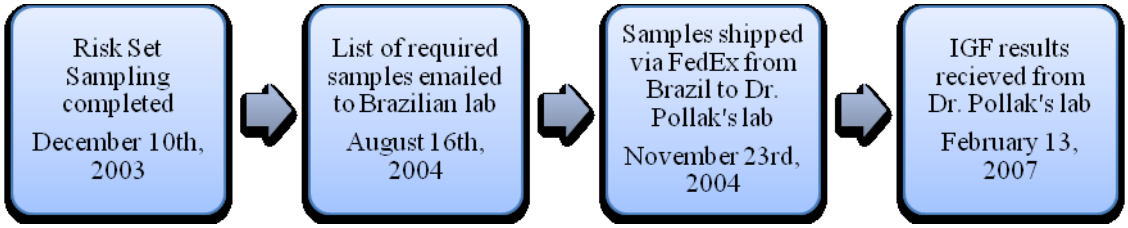

Figure 3: IGF sampling timeline.

Serum levels of IGF-1 and IGFBP-3 were assayed by enzyme-linked immunoabsorbant assay (ELISA) (Diagnostic Systems Laboratory, Webster, Texas). Before IGF-1 assay, IGFBP-3 was removed via acid-ethanol extraction. Each sample was tested in duplicate, and the mean was used for data analysis. If the mean relative difference between the two samples exceeded 10%, the assay was repeated. All assays were carried out in a blinded manner, and quality control samples were included within assay runs.

4.3.5. HPV DNA DETECTION

Cervical specimen DNA was extracted and purified following standard techniques. Briefly, cells were digested with 100μg/ml proteinase K for 3 hrs at 55°C, followed by organic extraction and ethanol precipitation. Specimens were tested for the presence of HPV DNA by a previously described polymerase chain reaction (PCR) protocol amplifying a highly conserved 450 bp segment in the L1 viral gene (flanked by primers MY09/11)(60, 61). Typing of the amplified products is performed by hybridization with individual oligonucleotide probes specific for all 27 HPV genital types whose nucleotide sequence probes within the MY09/11 fragment have been published in the literature. Twenty-three of these have received a taxonomic entry as HPV types: 6/11, 16, 18, 26, 31, 33, 35, 39, 40, 42, 45, 51, 52, 53, 54, 55, 56, 57, 58, 59, 66, 73, 82, 83, and 84.

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 (60)) to amplify a region of the β-Globin gene. All HPV assays were done blindly on coded specimens with no identification linking specimens from the same woman.

5.0 Definition of Variables

Variables of interest were extracted from the information collected during the questionnaire-based interviews. Variables were either 1) sociodemographic characteristics 2) sexual history & behavioural characteristics, or 3) reproductive and contraceptive history characteristics.

Sociodemographic or design variables included the woman's age at baseline, her educational attainment, religion, race, marital status, and the year of her baseline IGF-I/IGFBP-3 measurements. Sexual history and behavioural variables examined included the woman's age at first sexual intercourse, her cumulative number of sexual partners at baseline, STDs at baseline and drinking habits. Reproductive and contraceptive history variables of interest were the number of previous pregnancies, age at menarche, number of previous pap tests, and years of oral contraceptive use.

The coding of each variable was carefully assessed in the STATA database and compared to that of the questionnaire, to ensure it had been entered in correctly.

5.1 Sociodemographic Variables

Woman's age at entry into the study was obtained from the questionnaires as a continuous measure, but for the purpose of validating IGF-1 and IGFBP-3 measurements, was broken into decades.

Race was initially coded into white, mulatto, black, Asian, and native origin, but due to the relatively low proportion of individuals in some of the categories, race was changed into a dichotomous variable; white or non-white.

Educational attainment, as originally coded, had 7 categories: none, elementary incomplete, elementary complete secondary incomplete, secondary complete, college-technical-professional training, and university. The study population had very few women with an educational attainment greater than elementary school, and so the categories were re-organized into less than elementary school, elementary school completed, high school completed, and college/university completed.

A woman's monthly income was obtained as a continuous variable, which was converted into US dollars using monthly exchange rates to correct for the heavy inflation that was occurring in Brazil until June 1994. Conversion was necessary because of the economical instability, and because a new currency, the Real, was introduced to Brazil in 1994, making it impossible to compare incomes reported by women during and after this time based exclusively on the local currency. After converting to US dollars, annual income was categorized into quartiles: $55-$349, $350-$729, $730-$39,999, $40,000+. In this study, a woman's annual salary in USD was used as a proxy for a woman's ability to eat properly and therefore her nutritional status.

Marital status was recoded from its original 5 categories of single, married, widowed, separated, and unmarried but living with a partner, into a dichotomous variable. The two new categories were single/widowed/separated, and married/living with a partner.

Religion was also recoded due to a small proportion of women in some of the categories. Those that identified as either Protestant or Crente were grouped together, as Crente is a form of Protestantism. Catholic and Protestant were the two largest denominations, with all other women being categorized as being of 'other' religion.

5.2 Markers of sexual activity

Age at first intercourse was collected as a continuous variable, but was then categorized into quartiles: <16, 16-17, 18-19 and 20+.

5.3 IGF-I and IGFBP-3 Measurements and SILs

In cases, the serum concentrations of IGF-1 and IGFBP-3 used in the analysis were measured at the baseline visit, and at the visit in which they were identified as having either a LG-SIL or HG-SIL. The serum concentrations of IGF-1 and IGFBP-3 in controls were measured at the baseline visit, and at the visit in which their matched case was diagnosed.

IGF-1:IGFBP-3 molar ratios were calculated using the following formula: ([IGF-1]*0.13)/([IGFBP-3]*0.035) (62, 63).

To identify determinants of IGF-1 and IGFBP-3, the serum concentrations of these proteins, which were continuous, were dichotomized at the median value identified for the controls in each group.

The association between IGF-1 and IGFBP-3 and risk of SIL was analyzed using IGF-1 and IGFBP-3 categorized into quartiles based on the distribution in control subjects.

When analyzing whether IGF level was associated with risk of developing a SIL, women were classified as SIL positive if at any of their visits they presented with either a LG-SIL or a HG-SIL.

5.4 HPV Types

HPV types were classified as either being of high-risk (HR-HPV) oncogenic potential or low-risk (LR-HPV) oncogenic potential (64). HR-HPV types included types 16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, 59, 68, 73, and 82. LR-HPV types included 6, 11, 13, 74, 44, 55, 91, 7, 40, 32, 42, 54, 57, 61, 62, 71, 72, 81, 83, 84, and 89. Also examined was HPV status, defined in four ways; those who tested positive for any HPV

type; those who tested positive for at least one high-risk HPV type; those who tested positive for a LR-HPV type; those who tested positive for HPV-16; and those who tested positive for HPV-6 and/or HPV-11.

5.5 Nutritional Intake

There were several questions on the second questionnaire that aimed to determine the nutritional status of the study participants. These questions were designed originally to determine if various vitamins and dietary items impact HPV status/infection, and the questionnaire was not designed with the present nested case control study in mind. As such, because of its limited range of questions on diet, the available nutritional information was used only as an approximation of the woman's nutritional status.

All food intake variables in the questionnaire were graded on a 7 point scale, ordered from never eating those particular foods to eating them daily or more often. Those food variables that were approximating the same nutrient (eg: papaya, carrot, and pumpkin were related to/proxies for beta-carotene intake) were then grouped together and the rank sum for frequency of intake for each woman was calculated. These sums were used to determine the overall frequency of intake and were categorized into quartiles.

6.0 Statistical Analysis

6.1 Descriptive Statistics

The proportion of women in each variable category was calculated for all categorical variables. The mean and standard deviation was calculated for all continuous variables. The distribution of variables in the study population was compared to that of the original cohort study.

In most study populations, IGF-1 and IGFBP-3 serum concentrations are normally distributed, however, in the Ludwig-McGill study population, this was not the case (data not shown). In order to validate the measurements and to rule out lab/collection errors, IGFBP-3 was first regressed against IGF-1. IGF-1 and IGFBP-3 should have a linear relationship; as one increases, so should the other. The two proteins were also graphed against the woman's age at time of collection, as other studies have shown that IGF serum concentrations tend to decrease with age (21-24). The concentrations were also graphed against a woman's annual salary in USD, which was used as a proxy for proper nutrition and food/protein availability. This was a low-income study population, and so as salary increased, hypothetically the IGF-1 serum concentration should have increased as well.

6.2 Determinants of IGF-1 and IGFBP-3 serum concentrations

As stated above, a secondary objective of this research project was to determine which factors in this study population influenced the serum concentrations of circulating IGF-1 and IGFBP-3. An important caveat must be recognized in this respect. A nested case control study is not an ideal study design for exploring determinants of IGF, as the sample is not representative of the population – all cases are included, and only a subset

of the available controls are included, and they have been matched to the cases based on risk-set sampling. However, the Ludwig-McGill Cohort study was already not representative of the population of São Paulo as women who participated in the cohort had geographical reasons for doing so. While this is a drawback for some reasons, it is also a strength in that it let us have a high response rate and long follow-up. The purpose of this analysis is therefore simply to get an idea of potential determinants of IGFs in this sample, and the results cannot be externally generalized to the population of São Paulo or for the specific source population (because of the matching). Also, sensitivity analysis was done using only the controls, as they are slightly more representative of the average women in the source population.

Neither IGF-1 nor IGFBP-3 had a normal distribution, either linearly or after log-transformation, and so IGF-1/IGFBP-3 was dichotomized as either being above or below the median value. This also helped to increase the statistical power. Odds ratios (ORs) and their 95% confidence intervals (95% CI) were calculated using unconditional logistic regression to measure the associations between various sociodemographic, behavioural, sexual, and reproductive factors and serum concentrations of IGF-1 and IGFBP-3.

Variables that were significantly associated with IGF-1/IGFBP-3 were then put into a backward-selection estimation model. As age has consistently been shown to be an important determinant of IGF-I levels (21-24), it was chosen *a priori* to be included in the model. If the removal of any of the variables changed the OR for the relationship by more than 5%, they were considered to be confounders or mediators.

6.3 Serum IGF-1/IGFBP-3 concentrations as a determinant of SIL development

The main objective of this study was to determine whether baseline or diagnosis levels of IGF-1/IGFBP-3 are predictive of LG-SIL or HG-SIL. As the risk of developing

an SIL after having tested HPV DNA negative throughout the study is so small, women were not stratified on their HPV status. ORs and their 95% CIs were calculated using conditional logistic regression to measure the associations between IGF-1/IGFBP-3 variability as well as IGF-1:IGFBP-3 molar ratio with any SIL or HG-SIL. The analysis was done separately for prevalent cases, and the analysis included IGF-1/IGFBP-3 measured at baseline, at diagnosis and the highest measured IGF-1/IGFBP-3 concentration. Sets included one case and two controls, matched on age and time at risk. Variables were selected the same way as mentioned above.

6.4 IGF-1, IGFBP-3 and IGF-1:IGFBP-3 molar ratio as predictors of HPV infection

Another secondary objective of this study was to investigate whether a woman's serum concentration of either IGF-1 or IGFBP-3 influenced her susceptibility of acquiring an HPV infection. Again, an important caveat must be recognized here, because a nested case control study is not the appropriate study design to carry out this kind of analysis. However, for the reasons mentioned above in section 6.2. I went ahead with an exploratory analysis. The ORs in this analysis do not gauge the true magnitude of association that would be seen in the general population, but they are useful only for trying to identify associations. Sensitivity analysis was done for this analysis using controls only as they were the best approximation to the local source population of São Paulo.

For this analysis conditional logistic regression could not be used because the sets had been made based on SIL outcome, not HPV. ORs and their 95% CIs were calculated using unconditional logistic regression to measure the associations between IGF-

1/IGFBP-3 quartiles with any HPV, HR-HPV, LR-HPV, HPV-16, or HPV-6/11 infections.

The analyses used the maximum IGF-1 and IGFBP-3 values, because it is hypothesized that these values represent the woman's normal healthy IGF-I/IGFBP-3 levels. Analyses were age-adjusted, and any potential confounders identified in the previous section were adjusted for. Age was chosen to be included *a priori.*

6.5 Time to SIL diagnosis

A tertiary objective of this thesis was to determine if there were any associations between level of IGF-1, IGFBP-3 or IGF:IGFBP-3 molar ratio and speed of lesion development in those that did develop lesions. This analysis is not a true Kaplan-Meier analysis, as only cases are used; controls in matched sets had their follow-up time truncated after their paired case was diagnosed and could not be used. Had they been included, it would have led to all controls being censored at the time of last follow-up visit. Because the time on study for controls is underrepresented in this nested case-control design a formal Kaplan-Meier analysis would not have been informative. Therefore, I chose to conduct an analysis that was conditional on women having already become a case by the end of the study, and so the median time to event in this section is not the true median time to event that would be seen in the population. The purpose of this analysis is simply to see if level of IGF affects speed of lesion development in those that did develop lesions at the study closing date. For this reason, the point estimates do not matter, just the association of IGF with the rate of lesion development.

The baseline visit, t_0, was the starting point for all subjects, although it was not the same point in real time for all subjects. Time until SIL was measured as the number of years from the date of the baseline visit until the visit date during which the lesion was

diagnosed. The Kaplan-Meier or product-limit method (65) for survival analysis was used to obtain the distribution of events over time, stratified by whether IGF-1 concentration was above or low the median value. The log-rank test was used to compare the observed and expected distribution of time to event between women with high versus low IGF-1 serum concentration. Graphical representations are shown in terms of the survival function, S(t). Again, the caveat here is the conditionality of the time-to-event analysis, i.e., the estimates are conditional and exclusive to those who had the lesion event. Therefore, all time-to-event times are complete and thus not censored.

7.0 Results

7.1 Descriptive Results

Baseline characteristics of the nested case control study population were compared to those of the entire cohort (Table 3 & Table 4). The populations did not appear to have any major differences in the distribution of the various variables. It should be noted that the number of sexual partners in the entire cohort seems to be larger than that for the nested case control study. This is due to the measure being a mean and not a median. The mean is highly skewed by a few outliers, resulting in an exaggerated mean and a large standard error.

Table 3: Baseline socio-demographic characteristics of women selected for the nested case control study, and those of the entire Ludwig-McGill Cohort according to unit of analysis (case/control status).

Variable	Category	Cases			Controls	Entire Cohort
		HSIL (33)	LSIL (168)	Overall (201)	(402)	(2462)
		N(%)	N(%)	N(%)	N(%)	N (%)
Age at baseline	18-23	5 (15.2)	45 (26.8)	50 (24.5)	92 (22.9)	386 (15.7)
	24-29	10 (30.3)	42 (25.0)	52 (25.9)	98 (24.4)	534 (21.7)
	30-34	7 (21.1)	30 (17.9)	37 (18.4)	91 (22.6)	516 (21.0)
	35+	11 (33.3)	51 (30.4)	62 (30.9)	121 (30.1)	1026 (41.7)
Race	White	25 (75.8)	102 (60.7)	127 (63.18)	237 (59.0)	1585 (64.4)
	Non-White	8 (24.2)	66 (39.3)	74 (36.8)	165 (41.0)	874 (35.5)
Education at baseline	< Elementary	6 (18.2)	39 (23.2)	47 (18.5)	79 (19.7)	554 (22.5)
	Elementary	22 (66.7)	92 (54.8)	148 (58.3)	240 (59.7)	1438 (58.4)
	High School	5 (15.2)	32 (19.1)	54 (21.3)	73 (18.2)	397 (16.1)
	Col/Univ	0 (0.0)	5 (3.0)	5 (2.0)	10 (2.5)	70 (2.8)
Smoking at baseline	Never	12 (36.4)	70 (41.7)	111 (43.7)	187 (46.5)	1168 (47.5)
	Current	18 (54.5)	71 (42.3)	104 (40.9)	133 (33.1)	864 (35.1)
	Former	3 (9.1)	27 (16.1)	39 (15.4)	82 (20.4)	429 (17.4)
Income in USD at baseline	55-349	9 (27.3)	49 (29.2)	58 (28.9)	83 (20.7)	509 (20.7)
	350-729	9 (27.3)	42 (25.0)	51 (25.4)	108 (26.9)	685 (27.8)
	730-39,999	7 (21.2)	42 (25.0)	49 (24.4)	102 (25.4)	642 (26.1)
	40,000+	8 (24.2)	35 (20.8)	43 (21.4)	109 (27.1)	625 (25.4)

Table 4: Baseline sexual behaviour and reproductive health characteristics of women selected for the nested case control study, and those of the entire Ludwig-McGill Cohort.

Variable	Category	Cases			Controls	Entire Cohort
		HSIL(33)	LSIL(168)	Overall (201)	(402)	2462
		N(%)	N(%)	N(%)	N(%)	N(%)
Age at first intercourse	<16	15 (45.5)	49 (29.1)	91 (35.8)	124 (30.9)	676 (27.5)
	16-17	11 (33.3)	52 (31.0)	79 (31.1)	119 (29.6)	632 (25.7)
	18-19	5 (15.2)	42 (25.0)	56 (22.1)	73 (18.2)	518 (21.1)
	20+	2(6.1)	25 (14.9)	28 (11.0)	86 (21.4)	635 (25.8)
	Mean(SD)	15.9 (2.6)	17.0 (3.0)	16.5(2.9)	17.5 (4.0)	17.9 (3.9)
Number of lifetime sexual partners at baseline	0-1	9 (27.3)	61 (36.3)	70 (34.8)	192 (47.8)	1089 (44.3)
	2	7 (21.2)	34 (20.2)	41 (20.4)	82 (20.4)	513 (20.9)
	3+	17 (51.5)	73(43.5)	90 (44.8)	128 (31.8)	858 (34.9)
	Mean(SD)	2.2 (0.9)	2.1(0.9)	2.1 (0.9)	1.8 (0.9)	5.8 (89.0)
Total number of pregnancies at baseline	0-1	2 (6.1)	40 (23.8)	48 (18.9)	67 (16.7)	417 (16.9)
	2-3	12 (36.4)	52 (31.0)	92 (36.2)	181 (45.0)	1041 (42.3)
	4-6	12 (36.4)	60 (35.7)	90 (35.4)	107 (26.6)	737 (29.9)
	7+	7 (21.2)	16(9.5)	24 (9.5)	41 (10.2)	248 (10.1)
	Missing	0 (0.0)	0 (0.0)	0 (0.0)	6(1.5)	19 (0.8)
OC use at baseline	Never	3 (9.1)	37 (22.0)	46 (18.11)	74 (18.4)	401 (16.3)
	< 6 Years	18 (54.6)	86 (51.2)	148(58.3)	221 (55.0)	1349 (54.8)
	6+ Years	12 (36.4)	45 (26.8)	60 (23.6)	107 (26.6)	711 (28.9)
STDs at baseline	No STD	25 (75.8)	131 (78.0)	201(79.1)	322 (80.1)	1881 (76.4)
	HPV- STD	3 (9.1)	12 (7.1)	15 (5.9)	16 (4.0)	108 (4.4)
	Other- STD	5 (15.2)	24 (14.3)	37 (14.6)	61 (15.2)	463 (18.8)
	Missing	0(0.0)	1 (0.6)	1 (0.4)	3 (0.8)	9 (0.4)
Age at Menarche	8-12	14 (42.4)	76 (45.2)	100 (45.3)	201 (50.0)	1089 (44.2)
	13+	19 (57.6)	90(53.6)	119 (53.9)	199 (49.5)	1367 (55.5)
	Missing	0 (0)	2 (1.2)	2 (0.9)	2 (0.5)	6 (0.2)

The relationship between IGF-1 and IGFBP-3 serum concentration was examined and found to be correlated (correlation coefficient=0.83), and the function [IGF-1] = [IGFBP-3]^2 had an R^2 value of 0.74 at baseline and 0.71 at diagnosis (Figure 4). Both IGF-1 and IGFBP-3 serum concentrations were inversely correlated with the women's age at the time of measurement (correlation coefficients of -0.14 and -0.15 respectively), as shown in (Figure 5). As age increased, the median value of IGF-1 and IGFBP-3 serum concentration decreased. IGF-1 was also slightly correlated with salary (correlation coefficient of 0.28), as was IGFBP-3 (correlation coefficient of 0.30). As a woman's annual income increased, so did the serum concentrations of IGF-1 and IGFBP-3 (Figure 6).

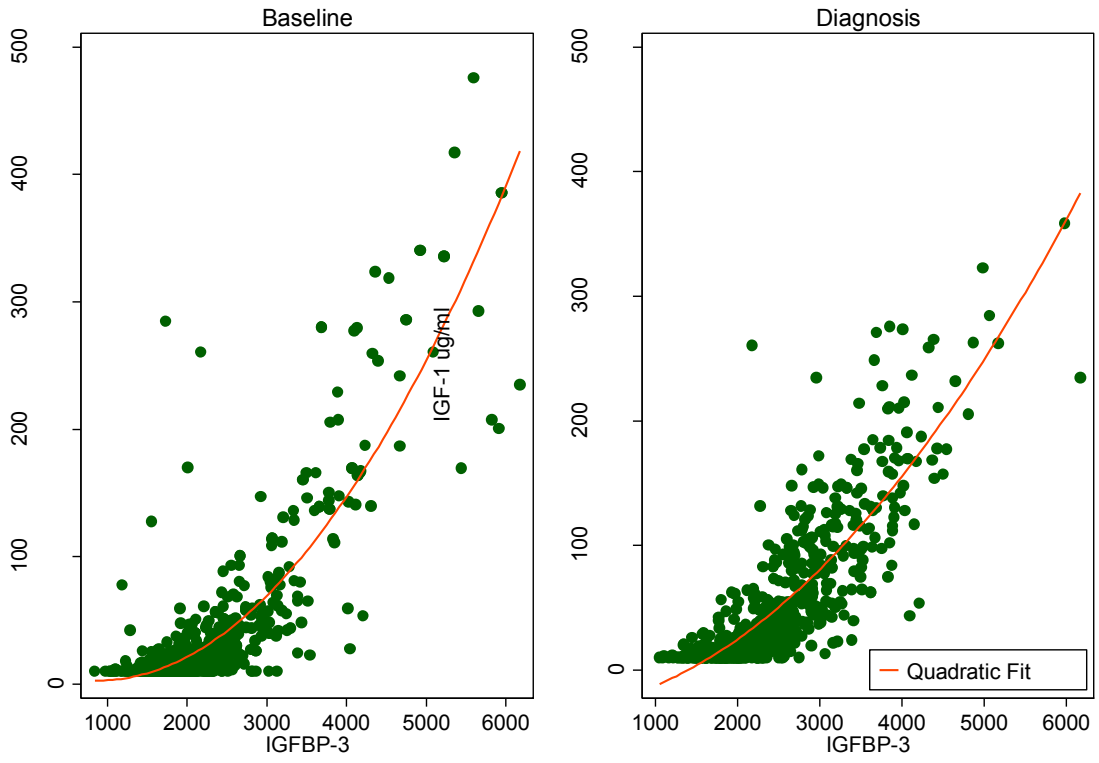

Figure 4: Correlation between IGF-I and IGFBP-3, measured at either baseline or at time of diagnosis.

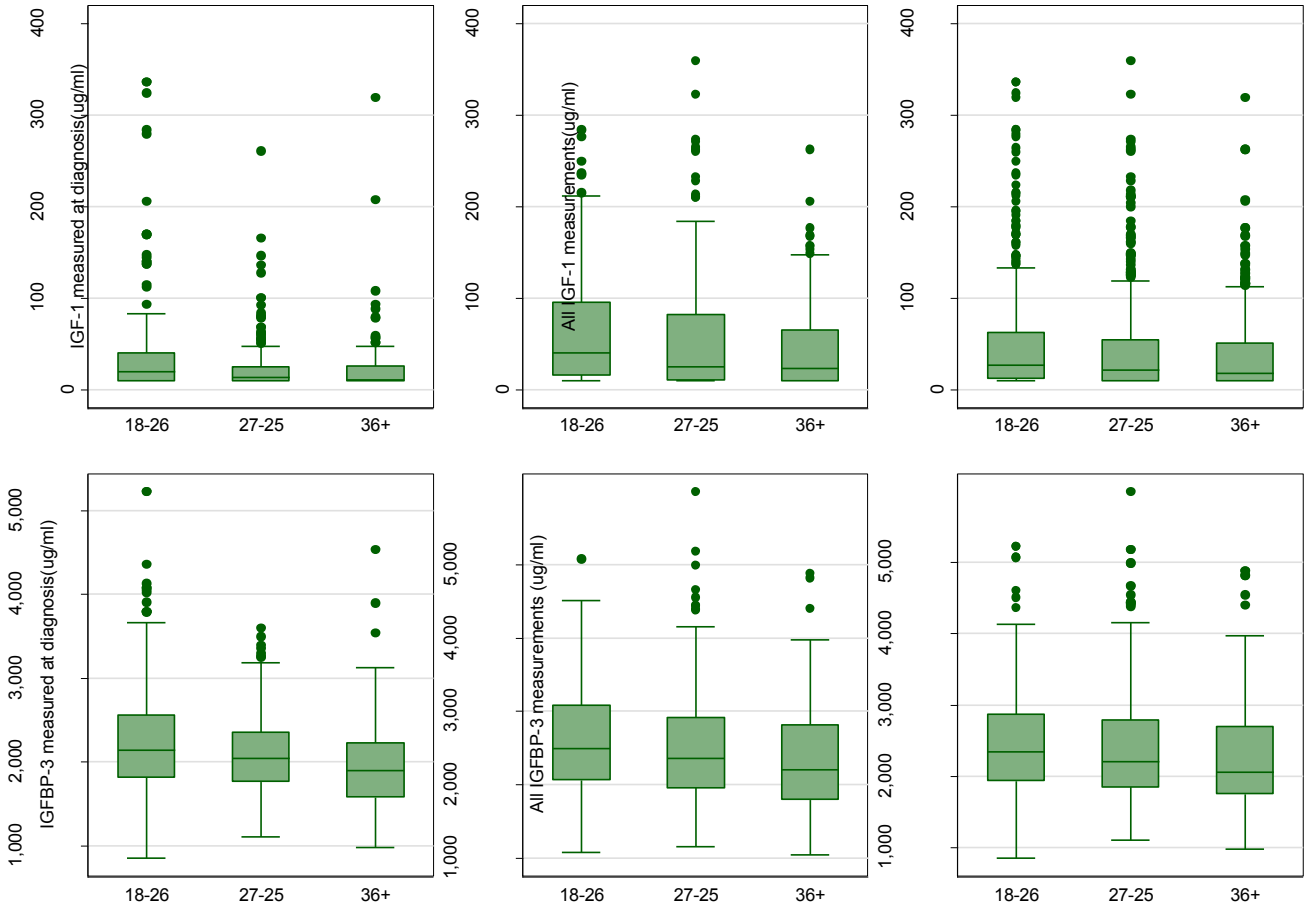

Figure 5: The relationship between serum concentrations of IGF-1 and IGFBP-3 and woman's age at time of measurement. Graphs on the far left show the relationship using IGF-1/IGFBP-3 samples taken at baseline, the graphs in the middle with samples taken at time of case diagnosis, and the graphs on the far right with all of a woman's samples.

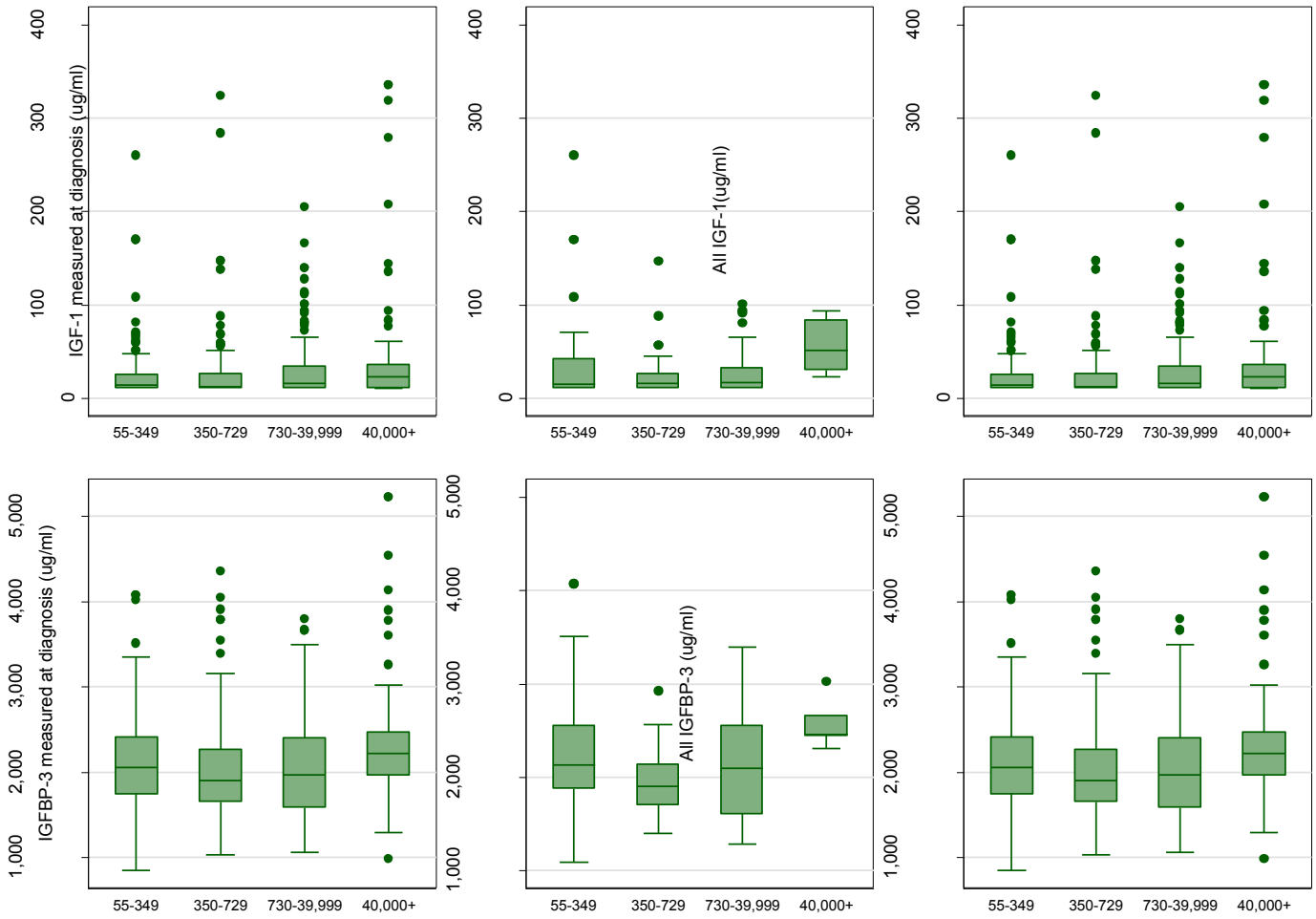

Figure 6: The relationship between serum IGF-1 and IGFBP-3 concentrations and the woman's annual salary in US dollars. Graphs on the far left show the relationship using IGF-1/IGFBP-3 samples taken at baseline, the graphs in the middle with samples taken at time of case diagnosis, and the graphs on the far right include all of the woman's samples.

The median baseline and diagnosis IGF-1 and IGFBP-3 serum concentration was compared between cases and controls (Table 4). The median concentrations of IGFs in women with HG-SIL or LG-SIL and controls were not significantly different either at baseline or at case diagnosis.

Table 4: Median serum IGF-1 and IGFBP-3 concentration and IGF-1:IGFBP-3 molar ratio of cases and controls. A nonparametric equality-of-medians test was conducted to determine if median IGF-1, IGFBP-3 or IGF-1:IGFBP-3 molar ratio, and p-values included.

Variable	Cases			Controls	P value
	HG-SIL(n=33)	LG-SIL(n=168)	Overall (n=201)	(n=402)	
	Median	Median	Median	Median	
IGF-1 Baseline	25.83	17.37	17.90	16.56	0.636
IGF-1 Diagnosis	12.59	36.99	33.119	27.12	0.056
IGFBP-3 Baseline	2402.59	2062.37	2074.24	2106.66	0.630
IGFBP-3 Diagnosis	2155.85	2394.91	2926.83	2358.55	0.441
IGF-1:IGFBP-3 Baseline	0.045	0.034	0.034	0.032	0.458
IGF-1:IGFBP-3 Diagnosis	0.028	0.057	0.054	0.047	0.068

A woman's IGF-I and IGFBP-3 serum concentration also changed over time (Figure 7). IGF levels measured during 1995 and 2001 are the lowest, where as samples taken during 1998 and 2002 are the highest.

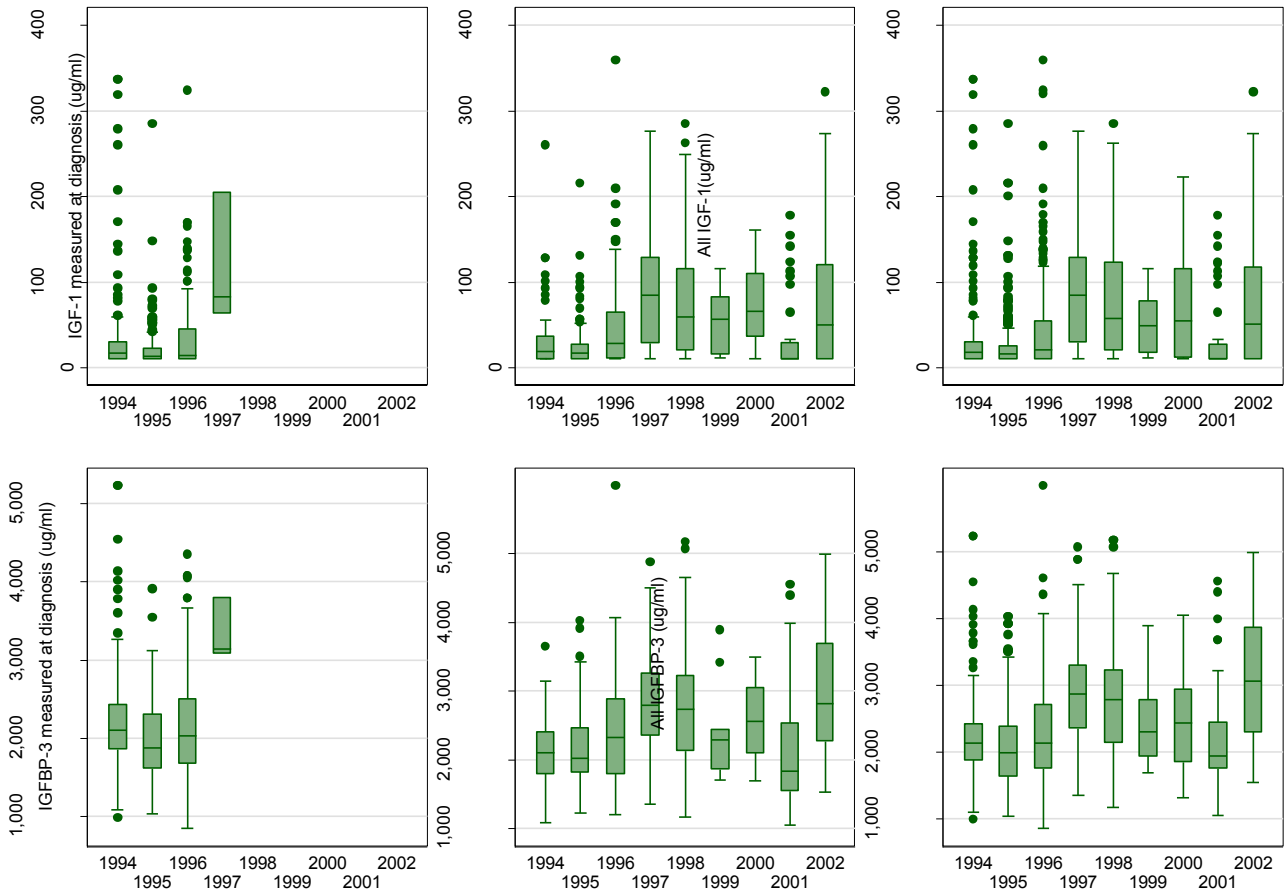

Figure 7: The relationship between serum IGF-1 and IGFBP-3 concentrations and the year the measurement was taken. Graphs on the far left show the relationship using IGF-1/IGFBP-3 samples taken at baseline, the graphs in the middle with samples taken at time of case diagnosis, and the graphs on the far right include all of a woman's measurements.

7.2 Determinants of IGF-1 and IGFBP-3

7.2.1 SOCIODEMOGRAPHIC AND DESIGN VARIABLES

Various sociodemographic, behavioural, reproductive, and nutritional variables were investigated as putative determinants of the IGF-1 and IGFBP-3 serum concentrations of the study participants.

Table 5 shows the associations between various sociodemographic and design variables and IGF-1/IGFBP-3 serum levels of cases and controls combined. Variables that were significantly related to IGF-I or IGFBP-3 concentration were a woman's age at time of measurement, educational attainment, annual salary, and year of baseline IGF measurement. When the analysis was done with just the controls (Table 5) results were comparable; age, salary and year of baseline measurement were significantly related to IGF measurements. Women who were 35 or older at the time of measurement had much lower levels of circulating IGF-I and IGFBP-3 at both baseline and at diagnosis than those who were 18-26 at baseline. There is a clear dose response relationship between age and IGF-I and IGFBP-3 serum concentration.

Educational attainment was significantly related to baseline IGF-1/IGFBP-3 levels for those women with a university or college education, but only when cases and controls were combined (OR= 4.03, 95% CI: 1.07-15.18 for IGF-1, OR=3.88, 95% CI: 1.03-14.64 for IGFBP-3). When the analysis was done with controls only, the relationship between education and IGF-I level was no longer significant, although the overall trend was the same for both groups (OR= 2.08, 95% CI: 0.49-8.83). Women who had the greatest educational attainment at baseline had higher protein levels than those in the lowest educational brackets. The effect of education on protein level was not seen in measurements taken at the time of SIL diagnosis.

A woman's annual salary had a similar effect of IGF-I levels as did education. Women earning the highest salaries at baseline, as compared to those earning the least amount of money had significantly higher IGF-I concentrations ($730-$39 999 US dollars a year OR=1.67, 95% CI:1.03-2.69, $40 000 or more US dollars a year OR=4.75, 95% CI:2.86-7.88). The effect remained significant in the $40,000 + group when the

analysis was restricted to controls only (OR=3.77, 95%CI: 2.03-7.00). The IGFBP-3 levels of women in the highest income quartile at baseline where also significantly higher than those of women in the lowest income quartile (OR=4.10, 95% CI: 2.47-6.80) and remained so when the analysis was restricted to controls (OR=3.51, 95% CI: 1.88-6.53). The effect of income on IGF-1/IGFBP-3 concentration was only seen in measurements taken at baseline.

The year that a woman's baseline serum concentration was measured had a significant impact on its level. If a woman had her baseline visit after 1995, her IGF-1 concentration was much higher (OR=2.04, 95% CI: 1.46-2.85). This was also true in the restricted analysis (OR=2.34, 95% CI: 1.55-3.51). The year of the baseline visit also impacted IGFBP-3 serum concentration; those who joined the study after 1995 had significantly higher concentrations of the binding protein at their baseline visit than those that joined prior to 1995 (OR=2.66, 95% CI: 1.90-3.73). This trend also held in the restricted analysis (OR= 3.35, 95% CI: 2.21-5.07).

Marital status, religion, and race were not significantly associated with the circulating serum concentrations of either IGF-1 or IGFBP-3.

Table 5: Age-adjusted univariate logistic regression between IGF-1 and IGFBP-3 measurements at either baseline or at time of SIL diagnosis and selected socio-economic, demographic, and design variables.

	Univariate Logistic Regression			
	IGF-1 Dichotomous at baseline	**IGF-1 Dichotomous at diagnosis**	**IGFBP-3 Dichotomous at baseline**	**IGFBP-3 Dichotomous at diagnosis**
	OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)
Age				
18-26	1.0 (ref)	1.0 (ref)	1.0 (ref)	1.0 (ref)
27-35	**0.45 (0.31-0.66)**	**0.63 (0.43-0.91)**	**0.62 (0.43-0.91)**	**0.67 (0.46-0.97)**
35+	**0.42 (0.28-0.65)**	**0.51 (0.33-0.78)**	**0.43 (0.28-0.66)**	**0.37 (0.24-0.56)**
Marital Status				
Single/Widowed/ Separated	1.0 (ref)	1.0 (ref)	1.0 (ref)	1.0 (ref)
Married/Living with partner	0.81 (0.55-1.19)	0.77 (0.52-1.13)	1.13 (0.76-1.66)	0.97 (0.66-1.44)
Education				
< Elementary School	1.0 (ref)	1.0 (ref)	1.0 (ref)	1.0 (ref)
Elementary School	0.96 (0.63-1.46)	0.82 (0.54-1.25)	0.96 (0.63-1.47)	1.08 (0.71-1.65)
High School	0.92 (0.54-1.58)	0.82 (0.48-1.40)	0.79 (0.46-1.36)	1.16 (0.68-2.00)
University/College	**4.03 (1.07-15.18)**	0.97 (0.33-2.87)	**3.88(1.03-14.64)**	1.22 (0.41-3.63)
Religion				
Catholic	1.0 (ref)	1.0 (ref)	1.0 (ref)	1.0 (ref)
Crente/Protestant	0.93 (0.58-1.46)	1.01 (0.64-1.59)	1.19 (0.75-1.88)	1.26 (0.80-2.00)
Other	0.78 (0.50-1.21)	0.89 (0.57-1.38)	1.02 (0.66-1.59)	1.01 (0.65-1.58)
Race				
White	1.0 (ref)	1.0 (ref)	1.0 (ref)	1.0 (ref)
Non-White	1.17 (0.84-1.63)	0.91 (0.65-1.27)	0.91 (0.65-1.27)	0.83 (0.60-1.16)
Salary (USD)				
55-349	1.0 (ref)	1.0 (ref)	1.0 (ref)	1.0 (ref)
350-729	1.13 (0.70-1.82)	0.82 (0.52-1.30)	0.82 (0.51-1.31)	0.92 (0.58-1.47)
730-39,999	**1.67 (1.03-2.69)**	0.88 (0.55-1.39)	1.24 (0.78-1.99)	0.91 (0.57-1.46)
40,000+	**4.75 (2.86-7.88)**	1.29 (0.81-2.06)	**4.10 (2.47-6.80)**	1.05 (0.66-1.68)
Year of Measurement				
Before 1995	1.0 (ref)	1.0 (ref)	1.0 (ref)	1.0 (ref)
After 1995	**2.04 (1.46-2.85)**	1.22 (0.88-1.68)	**2.66 (1.90-3.73)**	1.00 (0.72-1.38)

Table 6: Age-adjusted univariate logistic regression between the IGF-1 and IGFBP-3 measurements of controls only at either baseline or at time of SIL diagnosis and selected socio-economic, demographic, and design variables.

	Univariate Logistic Regression			
	IGF-1 Dichotomous at baseline	**IGF-1 Dichotomous at diagnosis**	**IGFBP-3 Dichotomous at baseline**	**IGFBP-3 Dichotomous at diagnosis**
	OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)
Age				
18-26	1.0 (ref)	1.0 (ref)	1.0 (ref)	1.0 (ref)
27-35	**0.38 (0.24-0.61)**	0.73 (46-1.15)	**0.56 (0.35-0.88)**	0.84 (0.53-1.32)
35+	**0.39 (0.23-0.66)**	**0.44 (0.26-0.75)**	**0.51 (0.30-0.87)**	**0.34 (0.20-0.58)**
Marital Status				
Single/Widowed/ Separated	1.0 (ref)	1.0 (ref)	1.0 (ref)	1.0 (ref)
Married/Living with partner	0.87 (0.52-1.43)	1.15 (0.70-1.90)	1.00 (0.61-1.65)	1.29 (0.78-2.13)
Education				
< Elementary School	1.0 (ref)	1.0 (ref)	1.0 (ref)	1.0 (ref)
Elementary School	0.71 (0.42-1.21)	0.78 (0.46-1.31)	0.76 (0.45-1.29)	1.05 (0.62-1.78)
High School	0.71 (0.37-1.39)	0.74 (0.38-1.43)	0.51 (0.26-0.99)	1.11 (0.57-2.15)
University/College	2.08 (0.49-8.83)	0.92 (0.24-3.49)	3.24 (0.64-16.5)	2.71 (0.64-11.5)
Religion				
Catholic	1.0 (ref)	1.0 (ref)	1.0 (ref)	1.0 (ref)
Crente/Protestant	0.89 (0.53-1.51)	1.13 (0.67-1.91)	1.25 (0.74-2.12)	1.23 (0.72-2.09)
Other	0.58 (0.33-1.02)	0.80 (0.46-1.40)	0.92 (0.53-1.60)	1.00 (0.58-2.09)
Race				
White	1.0 (ref)	1.0 (ref)	1.0 (ref)	1.0 (ref)
Non-White	1.09 (0.73-1.64)	0.83 (0.55-1.24)	0.98 (0.66-1.46)	0.78 (0.52-1.16)
Salary (USD)				
55-349	1.0 (ref)	1.0 (ref)	1.0 (ref)	1.0 (ref)
350-729	0.87 (0.48-1.59)	0.78 (0.43-1.39)	0.61 (0.34-1.10)	0.98 (0.55-1.76)
730-39,999	1.34 (0.74-2.44)	1.05 (0.58-1.89)	0.97 (0.54-1.75)	1.19 (0.66-2.16)
40,000+	**3.77 (2.03-7.00)**	1.40 (0.78-2.51)	**3.51 (1.88-6.53)**	1.17 (0.65-2.09)
Year of Measurement				
Before 1995	1.0 (ref)	1.0 (ref)	1.0 (ref)	1.0 (ref)
After 1995	**2.34 (1.55-3.51)**	1.16 (0.78-1.73)	**3.35 (2.21-5.07)**	1.00 (0.67-1.48)

7.2.2 Reproductive variables

Table 7 shows the associations between reproductive history variables and quartiles of IGF-1 and IGFBP-3 serum concentration (cases and controls combined). Of the reproductive variables examined, age at first menarche of 13 or greater was significantly associated with IGF-1 serum concentration. Oral contraceptive use of 6 years or greater was significantly associated with baseline IGFBP-3 serum concentration.

When the analysis was restricted to controls only, age at menarche and oral contraceptive use were no longer significantly associated with protein concentration (Table 7).

Having had prior Pap test (OR= 2.43, 95%CI: 1.11-5.32) and a STI other than HPV (OR=1.93, 95% CI: 1.08-3.43) became significantly associated with IGFBP-3 levels in the control group analysis.

Table 7: Age-Adjusted univariate logistic regression between IGF-1 and IGFBP-3 measurements at either baseline or at time of SIL diagnosis and selected reproductive history variables.

	Univariate Logistic Regression			
	IGF-1 Dichotomous at baseline	**IGF-1 Dichotomous at diagnosis**	**IGFBP-3 Dichotomous at baseline**	**IGFBP-3 Dichotomous at baseline**
	OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)
Pregnancies at Baseline				
0-1	1.0 (ref)	1.0 (ref)	1.0 (ref)	1.0 (ref)
2-3	1.01 (0.63-1.61)	0.76 (0.55-1.05)	1.48 (0.92-2.39)	1.03(0.74-1.42)
4-6	**0.55 (0.32-0.94)**	**0.63 (0.44-0.90)**	1.32 (0.77-2.25)	1.02 (0.71-1.46)
7+	1.00 (0.50-2.02)	0.94 (0.58-1.53)	1.54 (0.75-3.14)	1.17 (0.72-1.90)
Age at Menarche				
8-12	1.0 (ref)	1.0 (ref)	1.0 (ref)	1.0 (ref)
13+	**1.74 (1.24-2.44)**	**1.49 (1.18-1.87)**	1.39 (0.99-1.94)	1.23 (0.98-1.55)
STDs at Baseline				
No STDs	1.0 (ref)	1.0 (ref)	1.0 (ref)	1.0 (ref)
HPV	1.32 (0.64-2.76)	0.94 (0.54-1.64)	0.77 (0.36-1.63)	0.78 (0.44-1.36)
Other STD	0.84 (0.52-1.36)	1.09 (0.79-1.51)	0.65 (0.40-1.06)	0.94 (0.68-1.30)
Ever Pap Test				
No	1.0 (ref)	1.0 (ref)	1.0 (ref)	1.0 (ref)
Yes	1.50 (0.79-2.86)	1.11 (0.73-1.70)	1.70 (0.89-3.25)	1.32 (0.86-2.01)
Oral Contraceptive Use				
Never	1.0 (ref)	1.0 (ref)	1.0 (ref)	1.0 (ref)
< 6 years	0.99 (0.64-1.53)	0.75 (0.49-1.16)	1.55 (0.99-2.40)	0.92 (0.59-1.42)
> 6 years	0.89 (0.53-1.49)	0.82 (0.49-1.38)	**2.00 (1.18-3.40)**	0.82 (0.49-1.38)

Table 8: Age-Adjusted univariate logistic regression of IGF-1 and IGFBP-3 measurements of control women at either baseline or at time of SIL diagnosis and selected reproductive history variables.

	Univariate Logistic Regression			
	IGF-1 Dichotomous at baseline	**IGF-1 Dichotomous at diagnosis**	**IGFBP-3 Dichotomous at baseline**	**IGFBP-3 Dichotomous at baseline**
	OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)
Pregnancies at Baseline				
0-1	1.0 (ref)	1.0 (ref)	1.0 (ref)	1.0 (ref)
2-3	0.66 (0.37-1.19)	1.13 (0.63-2.01)	1.27 (0.71-2.27)	0.90 (0.50-1.62)
4-6	0.56 (0.29-1.09)	1.19 (0.62-2.30)	1.30 (0.68-2.51)	1.28 (0.66-2.47)
7+	0.51 (0.21-1.22)	1.48 (0.63-3.52)	1.38 (0.58-3.26)	1.32 (0.55-3.14)
Age at Menarche				
8-12	1.0 (ref)	1.0 (ref)	1.0 (ref)	1.0 (ref)
13+	1.37 (0.91-2.05)	1.32 (0.87-1.98)	1.25 (0.84-1.86)	1.29 (0.86-1.94)
STDs at Baseline				
No STDs	1.0 (ref)	1.0 (ref)	1.0 (ref)	1.0 (ref)
HPV	0.92 (0.33-2.56)	1.87 (0.65-5.36)	0.84 (0.31-2.32)	1.36 (0.48-3.82)
Other STD	0.68 (0.39-1.20)	1.50 (0.85-2.62)	0.56 (0.32-0.98)	**1.93 (1.08-3.43)**
Ever Pap Test				
No	1.0 (ref)	1.0 (ref)	1.0 (ref)	1.0 (ref)
Yes	0.96 (0.93-0.98)	0.80 (0.37-1.70)	**2.43 (1.11-5.32)**	1.41 (0.67-3.01)
Oral Contraceptive Use				
Never	1.0 (ref)	1.0 (ref)	1.0 (ref)	1.0 (ref)
< 6 years	0.87 (0.51-1.51)	0.85 (0.50-1.45)	1.39 (0.81-2.39)	1.00 (0.58-1.71)
> 6 years	0.85 (0.45-1.61)	0.74 (0.39-1.41)	1.82 (0.96-3.48)	0.78 (0.41-1.49)

7.2.3 BEHAVIOURAL & LIFESTYLE VARIABLES

The use of alcohol was significantly related to IGF-I and IGFBP-3 serum concentrations at diagnosis (OR=1.36, 95% CI: 1.06-1.73 for IGF-I and OR=1.47, 95% CI: 1.15-1.88 for IGFBP-3) (Table 9). This association remained when the analysis was restricted to controls only (OR=1.53 95% CI: 1.00-2.34 for IGF-I and OR=1.65, 95% CI: 1.07-2.52 for IGFBP-3). Women who were less than 15 years of age at first intercourse were significantly more likely to have lower IGFBP-3 levels at both baseline and time of diagnosis, and lower IGF-I levels at time of diagnosis, as compared to those women who over age 20 at the time of first intercourse. This association held only for women in the lowest age quartile. When the analysis was restricted to controls only, those in the

youngest age quartile had significantly lower levels of IGF-I at baseline (OR= 0.45, 95% CI: 0.24-0.84), and the association was no longer significant in regards to IGFBP-3.

Women who were former smokers had lower IGF-1 levels at both baseline and at time of diagnosis, as compared to women who had never smoked (OR=0.61, 95% CI: 0.38-0.98 at baseline, OR= 0.64, 95% CI:0.47-0.88 at time of diagnosis). When restricted to controls only, former smoking status was still significantly associated with IGF-I concentration at baseline (OR=0.50, 95% CI: 0.29-0.87). In the analysis including both cases and controls, being a former smoker was also associated with IGFBP-3 concentration at diagnosis (OR=0.68, 95% CI: 0.50-0.93), but this association did not persist in the controls only analysis. Duration of smoking, number of cigarettes/day and a woman's cumulative number of sexual partners were not related to IGF-I or IGFBP-3 serum concentrations.

Table 9: Age-adjusted univariate logistic regression of IGF-1 and IGFBP-3 measurements at either baseline or at time of SIL diagnosis and selected lifestyle and behaviour variables.

	IGF-1 Dichotomous at baseline	IGF-1 Dichotomous at diagnosis	IGFBP-3 Dichotomous at baseline	IGFBP-3 dichotomous at diagnosis
	OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)
Smoking History				
Never	1.0 (ref)	1.0 (ref)	1.0 (ref)	1.0 (ref)
Current	1.11 (0.71-1.72)	1.01 (0.74-1.37)	1.11 (0.71-1.73)	0.95 (0.70-1.28)
Former	**0.61 (0.38-0.98)**	**0.64 (0.47-0.88)**	0.75 (0.47-1.19)	**0.68 (0.50-0.93)**
Duration of Smoking	0.96(0.94-0.99)	1.00 (0.99-1.01)	1.00 (0.98-1.02)	1.00 (0.98-1.01)
Number of Cigarettes/Day				
Never/Former	1.0(ref)	1.0 (ref)	1.0 (ref)	1.0 (ref)
Up to 10/day	1.14 (0.75-1.73)	1.12 (0.84-1.50)	1.08 (0.71 – 1.64)	1.03 (0.77-1.38)
>10 /day	1.10 (0.69-1.74)	1.17(0.84-1.61)	1.12 (0.71-1.79)	1.24 (0.90-1.72)
Age at First Intercourse				
0-15	0.60 (0.36-1.02)	**0.68 (0.47-0.97)**	**0.54 (0.32-0.93)**	**0.63 (0.44-0.90)**
16-17	0.80 (0.47-1.36)	0.85 (0.60-1.23)	0.71 (0.42 – 1.20)	0.85 (0.59-1.22)
18-19	0.90 (0.51-1.56)	0.83 (0.57-1.22)	0.75 (0.43-1.32)	0.82 (0.56-1.20)
20-50	1.0 (ref)	1.0 (ref)	1.0 (ref)	1.0 (ref)
Number of Lifetime Sex Partners				
0-1	1.0(ref)	1.0 (ref)	1.0 (ref)	1.0 (ref)
2	1.02 (0.66-1.58)	1.00 (0.73-1.35)	0.94 (0.60-1.46)	1.01 (0.74-1.37)
3	0.84 (0.52-1.38)	0.81 (0.58-1.14)	0.80 (0.49-1.31)	0.78 (0.56-1.10)
4+	0.70 (0.44-1.09)	0.77 (0.57-1.04)	0.85 (0.55-1.32)	0.84 (0.62-1.13)
Ever Drink Alcohol				
No	1.0 (ref)	1.0 (ref)	1.0 (ref)	1.0 (ref)
Yes	1.01 (0.71-1.44)	**1.36 (1.06-1.73)**	1.14 (0.80-1.62)	**1.47 (1.15-1.88)**

Table 10: Age-adjusted univariate logistic regression of control women's IGF-1 and IGFBP-3 measurements at either baseline or at time of SIL diagnosis and selected lifestyle and behaviour variables.

	IGF-1 Dichotomous at baseline	IGF-1 Dichotomous at diagnosis	IGFBP-3 Dichotomous at baseline	IGFBP-3 dichotomous at diagnosis
	OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)
Smoking History				
Never	1.0 (ref)	1.0 (ref)	1.0 (ref)	1.0 (ref)
Current	1.05 (0.61-1.81)	0.98 (0.57-1.69)	1.32 (0.77-2.28)	0.76 (0.44-1.31)
Former	**0.50 (0.29-0.87)**	0.93 (0.55-1.59)	0.78 (0.46-1.33)	0.99 (0.58-1.69)
Duration of Smoking	1.01 (0.98-1.03)	0.99 (0.97-0.99)	1.01 (0.98-1.03)	0.99 (0.97-1.02)
Number of Cigarettes/Day				
Never/Former	1.0(ref)	1.0 (ref)	1.0 (ref)	1.0 (ref)
Up to 10/day	1.16 (0.69-1.94)	1.11 (0.66-1.86)	1.00 (0.60-1.68)	1.11 (0.66-1.87)
>10 /day	0.92 (0.50-1.68)	1.00 (0.55-1.84)	1.24 (0.68-2.27)	0.77 (0.42-1.42)
Age at First Intercourse				
0-15	**0.45 (0.24-0.84)**	0.60 (0.32-1.10)	0.72 (0.39-1.31)	0.53 (0.29-0.99)
16-17	0.60 (0.32-1.11)	0.63 (0.34-1.17)	0.99 (0.54-1.83)	0.74 (0.40-1.37)
18-19	0.98 (0.49-1.93)	0.58 (0.29-1.14)	0.81 (0.42-1.59)	0.69 (0.35-1.36)
20-50	1.0 (ref)	1.0 (ref)	1.0 (ref)	1.0 (ref)
Number of Lifetime Sex Partners				
0-1	1.0(ref)	1.0 (ref)	1.0 (ref)	1.0 (ref)
2-3	0.70 (0.44-1.10)	0.82 (0.52-1.30)	0.85 (0.54-1.33)	1.03 (0.66-1.63)
4+	0.52 (0.30-0.88)	1.23 (0.73-2.08)	0.74 (0.44-1.26)	1.31 (0.77-2.23)
Ever Drink Alcohol				
No	1.0 (ref)	1.0 (ref)	1.0 (ref)	1.0 (ref)
Yes	0.93 (0.61-1.42)	**1.53 (1.00-2.34)**	1.17 (0.77-1.79)	**1.65 (1.07-2.52)**

7.2.4 NUTRITIONAL UPTAKE VARIABLES

There were several questions on the second study questionnaire that aimed to quantify the participants' nutritional status and food intake. Vitamin C (orange, lemon, or vitamin C supplements), leafy greens and liver intake in any amount was not related to IGF-I or IGFBP-3 serum concentration (Table 11 and Table 12). Beta-carotene (carrot, pumpkin and papaya) intake in the highest tertile was associated with increased IGFBP-3 concentrations at baseline only (OR= 1.69, 95% CI: 1.11-2.58) in the full analysis, as did vitamin B supplements in the highest tertile OR= 2.13, 95% CI: 1.10-4.12, but neither of these two associations persisted in the analysis that was restricted to controls only. Baseline IGFBP-3 concentration was associated with dairy intake in a dose response manner; tertile 2 OR=1.62, 95% CI: 1.07-2.46, tertile 3 OR=1.81, 95% CI: 1.19-2.77, and the association persisted in the restricted analysis (OR=1.97, 95% CI: 1.18-3.28). IGFBP-3 at diagnosis was associated with multi-vitamin use in the highest tertile OR= 1.52, 95% CI: 1.06-2.19 in the full analysis only. IGF-1 was only associated with multi-vitamin use in the highest tertile, at baseline only OR=1.49, 95% CI: 1.03-2.14.

Table 11: Age-adjusted univariate logistic regression between IGF-1 and IGFBP-3 measurements at either baseline or at time of SIL diagnosis and selected nutritional intake variables.

	IGF-1 Dichotomous at baseline	IGF-1 Dichotomous at diagnosis	IGFBP-3 Dichotomous at baseline	IGFBP-3 dichotomous at diagnosis
	OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)
Carrot/Pumpkin/Papaya				
T1	1.0 (ref)	1.0 (ref)	1.0 (ref)	1.0 (ref)
T2	0.88 (0.59-1.31)	0.89 (0.60-1.32)	0.83 (0.56-1.24)	1.13 (0.76-1.68)
T3	1.12 (0.74-1.69)	0.91 (0.60-1.38)	**1.69 (1.11-2.58)**	1.19 (0.78-1.81)
Orange/Lemon/Vitamin C Supplement				
T1	1.0(ref)	1.0 (ref)	1.0 (ref)	1.0 (ref)
T2	0.70 (0.45-1.08)	0.89 (0.57-1.37)	1.08 (0.70-1.67)	0.94 (0.61-1.45)
T3	1.04 (0.70-1.55)	0.95 (0.64-1.42)	1.42 (0.95-2.13)	0.95 (0.64-1.43)
Leafy Greens				
T1	1.0(ref)	1.0 (ref)	1.0 (ref)	1.0 (ref)
T2	0.90 (0.59-1.38)	0.72 (0.48-1.10)	0.96 (0.63-1.46)	0.81 (0.53-1.23
T3	0.75 (0.49-1.16)	0.76 (0.49-1.18)	0.72 (0.46-1.11)	0.70 (0.45-1.08)
Dairy				
T1	1.0(ref)	1.0 (ref)	1.0 (ref)	1.0 (ref)
T2	0.86 (0.57-1.30)	1.21 (0.80-1.82)	**1.62 (1.07-2.46)**	1.09 (0.72-1.64)
T3	1.25 (0.82-1.90)	1.31 (0.86-1.99)	**1.81 (1.19-2.77)**	1.24 (0.82-1.89)
Liver				
T1	1.0(ref)	1.0 (ref)	1.0 (ref)	1.0 (ref)
T2	0.88 (0.57-1.37)	1.06 (0.69-1.63)	1.36 (0.88-2.10)	1.20 (0.78-1.86)
T3	0.70 (0.45-1.09)	0.98 (0.63-1.51)	1.34 (0.86-2.10)	1.54 (0.99-2.41)
Vitamin B Supplements				
T1	1.0(ref)	1.0 (ref)	1.0 (ref)	1.0 (ref)
T2	1.12 (0.55-2.27)	0.36 (0.17-0.78)	0.57 (0.28-1.19)	0.41 (0.19-0.86)
T3	1.51 (0.80-2.82)	1.12 (0.60-2.07)	**2.13 (1.10-4.12)**	1.10 (0.59-2.04)
Multivitamin				
T1	1.0(ref)	1.0 (ref)	1.0 (ref)	1.0 (ref)
T2	0.94 (0.49-1.80)	1.12 (0.59-2.15)	0.71 (0.37-1.38)	1.20 (0.63-2.31)
T3	**1.49 (1.03-2.14)**	1.24 (0.87-1.79)	1.40 (0.98-2.02)	**1.52 (1.06-2.19)**

Table 12: Age-adjusted univariate logistic regression of control women's IGF-1 and IGFBP-3 measurements at either baseline or at time of SIL diagnosis and selected nutritional intake variables.

	IGF-1 Dichotomous at baseline	IGF-1 Dichotomous at diagnosis	IGFBP-3 Dichotomous at baseline	IGFBP-3 dichotomous at diagnosis
	OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)
Carrot/Pumpkin/Papaya				
T1	1.0 (ref)	1.0 (ref)	1.0 (ref)	1.0 (ref)
T2	0.77 (0.47-1.26)	0.95 (0.58-1.55)	0.74 (0.45-1.21)	1.16 (0.71-1.90)
T3	1.04 (0.62-1.72)	0.98 (0.59-1.63)	1.57 (0.94-2.62)	1.20 (0.72-1.99)
Orange/Lemon/Vitamin C Supplement				
T1	1.0(ref)	1.0 (ref)	1.0 (ref)	1.0 (ref)
T2	0.64 (0.37-1.10)	0.94 (0.55-1.60)	1.13 (0.66-1.94)	0.89 (0.52-1.53)
T3	1.04 (0.63-1.71)	1.06 (0.65-1.74)	1.52 (0.93-2.49)	1.01 (0.62-1.67)
Leafy Greens				
T1	1.0(ref)	1.0 (ref)	1.0 (ref)	1.0 (ref)
T2	0.95 (0.56-1.61)	0.86 (0.51-1.45)	0.90 (0.53-1.53)	0.91 (0.54-1.55)
T3	0.71 (0.41-1.24)	0.92 (0.53-1.60)	0.64 (0.37-1.11)	0.74 (0.43-1.29)
Dairy				
T1	1.0(ref)	1.0 (ref)	1.0 (ref)	1.0 (ref)
T2	0.86 (0.52-1.43)	1.51 (0.91-2.50)	1.42 (0.86-2.35)	1.24 (0.75-2.06)
T3	1.42 (0.85-2.37)	**1.68 (1.01-2.80)**	**1.97 (1.18-3.28)**	1.53 (0.92-2.54)
Liver				
T1	1.0(ref)	1.0 (ref)	1.0 (ref)	1.0 (ref)
T2	1.17 (0.68-2.00)	1.23 (0.72-2.11)	1.43 (0.84-2.46)	1.28 (0.74-2.21)
T3	0.78 (0.45-1.36)	1.11 (0.64-1.93)	1.38 (0.80-2.39)	**1.76 (1.01-3.07)**
Vitamin B Supplements				
T1	1.0(ref)	1.0 (ref)	1.0 (ref)	1.0 (ref)
T2	1.58 (0.68-3.67)	0.22 (0.08-0.62)	0.75 (0.33-1.72)	0.25 (0.10-0.67)
T3	1.60 (0.72-3.57)	1.68 (0.74-3.79)	**2.89 (1.19-7.03)**	1.81 (0.78-4.17)
Multivitamin				
T1	1.0(ref)	1.0 (ref)	1.0 (ref)	1.0 (ref)
T2	1.05 (0.49-2.28)	1.13 (0.52-2.43)	0.54 (0.25-1.19)	1.02 (0.47-2.21)
T3	**1.66 (1.05-2.60)**	1.06 (0.68-1.67)	1.19 (0.76-1.86)	1.45 (0.93-2.28)

7.2.5 Determinants of IGF-I and IGFBP–3 in multivariate analysis

In the univariate analyses, several putative predictors of baseline IGF-I serum concentration were identified, including woman's age, her annual salary, education, the year of her baseline measurement, her age at menarche, and smoking history. However, these same potential determinants where generally no longer significantly associated with IGF-1 level at diagnosis (Table 13). Variables were put into a backward-selection estimation model, and age, salary, education, and age at menarche were significantly related to IGF-1 serum concentrations at baseline, and only age was a significant determinant of IGF-1 at diagnosis in the multivariate model. In the restricted analysis of controls only, smoking history, salary and age were significantly associated with baseline IGF-I level, and again, only age was significantly associated with IGF-I level at diagnosis.

Baseline IGFBP-3 serum concentration was significantly associated with age, educational attainment, annual salary, year of measurement and dairy intake in the univariate analysis. In the multivariate analysis, only age and salary changed the OR for the relationship by more than 5%. At time of diagnosis, only age was kept in the model. In the restricted analysis, age, salary, and alcohol consumption were associated with an increased IGFBP-3 level at baseline, and age and alcohol consumption were significantly related to IGFBP-3 level at diagnosis.

Table 13: Multiple logistic regression of IGF-1 and IGFBP-3 measured at either baseline or at time of SIL diagnosis and selected variables

Outcome	Variable	Cases and Controls OR (95% CI)		Controls Only OR (95% CI)	
		Baseline	Diagnosis	Baseline	Diagnosis
IGF-1	**Age**				
	18-26	1.0 (ref)	1.0 (ref)	1.0 (ref)	1.0 (ref)
	27-35	**0.37 (0.24-0.55)**	**0.63 (0.44-0.92)**	**0.31 (0.18-0.53)**	**0.38 (0.24-0.61)**
	36+	**0.29 (0.17-0.47)**	**0.49 (0.32-0.76)**	**0.32 (0.17-0.59)**	**0.39 (0.23-0.66)**
	Salary				
	55-349	1.0 (ref)		1.0 (ref)	
	350-729	1.21 (0.74-1.99)		1.15 (0.58-2.27)	
	730-39,999	**1.82 (1.11-3.01)**		1.38 (0.69-2.75)	
	40,000+	**5.09 (3.01-8.62)**		**5.08 (2.55-10.13)**	
	Education				
	< Elementary School	1.0 (ref)			
	Elementary School	1.05 (0.67-1.64)			
	High School	0.91 (0.51-1.63)			
	University/College	**4.61 (1.14-18.68)**			
	Age at Menarche				
	8-12	1.0 (ref)			
	13+	**1.70 (1.20-2.43)**			
	Smoking History				
	Never			1.0 (ref)	
	Current			0.78 (0.42-1.42)	
	Former			**0.48 (0.27-0.86)**	
IGFBP-3	Age				
	18-26	1.0 (ref)	1.0 (ref)	1.0 (ref)	1.0 (ref)
	27-35	**0.60 (0.40-0.88)**	**0.67 (0.46-0.97)**	**0.54 (0.33-0.88)**	0.83 (0.52-1.32)
	36+	**0.34 (0.21-0.53)**	**0.37 (0.24-0.56)**	**0.45 (0.25-0.80)**	**0.32 (0.18-0.56)**
	Salary				
	55-349	1.0 (ref)		1.0 (ref)	
	350-729	0.84 (0.52-1.34)		0.63 (0.34-1.15)	
	730-39,999	1.23 (0.77-1.98)		1.02 (0.55-1.87)	
	40,000+	**4.26 (2.55-7.09)**		**4.43 (2.31-8.50)**	
	Ever Drink Alcohol				
	No			1.0 (ref)	1.0 (ref)
	Yes			**1.76 (1.11-2.79)**	**1.82 (1.16-2.83)**

7.2.6 IGF-I AND IGFBP-3 SERUM CONCENTRATION AND RISK OF SIL

For the primary study aim, IGF-I and IGFBP-3 serum concentrations and IGF-I:IGFBP-3 molar ratio were not significantly associated with risk of cumulative SIL (Table 14). In general, the ORs for the association between IGF-I and IGF-I/IGFBP-3 molar ratio and risk of SIL were greater than 1.0, however, none of the 95% CIs were significant.

Table 14: Age-adjusted conditional logistic regression of any SIL and IGF-1/IGFBP-3. Measurements used are either from baseline, time of SIL diagnosis, or the highest measurement recorded. Sets are based on risk time and age, with two controls per case.

Variable	OR (95% CI) Baseline	OR (95% CI) Diagnosis	OR (95% CI) Highest measurement	OR (95% CI) Prevalent cases only
IGF-1				
Q1	1.0 (ref)	1.0 (ref)	1.0 (ref)	1.0 (ref)
Q2	0.98 (0.60-1.60)	1.50 (0.85-2.64)	1.38 (0.71-2.71)	0.89 (0.28-2.62)
Q3	1.33 (0.83-2.11)	**1.73 (1.03-2.93)**	1.66 (0.88-3.15)	1.44 (0.50-4.14)
Q4	1.12 (0.59-2.11)	1.46 (0.85-2.51)	1.52 (0.79-2.92)	2.91 (0.75-11.26)
IGFBP-3				
Q1	1.0 (ref)	1.0 (ref)	1.0 (ref)	1.0 (ref)
Q2	1.26 (0.77-2.07)	1.33(0.81-2.19)	1.06 (0.66-1.71)	1.45 (0.48-4.11)
Q3	1.02 (0.61-1.71)	1.14 (0.68-1.92)	0.92 (0.55-1.56)	1.18 (0.39-3.58)
Q4	1.26 (0.72-2.21)	1.11 (0.64-1.92)	0.96 (0.55-1.69)	1.67 (0.50-5.62)
IGF-1:IGFBP-3				
Q1	1.0 (ref)	1.0 (ref)	1.0 (ref)	1.0 (ref)
Q2	1.15 (0.72-1.82)	1.26 (0.74-2.13)	1.37 (0.72-2.60)	1.49 (0.59-3.78)
Q3	1.42 (0.89-2.27)	1.58 (0.94-2.66)	1.51 (0.80-2.85)	1.85 (0.69-4.96)
Q4	0.93 (0.49-1.75)	1.26 (0.74-2.16)	1.32 (0.68-2.54)	2.09 (0.59-7.45)

Serum concentration of IGF-I and IGFBP-3 and IGF-I/IGFBP-3 molar ratio were also not significantly associated with risk of HG-SIL, however this was expected due to the significant reduction in statistical power that resulted from restricting the analysis (HG-SIL cases=33) (Table 15). At time of diagnosis, higher levels of IGF-I and a higher ratio of IGF-I to IGFBP-3 resulted in ORs below the null, with the opposite trend being observed in the baseline measurements. Serum concentration IGFBP-3 levels in the

highest quartiles appeared to be protective at both baseline and time of diagnosis measurements, but again, none of the confidence intervals are significant.

Table 15: Conditional Logistic Regression of HG-SIL and IGF-1/IGFBP-3. Measurements used are either from baseline, time of SIL diagnosis, or the highest measurement recorded. Sets are based on risk time and age, with two controls per case.

Variable	OR (95% CI) Baseline	OR (95% CI) Diagnosis	OR (95% CI) Highest measurement
IGF-1			
Q1	1.0 (ref)	1.0 (ref)	1.0 (ref)
Q2	1.12 (0.27-4.61)	0.90 (0.29-2.75)	0.79 (0.21-2.91)
Q3	2.71 (0.71-10.32)	0.73 (0.19-2.78)	1.44 (0.32-6.41)
Q4	1.36 (0.28-6.72)	0.36 (0.09-1.50)	0.65 (0.13-3.35)
IGFBP-3			
Q1	1.0 (ref)	1.0 (ref)	1.0 (ref)
Q2	0.89 (0.22-3.57)	1.46 (0.45-4.71)	0.44 (0.11-1.79)
Q3	0.35 (0.06-2.00)	1.11 (0.28-4.42)	1.03 (0.21-5.10)
Q4	0.92 (0.20-4.23)	0.56 (0.11-2.84)	0.24 (0.03-1.79)
IGF-1:IGFBP-3			
Q1	1.0 (ref)	1.0 (ref)	1.0 (ref)
Q2	1.21 (0.36-4.04)	0.63 (0.21-1.94)	1.06 (0.30-3.71)
Q3	2.31 (0.62-8.62)	0.49 (0.13-1.82)	1.34 (0.27-6.61)
Q4	1.12 (0.21-5.96)	0.30 (0.07-1.35)	0.69 (0.11-4.45)

7.2.7 IGF-I, IGFBP-3 AND RISK OF CUMULATIVE HPV INFECTION

Maximum serum IGF-I concentrations in the two highest quartiles were significantly associated with an increased risk of acquiring any HPV infection over the course of follow-up. When the analysis was broken down by type of HPV infection, circulating IGF-I in the highest two quartiles was significantly associated with an increased risk of acquiring a LR-HPV infection, other than types 6 or 11 (Table 16). Circulating IGFBP-3 was also associated with HPV infection, however; only those with a concentration in the third quartile had a significantly increased risk of acquiring any HPV

infection and any LR-HPV infection. Power in the fourth quartile may have been too low for the positive association to reach statistical significance. Those in the highest quartile of circulating IGFBP-3 also had an increased risk of developing a HR-HPV infection, as compared to those with the lowest levels of the binding protein.

Women with IGF-1/IGFBP-3 molar ratios greater than the reference category had a significantly increased chance of developing any HPV infection, and when the analysis was broken down by HPV type, these women were significantly more likely to acquire a LR-HPV infection. This is the same trend that was seen when looking at IGF-I alone.

Table 16: Age and salary adjusted logistic regression of the maximum IGF value and risk of HPV infection

	Any HPV	HR-HPV	HPV-16	LR-HPV	HPV-6/11
IGF-1					
Q1	1.0 (ref)	1.0 (ref)	1.0 (ref)	1.0 (ref)	1.0 (ref)
Q2	1.33 (0.71-2.48)	0.88 (0.45-1.70)	0.47 (0.17-1.28)	2.26 (0.92-5.54)	0.55 (0.09-3.43)
Q3	**1.88(1.08-3.28)**	1.38 (0.78-2.44)	0.64 (0.29-1.43)	**2.57 (1.13-5.85)**	0.87 (0.21-3.61)
Q4	**1.84(1.08-3.15)**	1.45 (0.84-2.52)	0.75 (0.35-1.59)	**3.00 (1.35-6.67)**	1.08 (0.28-4.11)
IGFBP-3					
Q1	1.0 (ref)	1.0 (ref)	1.0 (ref)	1.0 (ref)	1.0 (ref)
Q2	1.19 (0.75-1.91)	1.24 (0.76-2.04)	0.69 (0.31–1.49)	1.15 (0.63-2.10)	0.80 (0.24-2.63)
Q3	**1.68 (1.05-2.69)**	1.47 (0.90-2.41)	1.31 (0.65-2.64)	**1.83 (1.04-3.23)**	0.83 (0.25-2.73)
Q4	1.37 (0.85-2.23)	**1.69 (1.03-2.78)**	0.94 (0.45-1.99)	1.34 (0.74-2.43)	0.85 (0.26-2.76)
IGF-1: IGFBP-3					
Q1	1.0 (ref)	1.0 (ref)	1.0 (ref)	1.0 (ref)	1.0 (ref)
Q2	**1.83 (1.02-3.28)**	1.15 (0.62-2.12)	0.51 (0.20-1.29)	**3.91 (1.61-9.53)**	0.99 (0.21-4.65)
Q3	1.66 (0.96-2.85)	1.29 (0.74-2.26)	0.69 (0.32-1.50)	**2.86 (1.20-6.79)**	0.84 (0.20-3.56)
Q4	**1.94 (1.15-3.28)**	1.43 (0.83-2.46)	0.71 (0.33-1.49)	**3.91 (1.69-9.07)**	1.10 (0.28-4.26)

Table 17: Age and salary adjusted logistic regression of the maximum IGF value of controls and risk of HPV infection

	Any HPV	HR-HPV	HPV-16	LR-HPV	HPV-6/11
IGF-1					
Q1	1.0 (ref)	1.0 (ref)	1.0 (ref)	1.0 (ref)	1.0 (ref)
Q2	1.43 (0.57-3.58)	0.69 (0.22-2.13)	0.54 (0.08-3.45)	2.24 (0.65-7.80)	--
Q3	1.88 (0.83-4.25)	1.04 (0.41-2.67)	0.39 (0.07-2.07)	2.39 (0.76-7.56)	1.37 (0.14-13.56)
Q4	**2.38 (1.08-5.22)**	1.71 (0.71-4.13)	1.25 (0.32-4.90)	2.88 (0.95-8.75)	1.45 (0.16-13.50)
IGFBP-3					
Q1	1.0 (ref)	1.0 (ref)	1.0 (ref)	1.0 (ref)	1.0 (ref)
Q2	1.19 (0.61-2.34)	1.92 (0.83-4.42)	1.34 (0.35-5.11)	0.53 (0.21-1.33)	0.47 (0.07-3.13)
Q3	**1.93 (1.01-3.69)**	1.81 (0.79-4.17)	1.37 (0.36-5.18)	1.80 (0.86-3.78)	0.79 (0.14-4.39)
Q4	1.62 (0.83-3.17)	**2.50 (1.09-5.72)**	1.05 (0.26-4.31)	0.87 (0.38-1.99)	0.83 (0.16-4.24)
IGF-1: IGFBP-3					
Q1	1.0 (ref)	1.0 (ref)	1.0 (ref)	1.0 (ref)	1.0 (ref)
Q2	2.09 (0.90-4.90)	1.09 (0.40-2.93)	0.40 (0.07-2.34)	2.86 (0.85-9.56)	0.70 (0.04-12.01)
Q3	1.78 (0.80-3.93)	1.07 (0.43-2.65)	0.48 (0.11-2.08)	2.66 (0.84-8.36)	1.16 (0.11-12.08)
Q4	**2.57 (1.20-5.50)**	1.67 (0.71-3.90)	0.90 (0.25-3.18)	**3.35 (1.10-10.17)**	1.77 (0.19-16.24)

7.2.8 IGF-I AND IGFBP-3 AND TIME UNTIL A LESION EVENT IN CASES

A conditional time to SIL analysis was conducted to determine if there were any differences in the time until SIL was detected amongst those who did develop a lesion, depending on their IGF-I or IGFBP-3 serum concentration or IGF-I/IGFBP-3 molar ratios. Figure 8 shows the Kaplan-Meier curves of the time until event analysis. Figure 9 shows that only by using IGF-I measurements taken at baseline does there appear to be a significant difference in the time until SIL event diagnosis (p =0.0013), as stratified by IGF-I level (Table 18). Once prevalent cases are removed from the analyses, there is no significant difference between time until SIL event in those above or below the median IGF-I concentration.

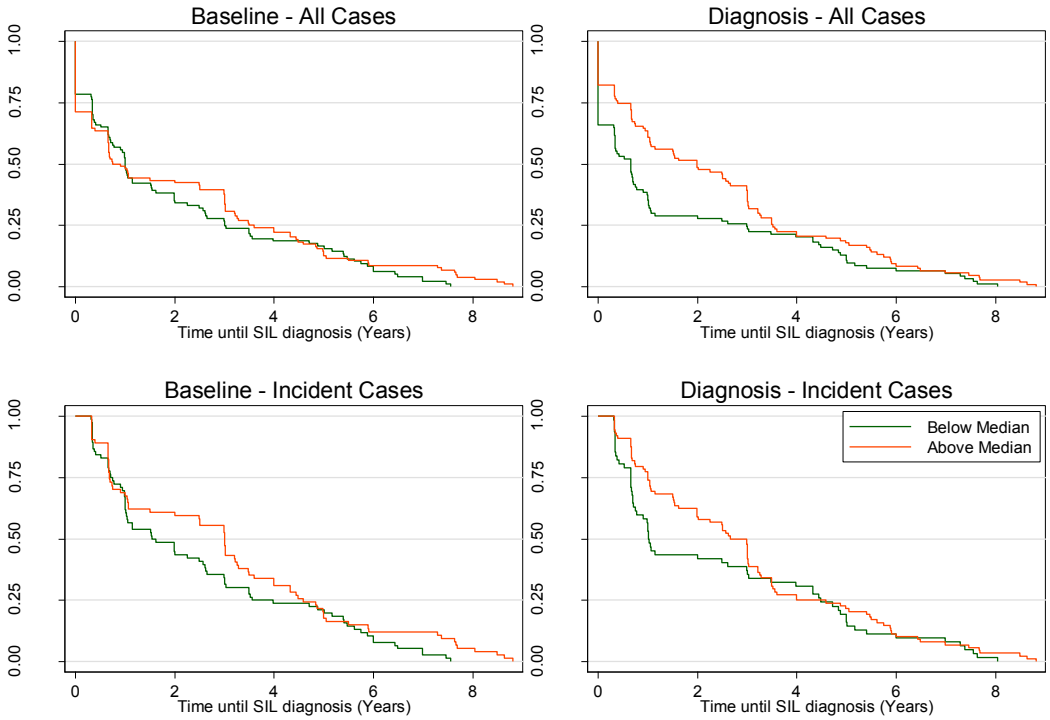

Figure 8: Kaplan-Meier analysis of time to SIL diagnosis stratified as being either above or below the median serum IGF-1 concentration. IGF-1 concentrations used in the analysis are either the woman's baseline measurement (left) or her measurement at time of diagnosis (right). The bottom graphs represent only incident SIL cases.

We also wanted to determine whether IGFBP-3 level influence time until SIL detection. Figure 9 shows the Kaplan-Meier analysis of time to SIL event stratified by IGFBP-3 level. There was no significant difference in the length of time before a SIL in any of the groups (Table 18).

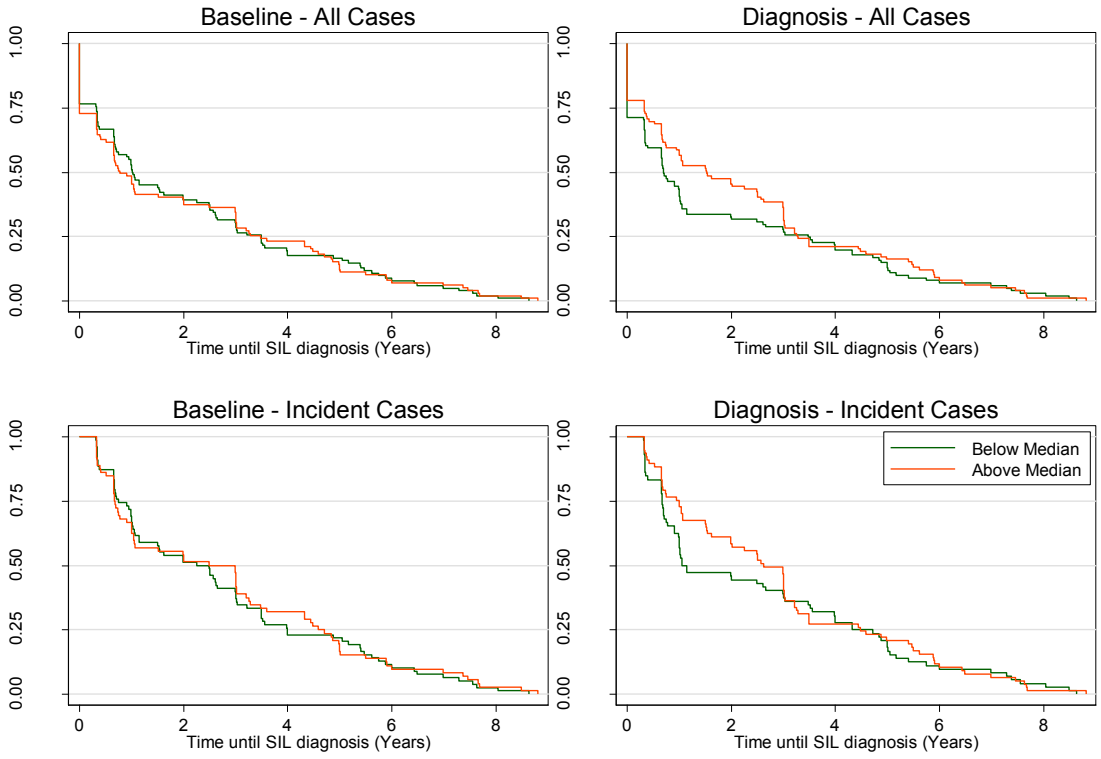

Figure 9: Kaplan-Meier analysis of time to SIL diagnosis stratified as being either above or below the median serum IGFBP-3 concentration. IGFBP-3 concentrations used in the analysis were either the woman's baseline measurement (left) or her measurement at time of diagnosis (right). The bottom graphs represent only incident SIL cases.

A Kaplan-Meier analysis of IGF-I to IGFBP-3 molar ratio and time until SIL detection was also carried out (Table 18). As was seen when stratifying by IGF-I level, when using all cases and molar ratio measured at time of diagnosis there is a significant

difference in time until SIL between those above or below the median ratio (p-value = 0.0221). This trend disappears when only including incident cases (Table 18).

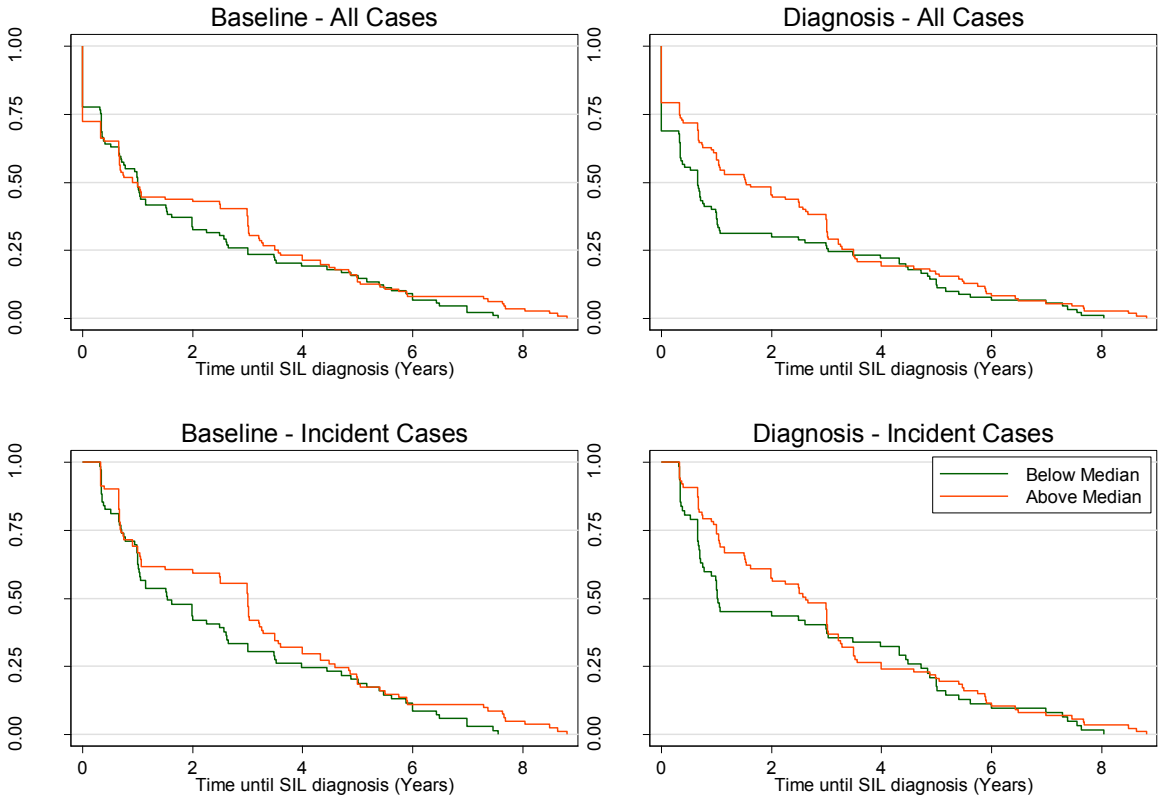

Figure 10: Kaplan-Meier analysis of time to SIL diagnosis stratified as being either above or below the median serum IGF-1:IGFBP-3 ratio. The molar ratios used in the analysis were either the woman's baseline ratio (left) or her molar ratio at time of diagnosis (right). The bottom graphs represent only incident SIL cases.

Table 18: Log-rank test results comparing observed and expected distribution of time to event between women with high versus low IGF-1, IGFBP-3, or IGF-1:IGFBP-3 molar ratio serum concentration.

Variable			P-Value
IGF-1	All Cases	IGF-1 Baseline	0.3439
		IGF-1 Diagnosis	0.0142
	Incident Cases Only	IGF-1 Baseline	0.0926
		IGF-1 Diagnosis	0.1700
IGFBP-3	All Cases	IGF-1 Baseline	0.9999
		IGF-1 Diagnosis	0.2958
	Incident Cases Only	IGF-1 Baseline	0.7567
		IGF-1 Diagnosis	0.5247
IGF-1:IGFBP-3 molar ratio	All Cases	IGF-1 Baseline	0.3066
		IGF-1 Diagnosis	0.0772
	Incident Cases Only	IGF-1 Baseline	0.1102
		IGF-1 Diagnosis	0.2583

8.0 Discussion

8.1 Determinants of IGF-I, IGFBP-3, and IGF-I/IGFBP-3 molar ratio

In an average population, IGF-I and IGFBP-3 serum concentration is normally distributed and relatively stable throughout a person's lifetime. The IGF-I and IGFBP-3 serum concentration profile for the women participating in the Ludwig-McGill Cohort Study is atypical; IGF-I concentration is highly skewed towards very low measurements, and the mean measures for the population are well below the expected average determined from other population studies (21, 23). It is therefore very important to explore possible determinants of IGF-I and IGFBP-3 in this population of women in São Paulo, and to try to elucidate the reason for such an abnormal distribution.

Between 1980 and 1993, Brazil experienced an incredibly high rate of inflation. From January to June 1993, the monthly inflation average in São Paulo fluctuated from

between 38 and 55 percent, and the total inflation for the year spanning June 1992 to June 1993 reached an incredible 6100 percent (66). After June 1993, the Brazilian government launched an anti-inflation plan called the *Plano Real*. As a result of this plan, which introduced a new currency, the *real,* the last 6 months of 1993 saw the inflation rate brought down to 1.5-2.0 percent per month, as compared to the 40 percent seen in June (66).

The reason that Brazil's economic climate in 1993 may be relevant to this study stems from the fact that the first IGF measurements were taken in this study in late 1993 and early 1994, a time directly affected by the high rates of inflation, and because circulating IGF-1 and IGFBP-3 serum levels are sensitive to calorie restriction. Turnover of epithelial cells requires energy and protein from diet and is increased by IGF-I (67). Times of protein and calorie restriction/malnutrition could lead to a down-regulation of IGF-I, and thus a decrease in cell turnover in order to conserve protein and energy. Brazil also has high social inequality by international standards. Income disparities peaked in 1989 and oscillated between 1989 and 1993, after which both income disparity and poverty began to decline (68). As the participants in this study include mostly women with a poor socioeconomic status, it is entirely possible that the abnormal fluctuation, distribution, and levels seen in the IGF levels could be attributed to fluctuation in caloric intake, due to the changes in the local economy. Another indication that changing nutritional status may be the correct cause of the fluctuations, as opposed to mechanical error, is the positive correlation between IGF-1 and IGFBP-3, and the negative correlation of serum concentrations with age. Interestingly, IGF-I and IGFBP-3 levels increased after June 1993, and decreased again in 1999 – the year that the Brazilian dollar, the Real, crashed in value (69).

Due to the temporal nature of the unique stressors experienced by our study population, it is necessary to identify determinants of IGF-I and IGFBP-3 concentrations measurements taken at baseline and at diagnosis separately. Confirmed determinants of IGF-I include age and sex (21-24), and its expression is decreased during long-term caloric and protein restriction (25, 26). It is understood that in countries such as Canada and the United States, income is not necessarily a good proxy for caloric restriction, as it is the calorie rich, nutrition poor foods that are the cheapest. However, after speaking with epidemiologists from Brazil, I believe that this was not the case for our study population. If the hypothesis that IGF-I measurements at diagnosis are closer to what is normally expected in a population is correct, then age should be the only determinant associated with IGF-I levels measured at diagnosis. This is indeed what was observed, for both IGF-I and IGFBP-3.

Potential determinants of IGF-I at baseline, other than age, were hypothesized to include measures that could act as proxies to caloric intake, such as salary. One should bear in mind that these analyses were conditional to membership in the nested case-control set. Women with the highest annual income did have greater odds of having higher IGF-I levels than those women in the lowest income quartile, and the same association was observed for women with the highest educational attainment. The association remained when the analysis was restricted to controls only, strengthening the hypothesis that it plays a role in IGF level independently of case-status. Educational attainment and salary may be linked, with women with greater education being able to earn a higher salary. When the analysis was restricted to only controls, the association between university/college education and IGF was no longer significant. This may potentially indicate that it was not a true association to begin with.

Baseline IGF-I concentration was also significantly associated with year of measurement, which was expected, and the association became stronger once the analysis was restricted to controls only. If a woman's baseline measurement was after 1995 it had a much higher odds of being above the median level than that of a women who had had their baseline visit before 1995. This can be attributed again to the economic crisis that ended late 1993, as the median IGF-I measurements increased every year until 1999 (Figure 7).

Age at menarche of 13 or older was associated with increased IGF-I levels, at both baseline and diagnosis measures, but only when the analysis contained both controls and cases. Once the analysis was restricted, this association ceased being significant, perhaps indicating that late age of menarche is somehow linked to disease status. This finding was unanticipated, as we had originally hypothesized that women with low body weights would experience menarche at a later age than women with higher body weights, and so this group of women should have lower IGF-I levels. One explanation for this trend may be that IGF-I levels peak during puberty (24), at about age 14.5 in girls, and begins to decline thereafter. Those that experience puberty later may have higher IGF-I levels than someone of the same age that experienced puberty earlier, because their levels have been declining for less time. This could also explain how age at menarche is potentially linked to disease status in this population; women who reach puberty at a later age have higher IGF-I levels throughout the time during which they are at risk of disease.

IGF-I levels were significantly lower in former smokers as compared to women who had never smoked before, for both baseline and diagnosis measurements. This association persisted in IGF-I at baseline after the analysis was restricted to controls only, but it decreased in magnitude. Observational studies are divided on whether or not

smoking is a determinant of IGF-I levels, with Juul *et al.* finding the same trend as seen in this study, but only significantly in men (24). In some experiments, smoking has been shown to increase GH release (70), and so long term smoking could lead to GH resistance, or down-regulation of the GH response. This could potentially explain why current smokers have higher IGF-I levels, as they may not have developed GH resistance yet, and those that have quit smoking may have been smokers for a much longer time before quitting. Despite this, it should be kept in mind that many observational studies have not found an association between cigarette smoking and IGF-I serum levels.

8.2 Dietary Correlates of IGF-I, IGFBP-3 and IGF-I:IGFBP-3 molar ratio

In our study, any alcohol consumption was associated with increased IGF-I and IGFBP-3 levels at diagnosis only. Goodman-Gruen *et al.* reported in their study that IGF-I levels were significantly higher in women who reported any alcohol use in the past year as compared to those who reported no alcohol use (22), a finding which is consistent with this study. The alcohol intake of the women in the Goodman-Gruen study was relatively low, and their findings were corroborated by Holmes *et al* (71), who found a non-linear association between alcohol intake and IGF-I. Holmes found that there was a somewhat positive association for low intakes of alcohol with IGF-I levels but an inverse association for the highest category of intake.

As IGF-I is produced in the liver, it was hypothesized at the beginning of this study that women who drink more would have lower IGF-I levels than those who did not. Animal studies have suggested that alcohol consumption decreases IGF-I levels (72), a finding that has been corroborated by several observational studies (73, 74). It could be that the women in the Ludwig-McGill Cohort Study were not heavy drinkers and so our

results are in line with those of Goodman-Gruen and Holmes *et al.* As to why moderate alcohol intake may increase IGF-I, a mechanism at this point is unknown, but it is thought that a decrease in protein concentration in heavy drinkers is due to liver damage, which somehow prevents hepatic synthesis of IGF-I. When the analysis was restricted to controls only, the magnitude of the association between IGF-I and any alcohol consumption increased, indicating it may be an important determinant of IGF level, independent of disease status. Alcohol intake was associated with increased IGFBP-3 levels in our study, and the magnitude of this association also increased upon restriction to controls. Holmes et al also observed this trend.

Dairy intake in the second and third tertile were associated with an increase in IGFBP-3 at baseline, and in the restricted analysis, the association remained significant in the highest tertile of dairy intake, and even increased in magnitude. Several studies have shown a relationship between higher milk intakes with an increase in serum IGF-I levels (71, 75), and one study found an increase in both IGF-I and IGFBP-3 in groups with higher milk intake (76). It is unusual that we found increased odds of higher protein levels in the IGFBP-3 group only and not in the IGF-I group at baseline as well, as the potential mechanistic reason behind dairy increasing protein levels should apply to both IGF-I and IGFBP-3. IGF-I was significantly associated with high dairy intake in the restricted analysis, but only in IGF measures taken at diagnosis. This finding may have been an artefact of having so many odds ratios being calculated, or that dairy was more easily accessible once the economy improved.

Studies have shown that it is milk in particular, and not other dairy products, such as yogurt, ice cream and cheese, that have a positive relationship with increased IGF-I and IGFBP-3 levels (76). It is possible that IGF-I in cows' milk or some other substance

in milk that is able to stimulate endogenous IGF-I/IGFBP-3 production is inactivated when milk is processed into other dairy products. Milk is rich in tropic factors such as hormones and cytokines, growth factors, and other bioactive peptides, which could potentially play a role in increasing levels of IGF-I and IGFBP-3 in the human body (76).

No relationship was seen between IGF-I or IGFBP-3 levels and β-carotene intake or vitamin C intake, which is corroborated by the findings of Holmes *et al*.(71).

When all of the potential determinants of IGF-I and IGFBP-3 were put into a backwards selection model only salary and age changed the baseline IGF-I model by more than 5%. At diagnosis, only age was an important determinant, and the same trends were seen when selecting determinants for the IGFBP-3 concentration model. When the analysis was restricted to the controls only, IGF-I at diagnosis was also only associated with age in the multivariate model. This confirms the hypothesis that age should be the only significant determinant of IGF in the population when the external pressures relating to the poor economy improved. Salary influenced IGF-I concentration at baseline for both the full and restricted analysis in the multivariate model, which also suggests that economic pressures influenced the study population's IGF-I levels, but once that pressure was alleviated, salary stopped having a significant impact.

It has to be kept in mind that a nested case control study design is not an ideal way to explore potential determinants of IGF-I/IGFBP-3 serum concentration levels, nor was the Ludwig-McGill population best suited for this undertaking. Because of these major methodological issues, I analysed the data on an exploratory basis – conditionally and with the necessary caveats recognized. The results of this analysis are not to be generalized to a larger population, such as the average woman on the streets of São Paulo. However, after conducting the sensitivity analysis, and consulting the literature, I

have confidence that the identified determinants of IGF levels, such as age, salary, year of test, alcohol and dairy intake, are real and are biologically plausible. A study design that is more appropriate for this type of analysis has to be done to confirm these results, and it would be best to conduct the analysis in a stable population that has not experienced such fluctuations in its protein concentrations.

8.3 IGF-I and IGFBP-3 serum concentration and risk of SIL

In this study, no association was seen between quartile of IGF-I or IGFBP-3 serum concentration, either at baseline or at diagnosis, and risk of cumulative SIL, which was the primary objective of this study. While the odds ratios of higher quartiles of circulating IGF-I showed a higher risk of developing SIL, as compared to the lowest quartile, none of the associations were significant. The same was found for both IGFBP-3 and the molar ratio IGF-I:IGFBP-3.

Some cancers have been linked to high serum IGF-I levels, such as prostate and colorectal cancers (27, 34), and an inverse association has been reported for gastric (77), endometrial (78, 79), liver (80), and lung cancer (81). There is considerable variation in the literature as to whether high or low IGF-I level plays a role in cervical cancer, or whether it plays any role at all.

Despite all the evidence that suggests high levels of circulating serum IGF-I should increase ones risk of developing cancer, Serrano *et al.*(82) and Schaffer *et al.* (51) found that low values of IGF-I and IGF-I:IGFBP-3 molar ratio were associated with cervical cancer/HG-CIN. Both of these studies were case-control studies that used ELISA

to measure IGF-I levels in the blood, however the Schaffer study used blood plasma where the Serrano study used the more common blood serum.

Wu *et al.* (63) reported that, in their study, higher IGF-I levels were associated with an increased risk of cervical cancer. The participants in the study by Wu *et al.* had lower mean IGF-I and IGFBP-3 levels that those found by Schaffer *et al.* and the participants were about ten years younger; they should have had higher mean IGF-I values. The fact that the participants in Wu's study were from an economically disadvantaged minority population and those from Schaffer's study were not may corroborate our findings that economic disadvantage leads to decreased IGF-I/IGFBP-3 levels. Wu also used the ELISA assay to determine serum IGF-I concentrations.

Huang *et al.* (54) reported that high-grade over-expression of IGF-IR was an independent predictor of cervical cancer death and recurrence, that preoperative serum concentration of IGF-I or IGFBP-3 levels failed to predict cervical cancer death and recurrence, and that there was a lack of correlation between circulating IGF-I and IGFBP-3, and IGF-IR over-expression in cervical cancer cells. Their findings indicate that there is a likely autocrine or paracrine IGF-IR stimulation (54). If cancerous cells can produce their own IGF-I, and stimulate IGF-IR through autocrine or paracrine mechanisms, then circulating IGF-I would not be a good marker for the level of stimulation that is occurring in the body. In such an instance, there would not be an association between the circulating levels and SIL.

Our study is at odds with the above-mentioned papers. We did not find either a positive or an inverse association with IGF-I level and risk of SIL. I speculate that potentially other papers that did agree with our findings may have been affected by publication bias, and were not published. Also, a case-control study of cervical cancer

and IGF-I levels would not be able to determine if IGF-I level was a risk factor or marker for disease, and so a prospective study was needed to answer this question. However, it is impossible to ethically carry out the study with cervical cancer as the outcome, as detection of preneoplastic lesions requires intervention. As such, our prospective study used SIL as the end point of interest, and this must be kept in mind when making comparisons to studies that used cancer as their outcome. It should also be kept in mind that the natural history of cervical cancer differs from that of other cancers, where sex hormones play a large role (breast, prostate cancers) and it was in these cancers that previous studies found an association with IGF-I concentration. Cervical cancer differs from these cancers in its development as it relies on the presence of a virus, HPV, to induce cellular changes. This may change the dynamics of neoplastic development as compared to other epithelial cancers.

8.4 IGF-I, IGFBP-3 and HPV infections

As mentioned earlier, our nested case control study was not set up with the intention of investigating a potential association between IGF-I level and risk of HPV infection, and so this study design is not an ideal one to explore this potential relationship. It was decided to do the analysis anyways, with a sensitivity analysis and with no matched sets, in order to explore the hypothesis that IGF-I levels may influence susceptibility of infection. The results have to be interpreted conditionally with many caveats.

In our study, we did not see any relationship between cumulative infection with any of the HR-HPV types and level of circulating serum IGF-I/IGFBP-3, despite

evidence in the literature indicating that this relationship may exist. For example, high molar ratios were associated with increased persistence of HR-HPV infections in one study (49).

IGF-I and IGFBP-3 proteins have been implicated in the development of certain cancers, and HR-HPV types are the carcinogenic types, so it is not surprising that no one has investigated the effect IGF-I level may have on LR-HPV infections and condylomata acuminata (genital warts caused by LR-HPV types) development. Our study did find an association between high levels of IGF-I and risk of cumulative LR-HPV infection. Condylomata acuminata are grossly visible florid lesions of the genital epithelium, and are most commonly found on the external genitalia and adjacent anal and perianal areas (83). They occur after HPV infects the genital basal epithelium, usually through a small abrasion or tissue disruption (84), and their lifecycle is tightly linked to the differentiation and proliferation of infected human keratinocytes (85). The growth and proliferation of keratinocytes could potentially be affected by IGF-I, a cellular growth factor.

Studies that investigated any potential link between IGF-I or IGFBP-3 levels and risk of HPV infection focused their analyses on infection with HR-HPV types only (51, 82, 86-88), and so there are no other studies with which to compare these findings.

8.5 IGF-I, IGFBP-3 and time until an SIL event

Among women in the cohort that did develop a SIL, whether or not they were above or below the IGF-I, IGFBP-3 or IGF-I:IGFBP-3 molar ratio median did not affect the time it took before the lesion was detected (Figure 8 - Figure 10). We had predicted that women with high levels of IGF-I would develop lesions faster than women who had

lower levels of the growth factor because of the proliferation stimulating effects of IGF-I but this was not the case. No effect may also have been seen because the IGF-I levels in this population are lower than that which is normally seen in a population, and could have been below a necessary threshold. For this reason conclusions made from this analysis may not be what would have been seen had the population not experienced CR. Further studies regarding IGF-I levels and time until lesion development should be conducted in a healthy population, with a cohort study design. As the above analysis was conducted using data from a nested case control study, and because the analysis used only cases, the survival estimates are not true estimates, and are not generalizable to any larger population.

9.0 Conclusions

This nested case control study has made important contributions to the understanding of the role serum concentrations of IGF-I, IGFBP-3, and the IGF-I:IGFBP-3 molar ratio play in the development of SIL and the acquisition of HPV infections. We did not find any significant association between IGF-I, IGFBP-3 or molar ratio level and risk of SIL, however, we did find a strong significant association between IGF-I level and risk of cumulative LR-HPV infection, and this is the first time that such an association has been documented. All previous studies on IGF-I and HPV infection included women with HR-HPV infections only, despite the biological plausibility of IGF-I influencing the development of proliferative skin lesions.

The results of this study contrast greatly with those obtained in previous studies; high levels of IGF-I have been associated with an increased risk of many cancers, including cervical cancer, and still other studies have shown an inverse association

between IGF-I and cervical cancer risk. Our study, the first nested case control study examining the relationship between this protein and risk of SIL found that there was no association at all. Unfortunately, our study population included women abnormally low serum concentrations of IGF-I and IGFBP-3, and so our results may not be generalizable to the general public until other prospective studies are carried out to corroborate our findings. Likewise, is it possible that by measuring lesion outcomes with Pap cytology, even as validated in a research cytopathology laboratory (as was done here), and not via histopathological ascertainment may have dampened the estimates of effect. However, the original intent of this cohort study was to study the natural history of cervical neoplasia without undue excisional interventions. It is interesting to note, however, how intimately caloric restriction is tied to IGF-I concentration. As there is little research on the relationship between IGF-I and IGFBP-3 and its role in cervical cancer and its precursors, and because there are contrasting results between studies in the literature, further research is needed in order to fully understand the role these factors play in the cancer's etiology. The relationship observed in this study between IGF-I and risk of cumulative LR-HPV infection needs to be further investigated, as this information could help us understand the natural history of LR-HPV infections more clearly, and to potentially be able to identify which LR-HPV infections are likely to persist.

References:

1. Ferlay J, Shin H, Bray F, et al. GLOBOCAN 2008: Cancer Incidence, Mortality and Prevalence Worldwide. Lyon, France: IARC CancerBase, 2010:http://www-dep.iarc.fr/ .
2. Parkin DM, Bray F. Chapter 2: The burden of HPV-related cancers. *Vaccine* 2006;24 Suppl 3:S3/11-25.
3. Sankaranarayanan R, Ferlay J. Worldwide burden of gynaecological cancer: the size of the problem. *Best Pract Res Clin Obstet Gynaecol* 2006;20(2):207-25.
4. Schlecht NF, Kulaga S, Robitaille J, et al. Persistent human papillomavirus infection as a predictor of cervical intraepithelial neoplasia. *JAMA* 2001;286(24):3106-14.
5. Franco EL, Rohan TE eds. Cancer Precursors: Epidemiology, Detection, and Prevention. Springer-Verlag New York Inc., 2002.
6. Schiffman MH, Brinton LA. The epidemiology of cervical carcinogenesis. *Cancer* 1995;76(10 Suppl):1888-901.
7. The 1988 Bethesda System for reporting cervical/vaginal cytologic diagnoses: developed and approved at the National Cancer Institute Workshop in Bethesda, Maryland, December 12-13, 1988. *Hum Pathol* 1990;21(7):704-8.
8. Muñoz N, Castellsagué X, de González AB, et al. Chapter 1: HPV in the etiology of human cancer. *Vaccine* 2006;24 Suppl 3:S3/1-10.
9. Trottier H, Franco EL. The epidemiology of genital human papillomavirus infection. *Vaccine* 2006;24(S1):S1/4-S1/15.
10. Giuliano AR, Papenfuss M, Abrahamsen M, et al. Differences in factors associated with oncogenic and nononcogenic human papillomavirus infection at the United States-Mexico border. *Cancer Epidemiol Biomarkers Prev* 2002;11(9):930-4.
11. Rousseau MC, Franco EL, Villa LL, et al. A cumulative case-control study of risk factor profiles for oncogenic and nononcogenic cervical human papillomavirus infections. *Cancer Epidemiol Biomarkers Prev* 2000;9(5):469-76.
12. Sellors JW, Karwalajtys TL, Kaczorowski J, et al. Incidence, clearance and predictors of human papillomavirus infection in women. *CMAJ* 2003;168(4):421-5.
13. Kjaer SK, Chackerian B, van den Brule AJ, et al. High-risk human papillomavirus is sexually transmitted: evidence from a follow-up study of virgins starting sexual activity (intercourse). *Cancer Epidemiol Biomarkers Prev* 2001;10(2):101-6.
14. Cogliano V, Baan R, Straif K, et al. Carcinogenicity of human papillomaviruses. *Lancet Oncol* 2005;6(4):204.
15. Bosch FX, Manos MM, Munoz N, et al. 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.
16. Walboomers JM, Jacobs MV, Manos MM, et al. Human papillomavirus is a necessary cause of invasive cervical cancer worldwide. *Journal of Pathology* 1999;189:12-9.
17. IARC Working Group. Human papillomaviruses. IARC Monographs on the Evaluation of Carcinogenic Risks to Humans, Vol. 90. Lyon: International Agency for Research on Cancer, 2007.
18. Macaulay VM. Insulin-like growth factors and cancer. *Br J Cancer* 1992;65(3):311-20.
19. Pollak M. Insulin-like growth factor physiology and cancer risk. *Eur J Cancer* 2000;36(10):1224-8.
20. Schaffer A, Koushik A, Trottier H, et al. Insulin-like growth factor-I and risk of high-grade cervical intraepithelial neoplasia. *Cancer Epidemiol Biomarkers Prev* 2007;16(4):716-22.
21. Kaklamani VG, Linos A, Kaklamani E, et al. Age, sex, and smoking are predictors of circulating insulin-like growth factor 1 and insulin-like growth factor-binding protein 3. *J Clin Oncol* 1999;17(3):813-7.

22. Goodman-Gruen D, Barrett-Connor E. Epidemiology of insulin-like growth factor-I in elderly men and women. The Rancho Bernardo Study. *Am J Epidemiol* 1997;145(11):970-6.
23. Landin-Wilhelmsen K, Wilhelmsen L, Lappas G, et al. Serum insulin-like growth factor I in a random population sample of men and women: relation to age, sex, smoking habits, coffee consumption and physical activity, blood pressure and concentrations of plasma lipids, fibrinogen, parathyroid hormone and osteocalcin. *Clin Endocrinol (Oxf)* 1994;41(3):351-7.
24. Juul A, Bang P, Hertel NT, et al. Serum insulin-like growth factor-I in 1030 healthy children, adolescents, and adults: relation to age, sex, stage of puberty, testicular size, and body mass index. *J Clin Endocrinol Metab* 1994;78(3):744-52.
25. Fontana L, Weiss EP, Villareal DT, et al. Long-term effects of calorie or protein restriction on serum IGF-1 and IGFBP-3 concentration in humans. *Aging Cell* 2008;7(5):681-7.
26. Smith WJ, Underwood LE, Clemmons DR. Effects of caloric or protein restriction on insulin-like growth factor-I (IGF-I) and IGF-binding proteins in children and adults. *J Clin Endocrinol Metab* 1995;80(2):443-9.
27. Chan JM, Stampfer MJ, Giovannucci E, et al. Plasma insulin-like growth factor-I and prostate cancer risk: a prospective study. *Science* 1998;279(9438850):563-6.
28. Johansson M, McKay J, Rinaldi S, et al. Genetic and plasma variation of insulin-like growth factor binding proteins in relation to prostate cancer incidence and survival. *The Prostate* 2009;69(12):1281-91.
29. Pollak M. Insulin-like growth factors and prostate cancer. *Epidemiol Rev* 2001;23(1):59-66.
30. Gunter M, Hoover D, Yu H, et al. Insulin, Insulin-Like Growth Factor-I, and Risk of Breast Cancer in Postmenopausal Women. *JNCI Journal of the National Cancer Institute* 2008;101(1):48-60.
31. Toniolo P, Bruning PF, Akhmedkhanov A, et al. Serum insulin-like growth factor-I and breast cancer. *Int J Cancer* 2000;88(5):828-32.
32. Hankinson SE, Willett WC, Colditz GA, et al. Circulating concentrations of insulin-like growth factor-I and risk of breast cancer. *Lancet* 1998;351(9113):1393-6.
33. Yu H, Spitz MR, Mistry J, et al. Plasma levels of insulin-like growth factor-I and lung cancer risk: a case-control analysis. *J Natl Cancer Inst* 1999;91(2):151-6.
34. Wolpin B, Meyerhardt J, Chan A, et al. Insulin, the Insulin-Like Growth Factor Axis, and Mortality in Patients With Nonmetastatic Colorectal Cancer. *Journal of Clinical Oncology* 2009;27(2):176-85.
35. Dal Maso L, Augustin LSA, Franceschi S, et al. Association between components of the insulin-like growth factor system and epithelial ovarian cancer risk. *Oncology* 2004;67(3-4):225-30.
36. Renehan AG, Zwahlen M, Minder C, et al. Insulin-like growth factor (IGF)-I, IGF binding protein-3, and cancer risk: systematic review and meta-regression analysis. *Lancet* 2004;363(9418):1346-53.
37. Pollak M, Schernhammer E, Hankinson S. Insulin-Like Growth Factors and Neoplasia. *Nature Reviews* 2004;4:505-18.
38. Chitnis MM, Yuen JSP, Protheroe AS, et al. The type 1 insulin-like growth factor receptor pathway. *Clin Cancer Res* 2008;14(20):6364-70.
39. Pollak M, Schernhammer ES, Hankinson S. Insulin-Like Growth Factors and Neoplasia. *Nature Reviews* 2004;4:505-18.
40. Shen M, Hsu Y, Hsu K, et al. Insulin-like growth factor 1 is a potent stimulator of cervical cancer cell invasiveness and proliferation that is modulated by alphavbeta3 integrin signaling. *Carcinogenesis* 2006;27(5):962-71.
41. Baserga R, Peruzzi F, Reiss K. The IGF-1 receptor in cancer biology. *Int J Cancer* 2003;107(6):873-7.

42. Nakamura K, Hongo A, Kodama J, et al. Down-regulation of the insulin-like growth factor I receptor by antisense RNA can reverse the transformed phenotype of human cervical cancer cell lines. *Cancer Res* 2000;60(3):760-5.
43. Karas M, Danilenko M, Fishman D, et al. Membrane-associated insulin-like growth factor-binding protein-3 inhibits insulin-like growth factor-I-induced insulin-like growth factor-I receptor signaling in ishikawa endometrial cancer cells. *J Biol Chem* 1997;272(26):16514-20.
44. Hochscheid R, Jaques G, Wegmann B. Transfection of human insulin-like growth factor-binding protein 3 gene inhibits cell growth and tumorigenicity: a cell culture model for lung cancer. *J Endocrinol* 2000;166(3):553-63.
45. Schwarze SR, DePrimo SE, Grabert LM, et al. Novel pathways associated with bypassing cellular senescence in human prostate epithelial cells. *J Biol Chem* 2002;277(17):14877-83.
46. Baege AC, Disbrow GL, Schlegel R. IGFBP-3, a marker of cellular senescence, is overexpressed in human papillomavirus-immortalized cervical cells and enhances IGF-1-induced mitogenesis. *J Virol* 2004;78(11):5720-7.
47. Mannhardt B, Weinzimer SA, Wagner M, et al. Human papillomavirus type 16 E7 oncoprotein binds and inactivates growth-inhibitory insulin-like growth factor binding protein 3. *Mol Cell Biol* 2000;20(10938125):6483-95.
48. Santer FR, Moser B, Spoden GA, et al. Human papillomavirus type 16 E7 oncoprotein inhibits apoptosis mediated by nuclear insulin-like growth factor-binding protein-3 by enhancing its ubiquitin/proteasome-dependent degradation. *Carcinogenesis* 2007;28(12):2511-20.
49. Harris T, Burk R, Yu H, et al. Insulin-Like Growth Factor Axis and Oncogenic Human Papillomavirus Natural History. *Cancer Epidemiology, Biomarkers & Prevention* 2008;17(1):245-8.
50. Harris T, Burk RD, Yu H, et al. Insulin-Like Growth Factor Axis and Oncogenic Human Papillomavirus Natural History. *Cancer Epidemiology, Biomarkers & Prevention* 2008;17(1):245-8.
51. Schaffer A, Koushik A, Trottier H, et al. Insulin-like Growth Factor-1 and Risk of High-Grade Cervical Intraepithelial Neoplasia. *Cancer Epidemiology, Biomarkers & Prevention* 2007;16(4):716-22.
52. Serrano ML, Romero A, Cendales R, et al. Serum levels of insulin-like growth factor-I and -II and insulin-like growth factor binding protein 3 in women with squamous intraepithelial lesions and cervical cancer. *Biomédica* 2006;26:256-68.
53. Wu X, Tortolero-Luna G, Zhao H, et al. Serum levels of insulin-like growth factor I and risk of squamous intraepithelial lesions of the cervix. *Clin Cancer Res* 2003;9(9):3356-61.
54. Huang Y, Shen M, Cheng Y, et al. Clinical implications of insulin-like growth factor 1 system in early-stage cervical cancer. *British journal of cancer* 2008;99:1096-102.
55. Franco E, Villa L, Rohan T, et al. 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. *Revista panamericana de salud p√∫blica = Pan American journal of public health* 1999;6(10572472):223-33.
56. Giuliano AR, Siegel EM, Roe DJ, et al. Dietary intake and risk of persistent human papillomavirus (HPV) infection: the Ludwig-McGill HPV Natural History Study. *The Journal of Infectious Diseases* 2003;188(14624376):1508-16.
57. Cancer Incidence in Five Continents, Vol. IX. In: Curado M, Edwards B, Shin H, et al., eds. *IARC Scientific Publications No 160*. Lyon: IARC, 2007.
58. Mirra A, Franco E. Cancer incidence in Sao Paulo, Brazil. In: Franco E, ed. *LICR Monograph Series in Cancer Epidemiology, Vol 1*: Ludwig Institute for Cancer Research, 1985.

59. Stafl A. Cervicography: a new method for cervical cancer detection. *Am J Obstet Gynecol* 1981;139(7):815-25.
60. Bauer HM, Ting Y, Greer CE, et al. Genital human papillomavirus infection in female university students as determined by a PCR-based method. *JAMA* 1991;265(4):472-7.
61. Hildesheim A, Schiffman MH, Gravitt PE, et al. Persistence of type-specific human papillomavirus infection among cytologically normal women. *J Infect Dis* 1994;169(2):235-40.
62. Ma J, Pollak MN, Giovannucci E, et al. Prospective study of colorectal cancer risk in men and plasma levels of insulin-like growth factor (IGF)-I and IGF-binding protein-3. *J Natl Cancer Inst* 1999;91(7):620-5.
63. Wu X, Tortolero-Luna G, Zhao H, et al. Serum levels of insulin-like growth factor I and risk of squamous intraepithelial lesions of the cervix. *Clinical cancer research* 2003;9(12960122):3356-61.
64. Munoz N, Bosch FX, de Sanjose S, et al. Epidemiologic classification of human papillomavirus types associated with cervical cancer. *N Engl J Med* 2003;348(6):518-27.
65. Kaplan EL, Meier P. Nonparametric Estimation from Incomplete Observations. *Journal of the American Statistical Association* 1958;53(282):457-81.
66. Tullio G, Ronci M. Brazilian inflation from 1980 to 1993: Causes, consequences and dynamics. *J Lat Am Stud* 1996;28:635-66.
67. Jones JI, Clemmons DR. Insulin-like growth factors and their binding proteins: biological actions. *Endocr Rev* 1995;16(1):3-34.
68. Ferreira F, Leite P, Litchfield J. The rise and fall of Brazilian Inequality: 1981–2004. *Macroeconomic Dynamics* 2008;12(S2):199-230.
69. Baer W. *The Brazilian Economy*. 6 ed. Colorado: Lynne Rienner Publishers, Inc.; 2008.
70. Attvall S, Fowelin J, Lager I, et al. Smoking induces insulin resistance--a potential link with the insulin resistance syndrome. *J Intern Med* 1993;233(4):327-32.
71. Holmes MD, Pollak M, Willet WC, et al. Dietary Correlates of Plasma Insulin-like Growth Factor I and Insulin-like Growth Factor Binding Protein 3 Concentrations. *Cancer Epidemiology, Biomarkers & Prevention* 2002;11:852-61.
72. Breese CR, Sonntag WE. Effect of ethanol on plasma and hepatic insulin-like growth factor regulation in pregnant rats. *Alcohol Clin Exp Res* 1995;19(4):867-73.
73. Lavigne JA, Wimbrow HH, Clevidence BA, et al. Effects of alcohol and menstrual cycle on insulin-like growth factor-I and insulin-like growth factor binding protein-3. *Cancer Epidemiol Biomarkers Prev* 2004;13(12):2264-7.
74. Karl JP, Alemany JA, Koenig C, et al. Diet, body composition, and physical fitness influences on IGF-I bioactivity in women. *Growth Horm IGF Res* 2009;19(6):491-6.
75. Norat T, Dossus L, Rinaldi S, et al. Diet, serum insulin-like growth factor-I and IGF-binding protein-3 in European women. *Eur J Clin Nutr* 2007;61(1):91-8.
76. Hoppe C, Molgaard C, Juul A, et al. High intakes of skimmed milk, but not meat, increase serum IGF-I and IGFBP-3 in eight-year-old boys. *Eur J Clin Nutr* 2004;58(9):1211-6.
77. Lee DY, Yang DH, Kang CW, et al. Serum insulin-like growth factors (IGFs) and IGF binding protein (IGFBP)-3 in patients with gastric cancer: IGFBP-3 protease activity induced by surgery. *J Korean Med Sci* 1997;12(1):32-9.
78. Rutanen EM, Stenman S, Blum W, et al. Relationship between carbohydrate metabolism and serum insulin-like growth factor system in postmenopausal women: comparison of endometrial cancer patients with healthy controls. *J Clin Endocrinol Metab* 1993;77(1):199-204.
79. Lacey JV, Jr., Potischman N, Madigan MP, et al. Insulin-like growth factors, insulin-like growth factor-binding proteins, and endometrial cancer in postmenopausal women: results from a U.S. case-control study. *Cancer Epidemiol Biomarkers Prev* 2004;13(4):607-12.

80. Stuver SO, Kuper H, Tzonou A, et al. Insulin-like growth factor 1 in hepatocellular carcinoma and metastatic liver cancer in men. *Int J Cancer* 2000;87(1):118-21.
81. Mazzoccoli G, Giuliani A, Bianco G, et al. Decreased serum levels of insulin-like growth factor (IGF)-I in patients with lung cancer: temporal relationship with growth hormone (GH) levels. *Anticancer Res* 1999;19(2B):1397-9.
82. Serrano M, Romero A, Cendales R, et al. Serum levels of insulin-like growth factor-I and -ll and insulin-like growth factor binding protein 3 in women with squamous intraepithelial lesions and cervical cancer. *Biomédica* 2006;26:256-68.
83. Habel LA, Van Den Eeden SK, Sherman KJ, et al. Risk factors for incident and recurrent condylomata acuminata among women. A population-based study. *Sex Transm Dis* 1998;25(6):285-92.
84. Hsueh PR. Human papillomavirus, genital warts, and vaccines. *J Microbiol Immunol Infect* 2009;42(2):101-6.
85. Zur Hausen H. *Infections causing human cancer*. Weinheim Chichester: Wiley-VCH ; John Wiley [distributor]; 2006.
86. Steller MA, Delgado CH, Bartels CJ, et al. Overexpression of the insulin-like growth factor-1 receptor and autocrine stimulation in human cervical cancer cells. *Cancer Res* 1996;56(8):1761-5.
87. Berger AJ, Baege A, Guillemette T, et al. Insulin-like growth factor-binding protein 3 expression increases during immortalization of cervical keratinocytes by human papillomavirus type 16 E6 and E7 proteins. *Am J Pathol* 2002;161(2):603-10.
88. Steller MA, Zou Z, Schiller JT, et al. Transformation by human papillomavirus 16 E6 and E7: role of the insulin-like growth factor 1 receptor. *Cancer Res* 1996;56(21):5087-91.

APPENDIX